

دانشگاه صنعتی شاهرود

دانشکده معدن و ژئوفیزیک

سمینار کارشناسی ارشد

کاربرد شبکه‌های عصبی مصنوعی در ژئوفیزیک

استاد راهنما:

دکتر علی مرادزاده

ارائه دهنده:

آرش کریمی

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تقديم به

عزیزانم

صفحه	عنوان
۱	مقدمه
۲	فصل اول - مبانی شبکه‌های عصبی مصنوعی
۲	۱-۱ مقدمه
۲	۱-۱-۱ محاسبات نرم (Soft computing)
۴	۲-۱ شبکه‌های عصبی
۸	۳-۱ شبکه‌های عصبی بیولوژیکی
۸	۱-۳-۱ فیزیولوژی نرون
۱۰	۴-۱- اختلاف‌ها و اشتراک‌های بین نرون‌ها و شبکه‌های عصبی
۱۰	۱-۴-۱- تنوع نرون‌های عصبی
۱۱	۵-۱- مفاهیم شبکه‌های عصبی مصنوعی
۱۳	۱-۵-۱- مدل پایه یک نرون
۱۵	۲-۵-۱- ساختار شبکه‌های عصبی
۱۵	۳-۵-۱- شبکه‌های پیشرو چند لایه
۱۶	۴-۵-۱- تعیین بهترین اندازه برای شبکه
۱۶	۶-۱- گونه‌های مختلف آموزش در شبکه‌های عصبی مصنوعی
۱۶	۱-۶-۱- آموزش با نظارت
۱۷	۲-۶-۱- آموزش تقویت یافته
۱۸	۳-۶-۱- آموزش بدون نظارت
۱۸	۴-۶-۱- آموزش رقابتی
۱۸	۵-۶-۱- الگوریتم‌های آموزش
۱۸	۱-۵-۶-۱- قانون دلتا
۱۸	۲-۵-۶-۱- قانون پس انتشار خطا یا کاهش گرادیان
۱۹	۳-۵-۶-۱- قانون آموزش هب
۱۹	۷-۱- ویژگی‌های شبکه عصبی
۲۰	۱-۷-۱- پارامترهای مهم شبکه‌های عصبی مصنوعی

- ۲۱-۲-۷-۱- قدرت تفکیک شبکه‌های عصبی مصنوعی
- ۲۱-۳-۷-۱- تعمیم یافتگی در مقابل حفظ
- ۲۱-۴-۷-۱- استخراج مشخصه
- ۲۲-۵-۷-۱- کنترل آموزش
- ۲۲-۶-۷-۱- مدرج کردن مشخصه‌های ورودی
- ۲۲-۸-۱- طبقه‌بندی شبکه‌های عصبی مصنوعی
- ۲۵-۱-۸-۱- مدل Mcculloch- Pitts
- ۲۵-۲-۸-۱- پرسپترون
- ۲۷-۱-۲-۸-۱- شبکه پرسپترون چند لایه
- ۲۹-۲-۲-۸-۱- الگوریتم پس انتشار خطا
- ۳۱-۳-۲-۸-۱- محدودیت الگوریتم پس انتشار خطا
- ۳۲-۴-۲-۸-۱- ملاحظات در مورد الگوریتم BP
- ۳۳-۵-۲-۸-۱- یادگیری به روش پس انتشار خطا و ارتباطش با یادگیری بیولوژیکی
- ۳۴-۶-۲-۸-۱- روش پس انتشار خطا
- ۳۷-۳-۸-۱- شبکه هاپفیلد
- ۳۸-۴-۸-۱- شبکه کوهنن
- ۳۸-۵-۸-۱- شبکه‌های عصبی مبتنی بر توابع پایه شعاعی (RBF)
- ۳۹-۶-۸-۱- مقایسه شبکه‌های عصبی پیشخور (پرسپترون) و شبکه‌های عصبی
- ۴۰-۹-۱- خلاصه
- ۵۵- فصل دوم- کاربرد شبکه‌های عصبی در ژئوفیزیک
- ۵۵-۱-۲- مقدمه
- ۵۷-۲-۲- شبکه‌های عصبی: ساختار و رفتار
- ۶۰-۲-۲- بهینه‌سازی شبکه
- ۶۰-۱-۳-۲- مسائل شناخته شده
- ۲-۳-۲- تخمین آموزش/وزن شبکه: یک مسأله بهینه‌سازی

۶۷	۳-۳-۲- عمومی سازی
۷۰	۴-۳-۲- بهینه‌سازی آرایش: پیش پردازش و هموار سازی وزنی
۷۴	۵-۳-۲- بهینه‌سازی آرایش: روشهای آسان سازی
۷۷	۶-۳-۲- بهینه‌سازی آرایش: روشهای سازنده
۸۰	۴-۲- یک استراتژی عمومی
۸۱	۵-۲- بحث و نتیجه‌گیری
۸۵	فصل سوم- کاربرد شبکه‌های عصبی در ژئوالکترونیک
۸۵	۱-۳- شبکه‌های عصبی مصنوعی برای تخمین پارامتر در ژئوفیزیک
۸۵	۱-۱-۳- خلاصه
۸۶	۲-۱-۳- مقدمه
۸۸	۳-۱-۳- شبکه عصبی پیشخور (FNN)
۸۹	۴-۱-۳- آموزش FNN
۹۱	۱-۴-۱-۳- آموزش FNN بوسیله روش پس انتشار و VFSA
۹۲	۵-۱-۳- تعبیر و تفسیر داده‌های سونداژ مقاومت ویژه
۹۳	۱-۵-۱-۳- آموزش FNN با داده‌های مصنوعی
۹۴	۲-۵-۱-۳- اثر اندازه FNN در پیش‌بینی
۹۶	۳-۵-۱-۳- آموزش FNN و پیش‌بینی با منحنی‌های VES چهار لایه‌ای
۹۷	۴-۵-۱-۳- تفسیر داده‌های صحرائی VES با FNNهای آموزش یافته
۹۹	۶-۱-۳- بحث و نتیجه‌گیری
۱۰۸	۲-۳- آشکارسازی صفحه ضخیم هادی توسط روش عنصر محدود و شبکه‌های عصبی
۱۰۸	۱-۲-۳- خلاصه
۱۰۸	۲-۲-۳- مقدمه
۱۱۰	۳-۲-۳- طراحی آرایش
۱۱۰	۴-۲-۳- نتایج مدلسازی عنصر محدود (Finite Element)
۱۱۰	۵-۲-۳- وارونه‌سازی شبکه عصبی
۱۱۱	۶-۲-۳- نتایج

- ۱۲۰-۳-۳- تفسیر ژئوفیزیکی توسط سیستم‌های عصبی مصنوعی
- ۱۲۰-۳-۳-۱- خلاصه
- ۱۲۰-۳-۳-۲- مقدمه
- ۱۲۱-۳-۳- سیستم‌های عصبی مصنوعی
- ۱۲۳-۳-۳-۴- کاربرد برای داده‌های سونداژ مقاومت ویژه
- ۱۲۴-۳-۳-۵- نتایج
- ۱۲۷-۳-۴- مثالی دیگر
- ۱۲۸- فصل چهارم- کاربرد شبکه‌های عصبی در الکترومغناطیس
- ۱۲۸-۴-۱- تشخیص الگوهای بیضی وار EM با شبکه‌های عصبی
- ۱۲۸-۴-۱-۱- خلاصه
- ۱۲۹-۴-۱-۲- اندازه‌گیری‌های صحرائی
- ۱۲۹-۴-۱-۳- برنامه آزمایشی شبکه عصبی
- ۱۳۴-۴-۱-۴- ارزیابی داده‌های الکترومغناطیس با استفاده از شبکه عصبی
- ۱۳۴-۴-۱-۴-۱- روش بررسی
- ۱۳۶-۴-۱-۴-۲- نتایج
- ۱۳۹-۴-۲- تعیین محل اهداف زیر زمین از داده‌های ژئوفیزیکی با استفاده از شبکه‌های عصبی
- ۱۳۹-۴-۱-۲- خلاصه
- ۱۳۹-۴-۲-۲- مقدمه
- ۱۴۰-۴-۲-۳- شبکه‌های عصبی
- ۱۴۱-۴-۲-۴- شرح آزمایش
- ۱۴۱-۴-۲-۴-۱- اکتساب داده‌ها
- ۱۴۳-۴-۲-۴-۲- طراحی شبکه عصبی
- ۱۴۵-۴-۲-۵- نتایج
- ۱۴۷-۴-۲-۶- خلاصه و نتایج
- ۱۵۳- فصل پنجم- کاربرد شبکه‌های عصبی در لرزه شناسی
- ۱۵۳-۵-۱- شبکه‌های عصبی برای تخمین و پیش بینی

۱۵۳	۱-۱-۵- وارونه‌سازی داده‌های لرزه‌ای توسط FNNها
۱۵۳	۱-۱-۱-۵- آزمایشات مصنوعی
۱۵۶	۲-۱-۵- بحث و نتیجه‌گیری
۱۶۱	۲-۵- تنظیم رد لرزه‌ای
۱۶۱	۱-۲-۵- کاربردهای دیگر
۱۶۴	۳-۵- کاربرد شبکه‌های عصبی برای داده‌های لرزه‌ای
۱۶۷	۴-۵- استفاده از شبکه‌های عصبی برای وارونه‌سازی داده‌های لرزه‌ای
۱۶۷	۱-۴-۵- خلاصه
۱۶۷	۲-۴-۵- مقدمه
۱۶۸	۳-۴-۵- شرح شبکه عصبی‌مان
۱۶۹	۴-۴-۵- مثالهای عددی
۱۶۹	۱-۴-۴-۵- وارونه‌سازی زمان‌های سیر
۱۷۰	۲-۴-۴-۵- مرتبط‌سازی یک لرزه‌نگار به یک پروفیل عمق - سرعت
۱۷۱	۵-۴-۵- نتایج
۱۷۶	فصل ششم- کاربرد شبکه‌های عصبی در ژئومغناطیس و گرانی‌سنجی
۱۷۶	۱-۶- تاریخچه
۱۷۸	۲-۶- تشخیص ساختار از میدانهای پتانسیلی با استفاده از شبکه‌های عصبی
۱۷۸	۱-۲-۶- خلاصه
۱۷۹	۲-۲-۶- مقدمه
۱۷۹	۳-۲-۶- تشخیص مشخصه
۱۸۰	۴-۲-۶- رده‌بندی ساختار شناسی
۱۸۲	۵-۲-۶- تشخیص خصیصه از میدان مغناطیسی هوابرد
۱۸۲	۱-۵-۲-۶- تشخیص مشخصه
۱۸۳	۲-۵-۲-۶- رده‌بندی ساختاری
۱۸۳	۶-۲-۶- نتیجه
۱۸۷	۳-۶- تحلیل مشخصه‌ها برای داده‌های میدانهای پتانسیلی با استفاده از شبکه‌های عصبی

۱۸۷	۶-۳-۱- خلاصه
۱۸۷	۶-۳-۲- آشکارسازی اولیه
۱۸۸	۶-۳-۳- شبکه‌های عصبی موازی (متقارن)
۱۹۰	۶-۳-۴- تحلیل مشخصه با استفاده از شبکه‌های عصبی
۱۹۱	۶-۳-۵- نتایج آزمایش
	۶-۴- تشخیص مشخصه‌ها برای تصاویر میدانهای پتانسیلی با استفاده از یک الگوریتم یادگیری برای ماشینهای بولتزمن
۱۹۴	۶-۴-۱- خلاصه
۱۹۴	۶-۴-۲- استخراج اطلاعات لبه‌ای
۱۹۵	۶-۴-۳- یادگیری بهینه
۱۹۶	۶-۴-۴- تشخیص مشخصه‌ها
۱۹۷	۶-۴-۵- نتایج
۱۹۹	۶-۵-۱- وارونه‌سازی داده‌های گرادئومتری گرانی‌سنجی با استفاده از یک شبکه عصبی
۱۹۹	۶-۵-۱- خلاصه
۱۹۹	۶-۵-۲- مقدمه
۱۹۹	۶-۵-۳- طراحی آزمایشی
۲۰۱	۶-۵-۴- نتایج
۲۰۴	فصل هفتم- کاربرد شبکه‌های عصبی در چاه‌پیمایی
۲۰۴	۷-۱- محاسبات عصبی در چاه‌پیمایی
۲۰۴	۷-۱-۱- نمونه آموزش شبکه
۲۰۹	۷-۲- وارونه‌سازی شبه دویعدی بلادرنگ داده‌های چاه‌پیمایی آرایش مقاومت ویژه
۲۰۹	۷-۲-۱- خلاصه
۲۰۹	۷-۲-۲- مقدمه
۲۱۲	۷-۲-۳-۱- کاهش بعدی
۲۱۵	۷-۲-۳-۲- شبیه‌سازی شبکه عصبی
۲۱۷	۷-۲-۴- مثالهای عددی

۲۱۹	۷-۲-۵- نتیجه
۲۲۰	ضمیمه A
۲۲۱	ضمیمه B
۲۳۳	فصل هشتم- قوتها و ضعفهای شبکه‌های عصبی
۲۳۵	فهرست منابع

مقدمه

شبکه‌های عصبی برای تعبیر و تفسیر مسائل در لرزه‌های چاه‌پیمایی، و کاوشهای ژئومغناطیسی، گرانی‌سنجی، مقاومت ویژه الکتریکی و الکترومغناطیسی سطحی به کار برده شده‌اند. اگرچه چاه‌پیمایی و کاوشهای هواپردی مقادیر زیادی از داده‌ها را تولید می‌کنند، اغلب تکنیکهای غیر لرزه‌های داده‌های کمتری از یک کاوش لرزه‌ای نوعی تولید می‌کنند. پیرادازش داده‌ای اندکی برای داده‌های غیر لرزه‌ای لازم است. بعد از برخی تصحیحات پایه‌ای که برای داده‌های مغناطیسی و گرانی‌سنجی بکار می‌روند، آنها شبکه‌بندی می‌شوند و به صورت خطوط میزان درآورده می‌شوند و مفسر با داده‌های کانتور شده کار می‌کند یا برخی مدلسازی‌های وارونه یا مستقیم نسبتاً ساده انجام می‌دهد. داده‌های مقاومت ویژه الکتریکی در شبه مقطع‌هایی برای تعبیر و تفسیر رسم می‌شوند و همچنین نوعاً به یک مدل ID یا 2D وارونه می‌گردند. داده‌های الکترومغناطیسی اغلب در پروفیلی برای هر فرکانس گردآوری شده (یا شبکه بندی شده و کانتور شده اگر داده‌های کافی گردآوری شده باشند) رسم می‌گردند و همچنین به یک مدل ID و 2D وارونه می‌شوند. کاربردهای شبکه عصبی که تخمین‌هایی از پارامترهای مدل زمین تولید می‌کند، مانند ضخامت و هدایت لایه، برای تولید مجموعه‌های آموزشی بر مدل‌های مستقیم تکیه دارد. از این رو، تولید مجموعه‌های آموزشی می‌تواند وقت‌گیر باشد و شمار مدل‌های آموزشی می‌تواند بزرگ باشد. برای کاربردها در جایی که اندازه مجموعه آموزش می‌تواند لازم باشد، "وارونه سازی" شبکه‌های عصبی می‌تواند به درستی و دقت وارونه سازی حداقل مربعات و اساساً سریعتر باشد. متعاقباً، شبکه‌های عصبی می‌توانند برای یادگیری جنبه مدل مستقیم مساله آموزش داده شوند و هنگامی که با وارونه سازی حداقل مربعات جفت شوند می‌توانند متعاقباً وارونه سازی خیلی سریعتری را نتیجه دهد.

امید آن می‌رود که این مجموعه بتواند علاقه دانشجویان و محققان این حیطه را برانگیزد تا با این مبحث آشنا شده و قابلیت‌های مختلف شبکه‌های عصبی و مصنوعی در این حیطه را کشف نمایند.

در خاتمه بر خود لازم می‌دانم تا از آقای دکتر مرادزاده استاد راهنمای محترم و کلیه افراد دیگری که مرا در تهیه این سمینار یاری نموده‌اند، تشکر و قدردانی نمایم.

فصل اول

مبانی شبکه‌های عصبی مصنوعی

۱-۱ مقدمه

۱-۱-۱ محاسبات نرم (Soft computing)

واژه محاسبات نرم اولین بار در سال ۱۹۹۲ توسط پروفسور زاده مطرح گردید. او محاسبات نرم را اینگونه تعریف کرد: "محاسبات نرم روش محاسبه‌ای است که به دنبال راهکارهایی برای تحقق مصنوعی قابلیت جالب توجه مغز انسان در استدلال و یادگیری در یک محیط نامعین و نادقیق می‌باشد." محاسبات نرم سه تکنیک منطق فازی، شبکه‌های عصبی، الگوریتم‌های ژنتیک و همچنین ترکیب آنها را شامل می‌شود که هر کدام ویژگیهای خاص خود را داشته و توانایی مدل کردن قسمتی از ویژگی‌های هوش انسانی را دارند.

شبکه‌های عصبی مصنوعی با ایجاد قابلیت یادگیری، آموزش، ذخیره محتوایی اطلاعات و پاسخ بلادرنگ و استفاده از تجربیات قبلی بصورت بلادرنگ بر همین اساس قابلیت برخورد با شرایط جدید که دانش قبلی نسبت به آن ندارد را می‌دهد.

در سالیان اخیر شاهد حرکتی مستمر از تحقیقات صرفاً تئوری به تحقیقات کاربردی علی‌الخصوص در پردازش اطلاعات برای مسائلی که یا برای آنها راه حلی موجود نیست و یا به راحتی

قابل حل نیستند، بوده‌ایم. با عنایت به این حقیقت علاقه فزاینده‌ای در توسعه تئوریک سیستمهای دینامیکی هوشمند مدل - آزاد که مبتنی بر داده‌های تجربی هستند، ایجاد شده است. شبکه‌های عصبی مصنوعی جزء این دسته از سیستمهای دینامیکی قرار دارند که با پردازش روی داده‌های تجربی دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کنند. بهمین خاطر به این سیستمها هوشمند گویند چرا که بر اساس محاسبات روی داده‌های عددی یا مثالها قوانین کلی را فرا می‌گیرند. این سیستمهای مبتنی بر هوش محاسباتی سعی در مدلسازی ساختار مغز بشر دارند.

شبکه‌های عصبی مدل‌های محاسباتی عظیمی، با ساختار موازی، هستند که جهت ذخیره و پردازش اطلاعات بکار می‌روند. همانطور که از اسم این شبکه‌ها فهمیده می‌شود، ساختار آنها از ساختمان عصبی و طرز کار مغز انسان الهام گرفته است. بخاطر وجود شباهتهای ساختاری بین این شبکه‌ها با ساختار عصبی مغز انسان، این شبکه‌ها همانند انسان دارای قابلیت‌ها و ظرفیتهای بی‌ظنیری در پردازش اطلاعات می‌باشند. می‌توان گفت که یادگیری از روی مثالها احتمالاً مهمترین ویژگی شبکه‌های عصبی می‌باشد. شبکه‌های عصبی توانایی آن را دارند که روابط و وابستگی‌های شدیداً غیر خطی و پیچیده را از یک حجم عظیم داده‌ها یاد بگیرند. اطلاعات و دانش آموخته شده توسط شبکه در اتصالات شبکه ذخیره شده و بازیافت این اطلاعات نیز براحتی و با ارجاع یکسری داده‌های ورودی جدید به شبکه انجام می‌شود. شبکه‌های عصبی همچنین مدل‌های محاسباتی بسیار مقاومی هستند (خرابی جزئی در بخش کوچکی از شبکه تأثیر چندانی بر عملکرد کلی شبکه ندارد)، این شبکه‌ها در برابر ورود داده‌های غلط یا ناقص نیز انعطاف‌پذیر بوده و در واقع سیستمهایی کاملاً انطباق‌پذیر با شرایط محیط می‌باشند.

این قابلیت‌های منحصربه‌فرد باعث شده است تا کاربرد شبکه‌های عصبی در گسترده وسیعی از مسایل هندسی بسیار مفید واقع شود، که منجمله می‌توان از کاربردهای شبکه‌های عصبی در زمینه پردازش صوت و تصویر، تشخیص الگو، کنترل و پیش‌بینی نام برد.

تحقیقات و تلاشهایی که در طول سه دهه گذشته انجام شده، عمدتاً بر گسترش و کاربرد روشهای عددی هر چه پیچیده‌تر متمرکز بوده است که از این روشهای عددی بطور شاخص می‌توان از روش آلمان محدود نام برد. گرچه در طول دهه‌های گذشته پیشرفتهای قابل توجهی نیز حاصل شده است، اما پر واضح است که هنوز در بسیاری موارد بین نتایج روشهای عددی و واقعیت فاصله

واقع تقلیدی است از عملکرد استدلالی و قیاسی مغز در برخورد با مسایل جدید. روش دوم بگونه‌ای پیچیده‌تر به این سیستم نزدیک می‌شود، در این روش با شبیه‌سازی از ساختمان بیولوژیکی مغز انسان شبکه‌های عصبی مصنوعی ساخته می‌شود.

شبکه‌های عصبی مصنوعی دستگاهها یا نرم‌افزارهایی هستند که بر اساس ساختمان عصبی مغز سازمان یافته‌اند و برخی رفتارها را از خود نشان می‌دهند که مشابه آن در کارکرد مغز انسان وجود دارد و یا آنکه قابل تفسیر به یکی از رفتارهای آدمی است.

نشان داده شده است که این شبکه‌ها قابلیت یادگیری، یادآوری، فراموش کردن، استنتاج، شناخت الگو، طبقه‌بندی اطلاعات و بسیاری دیگر از مهارت‌های مغز انسان را دارند. استفاده فن شناسی از این شبکه‌ها در جاهایی است که کامپیوترهای دیجیتال متداول امروزی علیرغم سرعت و دقت فوق‌العاده، هنوز قادر به رقابت با مغز انسان نیستند.

از این کاربردها می‌توان از شناخت تصاویر، شناخت صوت، خواندن دست‌نوشته‌ها، طبقه‌بندی داده‌ها نام برد. جالب است بدانیم یکی از دلایل برتری مغز انسان نسبت به کامپیوترهای فون نیومن دقت کمتر و یا به عبارت دیگر توانایی در مغز در صرف‌نظر کردن از اختلافات جزئی است. این عدم دقت باعث شده است که ما بتوانیم با درک تشابهات اشیاء یا مفاهیم، الگوهای کلی یا قوانین را از آنها استنتاج کنیم.

کامپیوترهای دیجیتال امروزی توانایی انجام میلیون‌ها عمل ضرب و تقسیم اعداد اعشاری را در کمتر از یک ثانیه دارند، یعنی بسیار قوی‌تر از آفرینندگان خود هستند. اما با این وجود در تقلید اعمال ساده مغز انسان که در کسری از ثانیه انجام می‌شود (مثل تشخیص یک چهره) نیاز به فرصت بسیار طولانی دارند. محققین علت این موضوع را در تفاوت ساختمان این دو سیستم جستجو می‌کنند.

در بررسی توانایی‌های مغز و کامپیوتر دیده می‌شود که مسایلی که توسط مغز به نحو مؤثری حل می‌شوند نوعاً دو مشخصه دارند، اولاً بخوبی قابل تعریف نیستند و ثانیاً نیازمند حجم زیادی از محاسبات هستند. آنچه باعث می‌شود که مغز بر این مشکلات قایل آید، ساختمان شبکه‌ای و امکان پردازش موازی اطلاعات در سیستم عصبی انسان می‌باشد.

شبکه عصبی انسان شامل حدود ۱۴ میلیارد سلول عصبی (نرون) است که در سازمانی بهم پیوسته، کنترل تمامی فرآیندهای حیاتی وی را در اختیار خود دارند. هر نرون از طریق گیرنده‌های

حسی خود (دندریتها) توسط تعداد زیادی از نرونهای دیگر تحریک می‌شود و در صورتی که میزان تحریک از یک حد آستانه‌ای فراتر رود، پیام عصبی در داخل شبکه عصبی پخش شده و پردازش می‌شود. در این میان هر نرون به تنهایی پردازش کوچکی بر روی اطلاعات (پیام یا تحریم عصبی) انجام می‌دهد: یا اینکه آیا تحریک به حد آستانه‌ای رسیده یا نه؟ اما بایستی توجه داشت که عمل این شبکه پیچیده و شدیداً در هم تنیده کارکردی بسیار فراتر را بوجود آورده است. البته همین جا بایستی یادآوری کرد که ساختمان واقعی مغز بسیار پیچیده‌تر از این است. هر نرون ممکن است با هزاران نرون دیگر در ارتباط باشد. شکل و رفتار نرونها در بخشهای مختلف مغز انسان متفاوت بوده و همچنین عمل هر نرون متأثر از بسیاری عوامل دیگر (نظیر شرایط محیطی، هورمونها، مواد شیمیایی، داروها و ...) است. به همین دلیل است که برآورد یا شبیه سازی حتی یک نرون نیز فوق‌العاده مشکل است. اما برای مشاهده رفتارهایی مشابه رفتارهای مغز انسان نیاز به در نظر گرفتن تمام پیچیدگیهای مغز نیست. در شبکه‌های عصبی مصنوعی تعداد نرونها، تعداد اتصالات و عملکرد هر نرون به مراتب ساده‌تر از شبکه عصبی مغز است.

شبکه‌های عصبی مصنوعی به دلیل تراکم زیاد ارتباط بین افراد محاسبه کننده ساده، قدرت و کارایی خوبی را در مقایسه با کامپیوترهای دیجیتالی فراهم می‌کنند. از این نظر، ساختمان شبکه‌های عصبی بر پایه شناخت فعلی از سیستمهای عصبی زیستی استوار است. توانایی مغز انسان در انجام پردازشهای پیچیده بر روی اطلاعات و پیدا کردن رابطه بین شناخته‌های جزئی از محیط پیرامون محرک اصلی در توسعه شبکه‌های عصبی مصنوعی یا بناهای پردازش کننده موازی بوده است.

مدلهای شبکه‌های عصبی بیشترین قدرت خود را در زمینه‌هایی نظیر شناخت صوت و تصویر نشان داده‌اند که در آنها تعداد زیادی از فرضها بطور موازی مطرح شده، نیاز به سرعت بالا بوده و بهترین سیستمهای موجود نیز بسیار نا کارتر از انسان عمل می‌کنند. بر خلاف کامپیوترهای فون نیومن که در آنها برای اجرای یک برنامه، دستور به دستور عمل می‌شود، در شبکه‌های عصبی فرضهای رقیب بصورت همزمان و به کمک سازمان موازی شبکه‌هایی که از طریق رابطه‌هایی با وزنهای متغیر بهم پیوسته‌اند، بررسی می‌شود.

با توجه به آنچه تا کنون گفته شد T.Kohonen تعریف زیر را از شبکه‌های عصبی مصنوعی

ارائه می‌کند:

«شبکه‌های عصبی مصنوعی، شبکه‌هایی کاملاً موازی و بهم پیوسته از اجزاء ساده (معمولاً تطبیق‌پذیر) همراه با سازمانهای تسلسلی‌اند که بتوانند با جهان واقعی به همان روش سیستم‌های عصبی زیستی بر هم کنش داشته باشند.»

به عبارت دیگر در این تعریف فرض بر این است که عملکرد شبکه‌های عصبی زیستی، حتی پیچیده‌ترین رفتارهای بشری نظیر ارتباط گفتاری، بایستی نتیجه دریافت تکانه‌های ورودی و باز ساخت پاسخهای خروجی باشد. هر چند که چنین تبدیلی نمی‌تواند همچون یک نگاشت ماتریسی بین انگیزش ورودی و بازتاب خروجی، ایستا و بی‌زمان باشد اما مطمئناً کاربردهای فنی از شبکه‌های عصبی مصنوعی بازساخت عملکردهای شناخته شده مغز انسان نیست بلکه این کاربردها معطوف به مسایلی است که با ساختمان موازی این شبکه‌ها سازگاری داشته باشند.

فرایند بالقوه شبکه‌های عصبی تنها به محاسبات سریع از طریق پردازش موازی محدود نمی‌گردد. شبکه‌های عصبی در مقایسه با کامپیوترهای متداول امروزی دارای تحمیل‌پذیری بیشتری در مقابل خطاهای داخلی می‌باشند، و از این لحاظ مقاوم‌ترند. علت این امر این است که این شبکه‌ها دارای تعدادی زیادی عناصر پردازشگر هستند که این عناصر هر کدام ارتباط محلی دارند. بدین ترتیب خرابی در چند نرون و یا چند مسیر ارتباطی، لزوماً کاهش چشمگیری در عملکرد کل شبکه بوجود نخواهد آورد. همچنین بسیاری از قوانین آموزش شبکه‌های عصبی، فرآیند آموزش و تغییر وزن شبکه را همزمان اجرا می‌کنند و کارایی شبکه را با توجه به آخرین اطلاعات موجود تنظیم می‌کنند. این قابلیت در بسیاری از کاربردهای شبکه‌های عصبی مانند کنترل و پردازش صوت و تصویر مساله‌ای حیاتی می‌باشد و خود به پایداری و مقاوم بودن شبکه کمک می‌کند، زیرا تغییرات جزئی خواص و مشخصات نرونهای شبکه را جبران می‌کند.

تاریخچه شبکه‌های عصبی به سال ۱۹۴۳ و مقاله Mc culloch و Pitts بر می‌گردد که در آن مدل اولیه نرون معرفی شده بود. در طی ۱۵ سال بعد مطالعات و تحقیقات وسیعی بر روی جزئیات مدل نرون و ارتباط شبکه‌های عصبی با برخی از سیستمهای منطقی مانند ماشینهای با حرارت محدود، انجام گرفت. در سال ۱۹۶۰ گروه Rosenblatt فعالیت‌هایی انجام دادند که هدف آن تعیین وزنهای مناسب شبکه برای انجام هر کار خاص بود. این گروه بیشتر بر روی نوعی از شبکه‌های عصبی به نام پرسپترون یک لایه متمرکز شده بودند. مشابه این کار نیز بر روی یک شبکه دیگر به نام آدلاین توسط Hoff و Widrow انجام گرفت. در روشهایی که توسط این افراد

معرفی شد و زندهای شبکه در طی یک فرآیند تکراری تغییر کرده و چنان تنظیم می‌گردد که خطای خروجی شبکه به حداقل برسد. نتایج این بررسی امیدواریهای فراوانی را در زمینه کاربرد شبکه‌های عصبی بدنبال آورد.

در سال ۱۹۶۹ بررسی‌های Papert و Minsky نشان داد که پرسپترون یک لایه تنها قادر به حل گروه کوچکی از مسایل می‌باشد. مثال معروفی که در این رابطه برای عدم کارایی این نوع شبکه ارایه شد حل مساله XOR می‌باشد. این موضوع باعث رکود و کندی تحقیقات مربوط به شبکه‌های عصبی برای حدود ۲۰ سال شد. در این دوره البته تحقیقاتی در مورد سایر انواع شبکه‌های عصبی توسط افرادی چون Kohonen و Grossberg انجام شد که نزدیکی بسیار زیادی به بررسی‌های انجام شده بر روی ساختار سیستم عصبی جانداران داشت. در دهه ۱۹۸۰ با ابداع روش پس انتشار خطا برای آموزش شبکه‌های پرسپترون چند لایه، بار دیگر اهمیت و کاربرد شبکه‌های عصبی مد نظر قرار گرفت. ولی کارهای اساسی بطور مستقل در سال ۱۹۸۵ توسط گروه Rumelhart تدوین و تکمیل گردید. این روش توانایی‌های بسیاری به شبکه پرسپترون چند لایه بخشیده است.

۱-۳ شبکه‌های عصبی بیولوژیکی

۱-۳-۱ فیزیولوژی نرون

واحد پایه و اساسی سیستم عصبی، خصوصاً مغز، نرون (سلول عصبی) می‌باشد. با وجود اندازه میکروسکوپی نرون، این واحد کوچک یک کارخانه شگفت‌انگیز پردازش اطلاعات الکتریکی و بیوشیمیایی می‌باشد. از دیدگاه کلاسیک، نرون یک واحد پردازش ساده است که سیگنالهایی را از نرونهای مجاورش و از طریق یکسری مسیرهای رشته‌ای دریافت و با هم ترکیب می‌کند. به این مسیرهای ورودی رشته‌ای دندریت می‌گویند (شکل ۱-۱).

شاخه‌های دندریتی به بدنه اصلی سلول عصبی (سوما) متصل می‌باشند. سوما هر می شکل یا استوانه‌ای شکل می‌باشد. مرز خارجی سلول عصبی غشاء نامیده می‌شود. درون سلول عصبی توسط سیال درون سلولی و خارج از سلول توسط سیال برون سلولی پر شده است. غشای نرون و مواد موجود در داخل و بیرون نرون در کارکرد و بقا (زیست) نرون نقش مهمی بازی می‌کنند. وقتی

که نرون تا یک حد معین، که به آستانه تحریک موسوم است، تحریک شود، نرون شلیک می‌کند. بدین معنا که یک سیگنال الکتریکی، که به پتانسیل فعالیت موسوم است، در امتداد یک مسیر، به نام اکسون، انتقال می‌یابد. به محل اتصال اکسون و سوما، برآمدگی اکسون می‌گویند. سمت دیگر اکسون شاخه‌شاخه شده و این مسیرهای رشته‌ای به دندریتهای نرونهای دیگر وصل می‌شوند. محل اتصال اکسون یک نرون به دندریت نرونی دیگر سیناپس نامیده می‌شود. بخشهای مختلف یک سیناپس عبارتند از: ترمینال پیش سیناپتیکی، اتصال سیناپتیکی، و ترمینال پس سیناپتیکی، که این بخشهای مختلف را می‌توان در شکل ۱-۲ مشاهده نمود

یک نرون به تنهایی ممکن است در حدود ۱۰۰۰ تا ۱۰۰۰۰ سیناپس داشته باشد و با حدود ۱۰۰۰ نرون در تماس باشد. اما بایستی توجه داشت که تمامی سیناپسها همزمان با هم تحریک نمی‌شوند بخاطر اینکه با رسیدن یک الگوی حسی به یک نرون، احتمالاً تنها درصد نسبتاً کمی از کل سیناپسهای نرون تحریک می‌شوند، لذا می‌توان نتیجه گرفت که تقریباً تعداد نامحدودی از این الگوها می‌توانند به یک نرون برسند (وارد شوند) بدون اینکه ظرفیت نرون مذکور اشباع شود. وقتی که پتانسیل فعالیت به یک انتهای اکسونی (انتهای یک رشته از اکسون) می‌رسد پیکهای شیمیایی آزاد می‌شوند. این پیکها در ساختار کروی بسیار کوچکی (شکل ۱-۲) ذخیره شده (قرار گرفته) و مسئولیت ارتباط اطلاعاتی مؤثر بین نرونها بر دوش آنها می‌باشد.

وقتی که یک پیک آزاد می‌شود به آهستگی در عرض اتصال سیناپتیکی حرکت کرده و شروع به بر هم زدن تعادل یونی غشا پس سیناپسیکی می‌کند، به عبارت دیگر توزیع یونی در سطح غشا تغییر کرده و بنابراین ولتاژ در عرض غشا نرون دریافت کننده تغییر می‌کند (در دو طرف اتصال سیناپتیکی اختلاف پتانسیل الکتریکی بوجود می‌آید و یک جریان یونی در عرض اتصال بوجود می‌آید) هر چه اتصال سیناپتیکی قوی‌تر باشد، تعداد بیشتری پیک شیمیایی به غشای پس سیناپتیکی می‌رسند. خود پیکهای شیمیایی انواع مختلفی دارند که بسته به نوع پیکهای آزاد شده، پتانسیل پس سیناپتیکی می‌تواند مثبت (برانگیزنده) یا منفی (بازدارنده) باشد.

جمع شدن تحریکات عصبی در سیناپس به دو صورت برآیند زمانی و برآیند مکانی می‌باشد. (شکل ۱-۳) در جمع‌زنی زمانی پتانسیل هر تحریک به مجموع پتانسیل‌های تحریکهای قبلی افزوده می‌شود. در جمع‌زنی مکانی مجموع پتانسیل‌تهای برانگیزنده یا بازدارنده که توسط تمامی نرونهای دیگر به نرون هدف ارسال شده‌اند، محاسبه می‌شود. شارژ پتانسیل کلی عبارتست از

برآیند مجموعهای زمانی و مکانی تحریکات رسیده به سیناپس‌های یک نرون. این برآیند در یک فاصله زمانی کوتاه و همزمان با ذخیره شدن این شارژها در غشای سلول، بوجود می‌آید. غشای سلولی در وهله اول همانند یک خازن عمل کرده و در مرحله بعد، وقتی که یکسری مکانیزمهای بیوشیمیایی پیچیده اتفاق می‌افتد، همانند یک فرستنده ثانویه داخلی عمل می‌کنند.

تمامی سیگنال‌های رسیده از غشای سلول در سوما با هم ترکیب می‌شوند، اگر دامنه سیگنال برآیند به حد آستانه‌ای نرون برسد، نرون به اصطلاح شلیک می‌کند (آتش می‌کند) و یک سیگنال خروجی تولید می‌شود. این سیگنال، که ممکن است تنها یک پالس یا توالی از پالسها با یک نرخ ثابت باشد، در امتداد اکسون نرون به رشته‌های اکسون منتقل می‌شود.

۱-۴-۱- اختلاف‌ها و اشتراکهای بین نرون‌ها و شبکه‌های عصبی

۱-۴-۱-۱- تنوع نرونهای عصبی

در دنیای واقعی شبکه‌های عصبی، تمامی نرونها دارای وظیفه یکسانی نبوده یا اینکه یک جور کار نمی‌کنند، این تنوع بر پیچیدگی شبکه‌های عصبی می‌افزاید. با وجود این تنوع، تمامی شبکه‌های عصبی ویژگی‌های مشخص زیر را به نمایش می‌گذارند.

- اتصالات موازی زیادی بین بسیاری از نرونها وجود دارد.

- تعداد زیادی از این اتصالات موازی مکانیزمهای پس‌خور را برای نرونها دیگر و یا خودشان فراهم می‌آورند.

- بعضی نرونها ممکن است باعث تحریک نرونهای دیگر شوند در حالیکه در همان حال موجب از کار انداختن فعالیت نرونهای دیگر می‌شوند.

- در بخشهایی از شبکه، اتصالات نرونها به همدیگر از پیش موجود است، در حالیکه در بخشهای دیگر این اتصالات ممکن است در حال شکل‌گیری یا توسعه باشند.

- خروجی شبکه لزوماً «بله - خیر» یا ۰ و ۱ نمی‌باشد.

- شبکه‌های عصبی کارکردی غیر همزمان دارند.

- شبکه‌های عصبی یک برنامه تسلسلی (قدم به قدم) را اجرا نمی‌کنند بلکه این برنامه کاملاً در کل شبکه پخش شده است.

- شبکه عصبی دارای یک پردازشگر مرکزی نمی‌باشد، بلکه پردازش نیز در کل شبکه و در واحدهای ساده پردازشگر انجام می‌شود.

۱-۵- مفاهیم شبکه‌های عصبی مصنوعی

در این بخش اصول اساسی را که توسط آنها می‌توان شبکه‌های عصبی را بصورت ریاضی توصیف کرد بیان می‌کنیم. این اصول از جهان بیولوژیکی سرچشمه می‌گیرند و با استفاده از ریاضیات سعی می‌کنیم تا رفتار بیولوژیکی نرونها و شبکه‌های آنها را حتی المقدور توصیف کنیم.

مشخصات اساسی یک شبکه عصبی مصنوعی به چهار بخش تقسیم می‌گردد:

الف- ساختار شبکه

ب- توابع عملکرد

ج- الگوریتم آموزش شبکه

د- شیوه بکارگیری شبکه‌های عصبی مصنوعی

در برخی متون (ب)، (ج) و (د) را تحت یک عنوان به نام ویژگیهای عملیاتی یا دینامیک نرونها بیان می‌کنند.

- ساختار شبکه معین می‌کند که شبکه از چند نرون تشکیل یافته و چگونه این نرونها در شبکه آرایش یافته و به چه شکل به همدیگر اتصال یافته‌اند. هر نرون یا عنصر پردازشگر مشابه با نرون طبیعی دارای تعدادی ورودی، قدرت سیناپس، عملکرد، تعدادی خروجی و بایاس می‌باشد.

تابع عملکرد خروجی یک نرون را به ازای یک ورودی معین مشخص می‌کند.

الگوریتم آموزش نحوه آموزش شبکه را برای یک سری الگوهای آموزش معین نشان می‌دهد.

شیوه بکارگیری شبکه عصبی مصنوعی نشان می‌دهد که چگونه می‌توان پاسخ شبکه یا خروجی آن را به ازای یک الگوی ورودی مشخص کرد.

ویژگی‌های اساسی شبکه‌های عصبی را می‌توان به معماری شبکه و خصوصیات کاری یا

نرودینامیک شبکه تقسیم کرد. معماری شبکه مشخص کننده ساختار شبکه یعنی تعداد نرونهای

مصنوعی شبکه و نحوه اتصال آنها می‌باشد. شبکه‌های عصبی از تعداد زیادی نرونهای متصل بهم یا واحدهای پردازشگر تشکیل شده‌اند. نرودینامیک شبکه‌های عصبی مشخص کننده ویژگی‌های شبکه

می‌باشد. منظور از ویژگی‌های شبکه یعنی اینکه چگونه شبکه عصبی یاد می‌گیرد (آموزش می‌بیند)، فرا خوانده می‌شود، تخصیص می‌یابد و به طور پیوسته اطلاعات جدید را با دانش موجود مقایسه می‌کند، چگونه شبکه اطلاعات جدید را دسته‌بندی کرده، و چگونه در صورت نیاز شبکه دسته‌بندی جدیدی را بوجود می‌آورد.

شبکه‌های عصبی اطلاعات را پردازش می‌کنند، ولی نه توسط یک الگوریتم گام به گام. این پردازش بر مبنای تجزیه (تلاش) موازی اطلاعات به درون واحدهای پایه پردازشگر استوار است. همانطور که رنگ ترکیبی را می‌توان به یک سری طول موجها و دامنه‌های پایه تجزیه کرد و سپس از نظر تئوری می‌توان مجدداً یک رنگ خاص را تولید کرد. بطور مشابهی یک شبکه عصبی اطلاعات پیچیده را به واحدهای پایه‌اش تجزیه می‌کند، این واحدها و رابطه بین آنها در حافظه مغز مانند شبکه ذخیره می‌شود. برای مثال وقتی که شما به یک تصویر نگاه می‌کنید، مغز شما و حافظه‌اش کاری همانند یک دوربین فیلمبرداری ویدئویی انجام نمی‌دهد، یعنی مشابه دوربین فیلمبرداری یک ماتریس از نقطه‌ها را ذخیره نمی‌کند، بلکه در عوض المانهای پایه تصویر مثل خطوط، نقطه‌های برجسته، اشکال، رنگها و موقعیت فضایی آنها نسبت به هم را ذخیره می‌کند.

شبکه‌های عصبی مصنوعی با وجود اینکه با سیستم عصبی طبیعی قابل مقایسه نیستند، ویژگی‌هایی دارند که به شرح زیر است:

- **قابلیت یادگیری:** یعنی توانایی تنظیم پارامترهای شبکه (وزن‌های سیناپتیکی) در مسیر زمان که محیط شبکه شرایط جدید را تجربه کند، با این هدف که اگر شبکه برای یک وضعیت خاص آموزش دید و تغییر کوچکی در شرایط محیطی شبکه (وضعیت خاص) رخ داد، شبکه بتواند با آموزش مختصر برای شرایط جدید نیز کارآمد باشد. دیگر اینکه اطلاعات در شبکه‌های عصبی در سیناپسها ذخیره و هر نرون در شبکه، بصورت بالقوه از کل فعالیت سایر نرونها متاثر می‌شود. در نتیجه، اطلاعات از نوع مجزا از هم نبوده، بلکه متاثر از کل شبکه می‌باشد.

- **پراکندگی اطلاعات** "پردازش اطلاعات بصورت متن": رابطه یک به یک بین ورودیها و وزنهای سیناپتیکی وجود ندارد. هر نرون در شبکه از کل فعالیت سایر نرونها متاثر می‌باشد. بر این اساس چنانچه بخشی از سلولهای شبکه حذف شوند و یا عملکرد غلط داشته باشند باز هم احتمال رسیدن به پاسخ صحیح وجود دارد.

- قابلیت تعمیم: شبکه تابع را یاد می‌گیرد، الگوریتم را می‌آموزد و یا رابطه تحلیلی مناسبی را برای تعداد نقاط در فضا بدست می‌آورد.

- پردازش موازی: سلولهایی که در یک تراز قرار می‌گیرند، می‌توانند بطور همزمان به ورودیهای آن تراز پاسخ دهند.

- مقاوم بودن: سلولها در یک روند همکاری خطاهای محلی یکدیگر را تصحیح می‌کنند.

۱-۵-۱- مدل پایه یک نرون

اکنون نظر خود را به واحد پایه یا بلوک ساختمانی شبکه‌های عصبی مصنوعی یعنی نرون (یا المان پردازشگر) معطوف می‌کنیم. منظور ما در اینجا نرون مصنوعی می‌باشد و بایستی بدانیم که این نرون با نرون بیولوژیکی واقعی فاصله بسیار زیادی دارد (مدل بسیار ساده شده‌ای از نرون واقعی می‌باشد). یک نرون در حالت کلی یکسری ورودی X_j دارد، که عدد j از ۱ تا n تغییر می‌کند و نشان دهنده منبع سیگنال ورودی می‌باشد. هر ورودی X_j قبل از رسیدن به بدنه اصلی واحد پردازشگر توسط مقاومت اتصال یا فاکتور وزنی X_j وزن دار می‌شود. (برای مثال X_j در W_j ضرب می‌شود) بعلاوه هر نرون دارای یک بایاس W_0 می‌باشد. همچنین یک مقدار آستانه‌ای برای هر نرون وجود دارد که در صورتی که فعالیت نرون به آن حد برسد یا از آن تجاوز کند، نرون شلیک می‌کند (یک سیگنال تولید می‌کند). یک تابع غیر خطی F نیز وجود دارد که بر روی سیگنال تولید شده (یا فعالیت) R ، عمل می‌کند و یک خروجی به نام O در اثر عملکرد تابع F بر روی R تولید می‌شود. O از طریق نرون مزبور به سمت نرونهای دیگر ارسال می‌شود و ورودی نرونهای بعدی محسوب می‌شود. وقتی که نرون جزئی از یک شبکه متشکل از نرونها بسیار باشد به هر نرون اصطلاحاً گره گفته می‌شود. در حالتی که m نرون در یک شبکه وجود داشته باشند، یک زیرنویس اضافی، i به منظور تشخیص هر نرون منفرد مورد نیاز می‌باشد. ورودی‌ها، وزنها، سیگنالهای فعالیت، خروجی، آستانه و تابع غیر خطی به ترتیب با علائم X_{ij} , W_{ij} , R_{ij} و F_i نشان داده می‌شوند. مدل پایه یک نرون در شکل ۱-۴ نمایش داده شده است. در شکل ۱-۵ مدل مصنوعی نرون با مدل بیولوژیکی (نرون بیولوژیکی) مقایسه شده است.

و نرون در صورتی شلیک می‌کند که رابطه زیر برقرار باشد.

$$\sum_{j=1}^n w_{ij} x_{ij} \geq \theta_i$$

تابع عملکرد این مدل پایه بصورت زیر بیان می‌شود:

$$y_i = F(\sum x_{ij} w_{ij})$$

$$Y_{m_i} = \sum x_{ij} w_{ij} = x_{0_i} w_{0_i} + \sum x_{ij} w_{ij}$$

$$= (1)b_i + \sum x_{ij} w_{ij}$$

که در آن i نشان‌دهنده شماره نرون می‌باشد و j نمایش‌دهنده ورودی‌های مختلف از نرونهای دیگر می‌باشد. شکل ۱-۶ طرز کار یک نرون را توضیح می‌دهد.

هدف از بکار بردن تابع غیر خطی این است که مطمئن شویم واکنش نرونها در برابر تحریکات مختلف (مثلاً تحریکات بیش از اندازه بزرگ یا بیش از اندازه کوچک) حد و مرزی دارد، به عبارت دیگر قابل کنترل است. این موضوع با دنیای بیولوژیکی نیز انطباق دارد. مثلاً به خوبی برای ما مشخص شده است که برای اینکه یک صدا را دو برابر بلندتر احساس کنیم، لازمست تا در واقعیت دامنه صدا در حدود ۱۰ برابر بزرگتر شود. به عبارت دیگر رابطه تقریباً لگاریتمی بین بلندی واقعی صدا و آنچه که ما احساس می‌کنیم برقرار است. نرونهای بیولوژیکی به طریق مشابهی شدت واکنش خود را در برابر تحریکات محدود می‌کنند. اما تابع غیر خطی که در اکثر نمونه‌های شبکه‌های عصبی مورد استفاده قرار می‌گیرد لزوماً یک تقلید محض از نسخه‌های بیولوژیکی نیست، بلکه اغلب این تابع فقط برای راحتی محاسبات و مدلسازی ریاضی مورد استفاده قرار می‌گیرد به همین خاطر انواع مختلفی از توابع غیر خطی استفاده می‌شوند. که نوع توابع غیر خطی استفاده شده بستگی به نوع شبکه عصبی و الگوریتم آموزش آن دارد. تعدادی از توابع غیر خطی که معمولاً مورد استفاده قرار می‌گیرند در شکل (۱-۷) مشاهده می‌شوند. علاوه بر این، توابع غیر خطی دیگری نیز مثل تابع \tanh ، مورد استفاده قرار می‌گیرند. محبوبترین توابع غیر خطی مورد استفاده تابع پله‌ای و تابع سیگموئید می‌باشند. توابع غیر خطی موجود در شکل ۱-۷ همگی در محدوده خاصی قرار دارند (دارای حد و مرز می‌باشند) این توابع دارای یک حد بالا یا یک حد پایین و یا هر دو حد می‌باشند. تابع سیگموئید محبوبترین تابع مورد استفاده می‌باشد چرا که این تابع یکنواخت است، دارای حد و مرز بوده و مشتق آن ساده است و در عین حال یک تابع غیرخطی نیز می‌باشد.

۱-۵-۲ - ساختار شبکه‌های عصبی

نرونها بصورت طبیعی به روش خاصی بهم اتصال می‌یابند تا یک شبکه عصبی را تشکیل دهند. این نحوه قرار گرفتن نرونها می‌تواند به گونه‌ای باشد که شبکه تک لایه یا چند لایه را بوجود آورد. تعداد دیواره‌های وزنه‌ای حائل بین دو لایه از نرونها مشخص کننده تعداد لایه‌هاست. در بسیاری از مدل‌های شبکه‌های عصبی از قبیل پرسپترون، تناظر خطی و مدل آموزش تشدید سازگار خروجی واحدهای یک لایه فقط اجازه تحریک نرونهای لایه بعد را دارد. اما در برخی مدل‌ها مانند شبکه هابفیلد خروجی هر واحد می‌تواند واحدهای همان لایه را تحریک نماید. قدرت محاسبات عصبی از تعداد نرونهای اتصال یافته در ساختار شبکه نشأت می‌گیرد. معمولاً شبکه‌های بزرگتر، توانایی محاسباتی بیشتری را عرضه می‌کنند. مرتب نمودن نرونها در لایه‌های یا طبقات مختلف تقلیدی از ساختار بخش معینی از مغز می‌باشد و شده است که تواناییهای شبکه‌های چند لایه از شبکه‌های تک لایه بیشتر است. متداولترین ساختار عصبی مورد استفاده همان شبکه‌های عصبی چند لایه با الگوریتم آموزش از نوع پس انتشار خطا می‌باشد.

در شکل ۱-۸ ساختار تعدادی از شبکه‌های عصبی نشان داده شده است که دایره‌ها بیانگر نرونها و خطوط فلش‌دار نشان‌دهنده وزنه‌های سیناپس می‌باشند. در اینجا برای نمونه شبکه پیشخورد (پیشرو) چند لایه را توضیح می‌دهیم.

۱-۵-۳ - شبکه‌های پیشرو چند لایه

در اینجا نوع دیگری از شبکه‌های پیشخورد که به شبکه پیشخورد چند لایه موسوم است را معرفی می‌نماییم. فرق این شبکه با شبکه تک لایه در این است که مابین لایه ورودی و لایه خروجی یک یا چند لایه به نام لایه‌های پنهان وجود دارد. وظیفه این لایه‌ها ارتباط دادن لایه ورودی با لایه خروجی می‌باشد. شبکه با داشتن این لایه پنهان قادر می‌گردد، روابط غیرخطی را از داده‌های عرضه شده به آن استخراج نماید.

شکل (۱-۹) نمایش یک شبکه پیشخورد دو لایه است.

وزنه‌ها در شبکه‌های عصبی مصنوعی مقادیری عددی به خود می‌گیرند و اطلاعات شبکه در همین وزنه‌ها ذخیره می‌گردد. هر شبکه عصبی باید مکانیزمی برای آموزش یا تعلیم داشته باشد

که بوسیله آن بتواند مجموعه بردارهای آموزشی ورودی را یاد بگیرد، یعنی بوسیله آنها مقدار وزنه‌های خود را تنظیم نماید. بردارهای آموزش ورودی باید حاوی اطلاعات مناسب راجع به مسأله مورد تعلیم را دارا باشند. یک شبکه عصبی به وسیله ساختار، الگوریتم آموزشی و الگوریتم پردازش خود مشخص می‌شود.

۱-۵-۴- تعیین بهترین اندازه برای شبکه

درجه آزادی در یک شبکه عصبی مصنوعی عبارت است از تعداد اتصالات داخلی شبکه که با تعداد نرونهای لایه پنهان در ارتباط مستقیم است. در حال حاضر روشی برای تعیین تعداد لایه‌های پنهان وجود ندارد و تنها روشی که پیشنهاد می‌گردد همان روش سعی و خطاست. با افزایش تعداد نرونهای لایه پنهان از یک مقدار کوچک به مقادیر بزرگ، ابتدا مجموع خطاها کاهش می‌یابد اما پس از رسیدن مقدار نرونها به اندازه معین این خطا شروع به افزایش می‌نماید. بدین روش می‌توان بهترین اندازه برای شبکه را انتخاب نمود.

۱-۶- گونه‌های مختلف آموزش در شبکه‌های عصبی مصنوعی

آموزش شبکه عصبی به این شکل تعریف می‌گردد که آموزش فرآیندی است که بوسیله آن شبکه عصبی خود را برای یک محرک سازگار می‌کند بنحوی که بعد از تعدیل مناسب پارامترهای شبکه پاسخ مطلوب را ارائه دهد. در خلال آموزش، شبکه پارامترهای خود یعنی وزنه‌های سیناپس را در پاسخ به ورودی محرک تعدیل می‌کند. بنحوی که خروجی واقعی شبکه به خروجی مطلوب همگرا شود. همانطور که یک شیوه آموزش خاص برای افراد معین مناسب است، برای هر شبکه عصبی مصنوعی نیز روشهای خاصی مناسب است. در ذیل گونه‌های مختلف آموزش آمده است.

۱-۶-۱- آموزش با نظارت

در این شیوه آموزش ورودیهای شبکه و خروجیهای متناظر با آنها از قبل مشخص است. حال در زمان آموزش شبکه یک ورودی را به شبکه اعمال می‌کنیم. شبکه در پاسخ به آن ورودی محرک پاسخ خروجی را نتیجه می‌دهد. این خروجی با خروجی مطلوب متناظر با همین ورودی

مقایسه می‌شود. حال اگر خروجی واقعی با خروجی مطلوب مغایرت داشته باشد، شبکه یک سیگنال خطا تولید می‌کند که از این سیگنال فقط برای محاسبه میزان تغییری که باید بر وزنه‌های سیناپس اعمال شود، استفاده می‌گردد. این روند آنقدر انجام می‌گردد تا خروجی واقع همان خروجی مطلوب یا نزدیک به آن شود. این روند حداقل نمودن خطا نیازمند به یک مدار ویژه به نام معلم یا ناظر دارد که با توجه به این میزان سیگنال خطا تعدیل نماید تا اینکه آموزش لازم را ببیند. به همین دلیل به این شیوه آموزش، آموزش با نظارت گفته می‌شود.

توجه به این نکته مهم است که در خلال آموزش وزنه‌ها به نحوی تعدیل می‌یابد که خطا به حداقل برسد. در زمان آموزش ممکن است به مقادیری برای وزنه‌ها برسیم که ظاهراً حداقل خطا را در خروجی ایجاد می‌نماید اما اگر روند آموزش را ادامه دهیم ابتدا خطا افزایش یافته و از نو به یک حداقل دیگر می‌رسد که از حداقل اولی کمتر است. به آن حداقل اول، حداقل موضعی می‌گوییم. حال اگر روند آموزش را ادامه دهیم و میزان خطا هرگز از میزان خطای دومی کمتر نشود به این حداقل خطا، حداقل مطلق می‌گوییم. این مطلب در شکل (۱-۱۰) نشان داده شده است. تمام الگوریتمهای آموزشی سعی‌شان بر این است که به این حداقل برسند.

۱-۶-۲- آموزش تقویت یافته

در این شیوه نیز مانند روش آموزش با واسطه نیاز به یک معلم می‌باشد، اما در این روش معلم تعیین نمی‌کند که خروجی واقعی چه اندازه با خروجی مطلوب تفاوت دارد، بلکه مشخص می‌کند که خروجی واقعی همان خروجی مطلوب است یا خیر (pass/fail). که pass نشان می‌دهد که خروجی واقعی همان خروجی مطلوب است و fail نشان‌دهنده مغایرت این دو است. بنابراین سیگنال خطای تولیدی در این روش باینری می‌باشد. اگر سیگنال خطا وضعیت عدم تطابق دو خروجی (واقعی و مطلوب) را نشان دهد، آنگاه شبکه پارامترهای خود را بطور مکرر تعدیل می‌نماید تا اینکه به خروجی مطلوب برسد.

در این روش نشان‌دهنده وجود ندارد که معین کند پاسخ خروجی در جهت صحیح در حال حرکت است یا اینکه پاسخ خروجی به چه میزانی به پاسخ مطلوب نزدیک است. هنگام استفاده از این روش حدود معینی باید بنا نهاده شود تا اینکه شبکه بطور نامحدود سعی در بدست آوردن خروجی مطلوب ننماید.

۱-۶-۳ - آموزش بدون نظارت

در این روش برخلاف شیوه آموزش با نظارت نیاز به معلم نیست یعنی خروجی هدف وجود ندارد. در خلال آموزش، شبکه الگوهای آموزشی خود را از طریق ورودیهایش دریافت می‌کند و به شکل دلخواه آنها را در طبقه‌های مختلفی طبقه‌بندی می‌کند. هنگامیکه شبکه یک ورودی را دریافت می‌کند، پاسخی در خروجی ظاهر می‌کند که نشاندهنده طبقه‌ای است که آن ورودی بدان تعلق دارد. اگر طبقه‌ای برای این ورودی یافت نشود آنگاه یک طبقه جدید تشکیل می‌شود.

۱-۶-۴ - آموزش رقابتی

آموزش رقابتی شکلی از آموزش بدون نظارت است. شبکه‌هایی که از این شیوه آموزش استفاده می‌کنند در لایه خروجی خود از چند نرون تشکیل یافته‌اند، هنگامیکه یک ورودی به شبکه اعمال می‌شود تمام نرونهای خروجی با یکدیگر رقابت می‌کنند تا یکی از نرونهای خروجی مقدار غیر صفر داشته باشد که به آن نرون برنده می‌گویند که معین کننده نرونی از نرونهای لایه ورودی یا بیشترین ورودی است و برای ورودی دیگر نرون دیگری از نرونهای لایه خروجی برنده می‌شود.

۱-۶-۵ - الگوریتمهای آموزش

در این بخش سه قانون معروف برای آموزش شبکه‌های عصبی به اختصار معرفی می‌گردد که عبارتند از:

۱-۶-۵-۱ - قانون دلنا

این قانون بر مبنای ایده تعدیل پیوسته وزنها بنا نهاده شده است به نحوی که خطا (دلنا) مابین خروجی مطلوب و خروجی واقعی کاهش یابد. این قانون همچنین به قانون آموزش Windrow-Hoff معروف است.

۱-۶-۵-۲ - قانون پس انتشار خطا یا کاهش گرادیان

در این شیوه وزنها بوسیله یک مقدار که متناسب با مشتق (گرادیان) اول خطای بین خروجی مطلوب و خروجی واقعی و با توجه به مقدار وزنه تعدیل می‌یابد. در این روش هدف کاهش تابع خطا و رسیدن به حداقل مطلق است.

۱-۶-۵-۳- قانون آموزش هب

در سال ۱۹۴۹ دونالد هب بیان کرد هنگامی که اکسون عصبی A، به اندازه کافی به سلول عصبی B نزدیک باشد، بنحوی که بطور مکرر آن را تحریک کند، آنگاه تعدادی تغییرات متابولیکی در یک یا هر دو سلول رخ می‌دهد بنحوی که قدرت سلول A (بعنوان سلولی که موجب آتش کردن یا تحریک سلول B شده است) افزایش می‌یابد. بنابراین قدرت سیناپس (معروف به وزنه w) بین دو سلول A و B مطابق درجه عملکرد متقابل بین ورودی و خروجی تعدیل می‌یابد.

۱-۷- ویژگیهای شبکه عصبی

شبکه‌های عصبی مصنوعی همانند نوع طبیعی باید ویژگیهای زیر را دارا باشند:

۱- محاسبات گروهی و یارانه‌ای

-محاسبات بصورت گروهی و یارانه‌ای توسط نرونها انجام می‌شود.

-پردازش بین تمام نرونها توزیع یافته و در نقطه‌ای خاص انجام نمی‌شود.

۲- آموزش

-شبکه بطور خودکار پیوند بین ورودی و خروجی را انجام می‌دهد.

-برنامه لازم برای انجام عملی خاص در حین آموزش توسط شبکه بوجود می‌آید.

-شبکه عصبی خود را بوسیله یا بدون معلم سازگار می‌کند و برای این کار به حضور برنامه‌نویس

نیازی ندارد.

۳- استحکام

-عملیات در شبکه نسبت به نقصهای جزئی در آن غیرحساس است.

-عملیات در شبکه نسبت به ورودی که دارای چند جزء غیردقیق هستند، غیرحساس است.

۴- انجام عملیات به شکل غیر همزمان

شبکه عصبی طبیعی ساعت مشخص برای همزمان سازی عملیات ندارد، در حالیکه تعدادی از

شبکه‌های عصبی مصنوعی به ساعت نیاز دارند.

۱-۷-۱- پارامترهای مهم شبکه‌های عصبی مصنوعی

راندمان در شبکه‌های عصبی مصنوعی به شکل تعداد الگوهای بدرستی فراخوانده شده به الگوهای ورودی هنگامیکه به شکل کاملاً درست یا همراه با نویز یا بصورت ناقص باشند، تعریف می‌گردد. در هنگام طراحی شبکه‌های عصبی مصنوعی باید نکات زیر را مدنظر قرار گیرد:

- ساختار شبکه
- تعداد لایه‌های شبکه
- تعداد نرونها در هر لایه
- نوع الگوریتم آموزش
- تعداد تکرارها برای هر الگو در خلال آموزش
- تعداد محاسبات در هر تکرار
- سرعت فراخوانی یک الگو
- راندمان شبکه
- میزان انعطاف شبکه (یعنی تعداد نرونهای خراب و میزان کارایی شبکه)
- ظرفیت شبکه یا تعداد الگوهایی که شبکه می‌تواند فراخوانی کند.
- میزان سازگاری شبکه (یعنی اینکه شبکه بعد از آموزش تا چه وسعتی می‌تواند خود را با شرایط جدید سازگار کند)
- مقادیر بایاس
- مقادیر آستانه
- حدود وزنه‌ها
- انتخاب تابع عملکرد غیر فعلی و تعیین حدود عملکرد نرون
- میزان ایمنی شبکه در مقابل نویز (یعنی تعداد مؤلفه‌های اشتباه در الگوی ورودی که شبکه با توجه به وجود آن می‌تواند خروجی مطلوب را فراخوانی کند).
- حالت ماندگار یا مقادیر نهایی وزنه‌ها که نشاندهنده برنامه ذخیره شده در شبکه است.

۱-۷-۲- قدرت تفکیک شبکه‌های عصبی مصنوعی

شبکه‌های عصبی به منظور تفکیک بین الگوهای ورودی آموزش می‌بینند تا بتوانند به هر ورودی پاسخ مطلوب متناظر آن را در خروجی ارائه دهند. در کاربردهای متعددی شبکه‌های عصبی موظف هستند که علاوه بر تشخیص الگوی ورودی طبقه یا کلاسی که آن الگو بدان تعلق دارد را معین کنند. این ویژگی شبکه‌های عصبی به قدرت تفکیک، یعنی تعداد الگوهایی که شبکه می‌تواند از هم تمیز دهد معروف است و هنگامی برتری خود را نسبت به روشهای متداول نشان می‌دهد که تعداد الگوهای موجود برای طبقه‌بندی زیاد باشد. قدرت تفکیک شبکه‌های عصبی مصنوعی بستگی به مشخصات خطی یا غیرخطی آن دارد که معین می‌کند یک شبکه خطی، چند خطی و یا غیر خطی است.

۱-۷-۳- تعمیم یافتگی در مقابل حفظ

یکی از ویژگیهای مهم شبکه عصبی قدرت تعمیم یافتگی آنهاست. تعمیم یافتگی یعنی توانایی شبکه در پاسخگویی درست به الگوهایی که تابحال با آنها مواجه نشده است. البته شرط صحت پاسخ شبکه این است که روابط درونی بین مؤلفه‌های بردار ورودی همان رابطه یا روابط موجود در بردارهای آموزشی از قبل آموخته باشد. توانایی درونیابی از میان داده‌های آموزشی لزوماً به مفهوم خوب تعمیم یافتگی نیست. یک شبکه طبقه‌بندی کننده‌ای که به شکل مناسب آموزش دیده، باید به یک داده آزمایشی با همان دقتی که به داده آموزشی پاسخ می‌دهد، جواب بدهد. اما اگر با دقت کمتری به داده آموزش پاسخ بدهد آنگاه این نتیجه قابل استنباط است که درجه آزادی شبکه به میزان درستی اختیار نشده است.

۱-۷-۴- استخراج مشخصه

یکی از مسائل درجه اول در تشخیص الگو این است که چگونه می‌توان مشخصه تفکیک کننده الگوها را از میان داده‌ها استخراج نمود. روش ریاضی برای انتخاب مشخصه این است که

خواص مشترک الگوها را تعیین کنیم و از این خواص برای کاهش ابعاد بردارهای الگو استفاده نماییم.

۱-۷-۵- کنترل آموزش

برای کنترل آموزش شبکه می‌توان از شرایط کنترل مختلفی استفاده کرد که در اینجا متداولترین آنها ذکر می‌شود:

- اگر تعداد تکرارها از حد معینی بیشتر شود، آنگاه آموزش متوقف شود (در اینجا منظور از یک تکرار عرضه تمامی الگوها به شبکه است).

- اگر a (در اینجا a همان سرعت آموزشی است) کمتر از حد معینی بشود، آنگاه آموزش را متوقف کنید.

- اگر مجموع مربعات خطا در خروجی کمتر از حد معینی بشود، آنگاه آموزش را متوقف شود.

- اگر تغییرات تمامی وزنه‌ها از حد معینی کمتر گردد، آنگاه آموزش را متوقف کنید.

۱-۷-۶- مدرج کردن مشخصه‌های ورودی

مشخصه‌ها باید بگونه‌ای مدرج شوند که وزنه‌های شبکه بتوانند تعدیل شوند و اهمیت مشخصه‌ها از بین نرود. این مسأله زمانی مورد توجه قرار می‌گیرد که داده‌های ورودی بسیار بزرگ بوده و در نتیجه برای کاهش اندازه آنها به منظور سادگی پردازش آنها را مدرج می‌نمایند.

۱-۸- طبقه‌بندی شبکه‌های عصبی مصنوعی

شبکه‌های عصبی دارای انواع مختلفی هستند که بعضی از آنها از ساختار موجود در سیستم اعصاب جانداران مشتق شده‌اند ولی برخی نیز بطور کاملاً مستقل و بدون هیچ منشاء فیزیولوژیکی می‌باشند و از کارهای مهندسان و ریاضیدانان منتج شده‌اند. ویژگی‌های هر شبکه عصبی را در سه راستا می‌توان عنوان نمود: ساختار و آرایش شبکه، مشخصه نرون‌ها شبکه، قوانین و روابطی که تحت آن قوانین شبکه آموزش داده می‌شود.

قوانین آموزش هر شبکه عصبی مقدار اولیه وزنه‌های شبکه را مشخص کرده و بیان می‌دارد که چگونه باید از وزنها را تغییر داد تا عملکرد شبکه به رفتار مطلوب نزدیک گردد. به عبارت دیگر قوانین آموزش با توجه به نوع شبکه عصبی راه یادگیری مناسب و فراخور حال آن شبکه را ارائه می‌دهند. ساختار و آرایش شبکه نیز با توجه به الگوهای مورد مطالعه طراح و زوایایی که طراح مدنظر دارد پیش‌بینی و ارائه می‌گردد. در واقع ارتباط بین اجزاء شبکه و وضعیت آنها نسبت به هم ساختار شبکه را مشخص می‌کند. طراحی ساختار مناسب و بیان قوانین آموزش شبکه عصبی از مسایل هستند که امروزه مورد بحث و بررسی قرار می‌گیرند. با عنایت به نکات فوق وجود انواع مختلف شبکه‌های عصبی امری طبیعی بنظر می‌رسد. با توجه به تنوع فراوان شبکه‌های عصبی، می‌توان این شبکه‌ها را بوسیله چند مشخصه مهم آنها طبقه‌بندی کرد. این مشخصه‌ها عبارتند از:

- نوع ورودی: ورودی شبکه‌های عصبی ممکن است پیوسته (آنالوگ) و یا گسسته (دیجیتالی) باشند. شبکه‌ها دارای ورودی گسسته معمولاً در کاربردهای مربوط به دسته‌بندی و تفکیک اطلاعات استفاده می‌شوند. حال آنکه کاربردهای مربوط به پردازش صوت و تصویر شناسایی و کنترل سیستمها اغلب با شبکه‌های با ورودی پیوسته سروکار دارند.

- آموزش با ناظر یا بدون ناظر: ساختار هر شبکه عصبی قانون یادگیری خاص خود را دارا می‌باشد. در آموزش با ناظر برای هر ورودی، خروجی مورد شبکه نیز موجود بوده که اطلاعات آن در فرایند آموزش بکار گرفته می‌شوند. در طی فرایند آموزش یک شبکه عصبی، اعمال هر تحریک ورودی باعث تولید یک عکس‌العمل خروجی می‌شود. این واکنش با واکنش مطلوب (خروجی مطلوب) شبکه مقایسه می‌شود. اگر خروجی واقعی با خروجی مطلوب تفاوت داشته باشد مقدار خطا محاسبه شده و در محاسبات مربوط به تصحیح وزنه‌های اتصالات شبکه اعمال می‌شود. هدف از این محاسبات به حداقل رساندن خطا و در صورت امکان به صفر رساندن آن می‌باشد. فرایند به حداقل رساندن خطا نیاز به مسیر (مدار) ویژه دارد، که به آن «معلم» یا «ناظر» می‌گویند.

در مقایسه با روش آموزش با ناظر، آموزش بدون ناظر به معلم نیازی ندارد. این موضوع بدین معناست که در این روش خروجی مطلوب ندارد. به عبارت دیگر در آموزش بدون ناظر، مقدار از پیش تعیین شده و مورد نظر برای خروجی، حتی در صورت وجود، تأثیری در فرایند آموزش شبکه ندارد. در این روش شبکه عصبی در طی فرایند آموزشش، الگوهای ورودی یا تحریکات

مختلفی را دریافت می‌دارد و سپس به صورت دلخواه این الگوها را در گروه‌های مختلف جای می‌دهد. وقتی که یک تحریک جدید به شبکه ارجاع داده می‌شود، شبکه در پاسخ یک خروجی تولید می‌کند که این خروجی گروهی را که تحریک بدان متعلق است، مشخص می‌کند. و اگر گروهی برای این تحریک جدید یافت نشود، یک گروه جدید توسط شبکه تولید می‌شود.

با وجود اینکه در آموزش بدون ناظر به معلم نیاز می‌باشد، ولی در عین حال یکسری راهنمایی‌های کلیدی درباره اینکه گروهها بر چه اساس شکل بگیرند، مورد نیاز می‌باشد. دسته‌بندی ورودی‌های ممکن است بر اساس شکل، رنگ، غلظت مواد یا بر اساس ویژگی‌های دیگر انجام شود. شبکه‌های بدون ناظر در طبقه‌بندی و تفکیک اطلاعات کاربرد دارند. ولی شبکه‌های دارای آموزش با ناظر در مجموع کارایی بیشتری دارند، و در عین حال نحوه آموزش این شبکه‌ها پیچیده و دشوار می‌باشد.

- شبکه‌های دارای پس‌خور و بدون پس‌خور: بعضی از انواع شبکه‌ها دارای پس‌خوری از خروجی به ورودی می‌باشند. این پس‌خور می‌تواند به انواع مختلف صورت پذیرد. این شبکه‌ها در واقع دارای ساختار دینامیکی بوده و توسط مجموعه‌ای از معادلات تفاضلی قابل بیان می‌باشند. این خانواده از شبکه‌های عصبی را به اختصار شبکه‌های بازگشتی می‌گویند. در مقابل گروه دیگری از شبکه‌های عصبی که فاقد پس‌خور می‌باشند و هیچگونه دینامیکی را دارا نمی‌باشند، موجود هستند که رابطه بین ورودی و خروجی اینگونه شبکه‌ها توسط یک تابع جبری استاتیکی قابل بیان است. از آنجائیکه به تدریج انواع پیچیده‌تری از شبکه‌های عصبی پا به عرصه می‌گذارند، و شبکه‌های جدید بیشتر دارای ساختارهای مرکب هستند، طبقه‌بندی بر اساس مشخصه‌های یاد شده تنها شامل انواع اصلی و ساده شبکه‌های عصبی می‌گردد. برای مثال شبکه عصبی حافظه‌دار شبکه‌ای است که ساختار کلی آن فرم شبکه‌های پیش‌خور را دار می‌باشد، در حالیکه بطور محلی از انواع بازگشتی می‌باشد.

پرکاربردترین شبکه‌های عصبی مصنوعی شبکه‌های هستند که دارای قانون آموزش با ناظر می‌باشند. این نوع شبکه‌ها به دلیل قابلیت فراوان در تخمین توابع و نگاشت‌های غیر خطی کاربردهای فراوانی در حیطه مسایل مهندسی پیدا کرده‌اند. از مشهورترین انواع این شبکه‌های عصبی می‌توان به شبکه‌های پرسپترون چند لایه و شبکه‌های هاپفیلد اشاره نمود. شبکه پرسپترون چند لایه، شبکه‌ای بدون پس‌خور و شبکه هاپفیلد شبکه‌ای با پس‌خور می‌باشد.

از جمله شبکه‌هایی که دارای قانون آموزش بدون ناظر هستند نیز می‌توان به شبکه‌های کوهنن و شبکه ART اشاره نمود. در ادامه این بخش به معرفی نمونه‌های مهمتر شبکه‌های عصبی می‌پردازیم؛ در این راستا و به منظور شروع کار مدل ساده Mcculloch- Pitts را معرفی می‌کنیم. سپس از آنجایی که تا به امروز پرکاربردترین و کارآمدترین شبکه عصبی مورد استفاده در حیطه مسایل مهندسی شبکه پرسپترون چند لایه با الگوریتم آموزش پس‌انتشار خطا بوده است، به تشریح مفصل این شبکه‌ها و روش آموزش پس انتشار خطا (BP) می‌پردازیم. در ادامه نیز بطور مختصر شبکه‌های هاپفیلد و کوهنن را معرفی می‌کنیم.

۱-۸-۱- مدل Mcculloch- Pitts

Warren Mcculloch نروبیولوژیست و Walter Pitts ریاضیدان در سال ۱۹۴۳ اولین مدل ریاضی از یک نرون بیولوژیکی ایده‌آل را ارائه کردند. این مدل که مدلی بسیار ساده می‌باشد، پایه و اساس و محرکی برای کارهای بعدی در زمینه شبکه‌های عصبی بوده است (شکل ۱-۱۱) در این مدل هر ورودی نرون X_j توسط مقادیر W_{ij} که نشان‌دهنده وزنه‌های اتصال می‌باشند، وزن دار می‌شود. تمام ورودیهای وزن دار شده با هم جمع شده و اگر این مجموع به یک سطح آستانه‌ای معین برسد، خروجی نرون تولید می‌شود. این خروجی از طریق اعمال تابع غیر خطی f بر روی مجموع فوق بدست می‌آید. در این رابطه X_{ij} سیگنال ورودی یا تحریک زام اعمال شده بر نرون i ام می‌باشد، تابع f تابع غیر خطی و O_i واکنش خروجی نرون i ام می‌باشد. در این مدل فرض می‌شود که مقدار بایاس و وزنه‌های اتصالات W_{ij} به حالت پایدار رسیده‌اند. این بدان معناست که فرایند آموزش یا تطبیق نرون به اتمام رسیده است.

۱-۸-۲- پرسپترون

در بخشهای قبل سعی بر این شد تا با دیدی کلی مطالبی هم راجع به بیولوژی شبکه‌های عصبی طبیعی و نحوه مدل‌سازی یک شبکه عصبی مصنوعی براساس نمونه طبیعی آن ارائه گردد و هم نحوه آموزش و عملکرد یک شبکه عصبی مصنوعی و پارامترهای مهم در طراحی آن مورد توجه

قرار گیرد. تا بدین وسیله آمادگی لازم برای تحلیل یک شبکه عصبی مصنوعی به نام پرسپترون حاصل شود.

پرسپترون اولین شبکه عصبی مصنوعی بود که در سال ۱۹۵۸ توسط Fran Rosenlatt ارائه گردید: وی روانشناس و متخصص در بیولوژی نرونی بود. او سعی کرد تا با مدل خویش مدلی ریاضی ارائه دهد که بتواند کار شناسایی الگوی سیستم بینایی انسان را شبیه‌سازی کند. بر همین اساس در حال حاضر از پرسپترون در کاربردهایی نظیر طبقه‌بندی اشکال، شناسایی نمادها و سیستم بینایی ربات استفاده به عمل می‌آید. شکل (۱-۱۲) ساختار اولیه پرسپترون را نشان می‌دهد که اولین لایه وزنه‌ها (بین واحدهای ورودی و واحدهای لایه پنهان) دارای مقادیر تصادفی ثابتی هستند و در حین آموزش تغییر نمی‌کنند. در حالی که دومین لایه وزنه‌ها (بین واحدهای لایه پنهان و واحد لایه خروجی) تعلیم پذیر هستند. در پرسپترونهای چند لایه بجز اولین وزنه‌ها، لایه‌های دیگر وزنه‌ها تعلیم پذیر هستند. شکل (۱-۱۳) ساختار یک پرسپترون چند لایه را نشان می‌دهد. ساختار دیگری برای پرسپترون وجود دارد که به نام ساختار بدون لایه با اتصالات میان بر معروف است. در این ساختار لایه ورودی مستقیماً به لایه خروجی اتصال می‌یابد و وزنه‌های این اتصالات نیز ثابت هستند. شکل (۱-۱۴) این ساختار را نشان می‌دهد که اتصالات با خطوط پرنرنگ نشان‌دهنده وزنه‌های ثابت هستند.

پرسپترون تک لایه با وزنه‌های تعلیم پذیر نوع خاصی از پرسپترون اولیه است که در آن توابع عملکرد واحدهای لایه پنهان همگی توابعی همانی هستند.

پرسپترون تک لایه قادر است الگوهای بطور خطی تفکیک پذیر را طبقه‌بندی کند. در حالیکه پرسپترون چند لایه می‌تواند الگوهای بطور غیرخطی تفکیک پذیر را نیز طبقه‌بندی نماید. به همین دلیل در اکثر کاربردهای امروزی به منظور شناسایی الگو و طبقه‌بندی آنها از پرسپترون چند لایه به خاطر قدرت آن در تفکیک‌پذیری استفاده می‌شود. برای آموزش پرسپترون چند لایه از الگوریتم آموزشی به نام آموزش پس انتشار خطا یا قانون دلتای تعمیم یافته استفاده می‌گردد. در هنگام استفاده از الگوریتم پس انتشار خطا تمام توابع عملکرد در تمام لایه‌ها (لایه ورودی می‌تواند هر تابع عملکرد دلخواهی داشته باشد) باید دارای خصوصیات زیر باشند:

الف: پیوسته باشند.

ب: مشتق پذیر باشند.

ج: یکنوای افزایشی باشند.

علاوه بر موارد فوق برای ارتقای بازدهی محاسباتی مطلوب است که مشتق آن نیز براحتی قابل محاسبه باشد. یعنی بتوان مقدار مشتق تابع در یک نقطه را براساس مقدار تابع در همان نقطه بیان نمود. معمولاً انتظار می‌رود که تابع عملکرد از نوع اشباع‌پذیر باشد، یعنی بطور مجانبی به یک حداکثر و حداقل معین و محدود میل کند. در این پروژه از تابع عملکرد تانژانت هیپربولیک به خاطر داشتن تمام موارد فوق و همچنین به خاطر دو قطبی بودنش که موجب افزایش سرعت تعلیم می‌گردد، استفاده شده است. الگوریتم پس انتشار خطا شامل سه مرحله است:

مرحله اول: تزریق ورودیهای آموزشی و محاسبه خروجی شبکه

مرحله دوم: محاسبه خطاها و پس انتشار آن از لایه خروجی به سمت لایه ورودی

مرحله سوم: تعدیل وزنها

۱-۸-۲-۱- شبکه پرسپترون چند لایه

محبوبترین ساختار (نوع) شبکه‌های عصبی، شبکه پرسپترون چندلایه (MLP) می‌باشد. شبکه پرسپترون چند لایه شبکه‌ای با ساختار لایه‌لایه است، بطوریکه هر لایه شامل چند نورن می‌باشد که ورودیهای آنها تنها به لایه قبلی و خروجی آنها به لایه بعدی متصل می‌گردد. این شبکه از یک لایه ورودی، تعدادی لایه مخفی (بطور معمول تنها از یک یا دو لایه مخفی استفاده می‌شود) و یک لایه خروجی مطابق شکل (۱-۱۵) تشکیل شده است. لایه‌های مخفی توسط وزن‌هایی اطلاعات را از لایه قبلی دریافت و پس از انجام عملیات مورد نظر بر روی آنها، توسط وزنهای دیگری این اطلاعات را به لایه بعدی منتقل می‌کنند، تا سرانجام به لایه خروجی منتهی گردد.

توابع فعالیت سیگموئیدی متداولترین توابع مورد استفاده در گروههای شبکه پرسپترون چند لایه می‌باشند. خروجی‌های هر گره در یک لایه به تمام ورودی‌های گره‌های لایه‌های بعدی وصل می‌شوند. اطلاعات در داخل شبکه فقط در یک جهت انتشار می‌یابند که این جهت از طرف ورودی به سمت خروجی شبکه می‌باشد. به همین خاطر این نوع از شبکه‌ها را شبکه‌های لایه پیش‌خور نیز می‌گویند.

وجود بیش از یک لایه در شبکه پرسپترون، رمز اصلی توانایی و قدرت این نوع شبکه عصبی است. شبکه پرسپترون یک لایه که پیش از پرسپترون چند لایه مطرح شده بود. با

نارسایی‌ها و محدودیتهایی روبرو شد که نه تنها کاربرد و توانایی این نوع شبکه، بلکه کل تحقیقات در زمینه شبکه‌های عصبی را تا مدتی با رکود مواجه ساخت. محدودیت شبکه پرسپترون یک لایه در این است که این شبکه فقط می‌تواند زیر مجموعه کوچکی از مسایل طبقه‌بندی و تخمین توابع را حل کند. این زیر مجموعه شامل مسایلی است که، طبقه‌بندی در فضای چند بعدی توسط یک صفحه از این فضا ممکن باشد. و یا نگاشت تابع غیر خطی مورد نظر بصورت یک نیم فضا قابل بیان باشد. پرسپترون چند لایه در مقابل می‌تواند هر نوع طبقه‌بندی را انجام داده، و هر نوع تابع غیر خطی را تقریب بزند. همچنین ثابت شده است که برای این منظور به بیش از سه لایه نیاز نمی‌باشد. یعنی با شبکه پرسپترون سه لایه می‌توان هر تابع غیر خطی را تخمین زد. لازم به یادآوری است که برای رسیدن به توانایی مورد بحث، لازم است که نرونهای شبکه دارای توابع غیر خطی باشند. (نرونهای لایه خروجی را می‌توان خطی انتخاب کرد) در صورتی که تمام نرونهای شبکه، خطی باشند، بسادگی می‌توان نشان داد که شبکه هر تعداد لایه مخفی هم که داشته باشد، معادل یک شبکه پرسپترون یک لایه خواهد بود.

قدرت دسته‌بندی و تخمین یک شبکه با توجه به تعداد لایه‌های آن در شکل ۱-۱۶ آمده است. ملاحظه می‌گردد که شبکه‌های تک لایه تنها قدرت دسته‌بندی خطی را دارا می‌باشند و در مقابل مثالهایی نظیر XOR ناتوان می‌مانند. اگر شبکه عصبی مورد نظر دارای دو لایه باشد، این شبکه قادر خواهد بود دسته‌بندیهایی را که در قالب فضاهای محدب بگنجد، جداسازی و مشخص نماید. شبکه سه لایه قادر به حل هر مسأله‌ای با فضای دلخواه می‌باشد.

همانطور که عنوان شد، یک جنبه بسیار مهم شبکه‌های پرسپترون چند لایه این است که این شبکه‌ها می‌توانند هر تابع غیر خطی پیوسته را با دقت زیادی تخمین بزنند. بنابراین شبکه‌های پرسپترون چند لایه دارای پتانسیل قابل ملاحظه‌ای برای بسیاری از کاربردها، از جمله مدلسازی و کنترل فرآیندهای غیر خطی واقعی می‌باشند.

در آموزش شبکه پرسپترون چندلایه، هدف این است که با یافتن بهترین مقدار برای وزن‌های شبکه عصبی، بتوان رابطه تابعی معیار مطلوب را که بر اساس حداقل سازی خطای یادگیری شبکه بنا نهاده شده است، حداقل کرد. معروفترین الگوریتم آموزش شبکه پرسپترون چند لایه، روش پس انتشار خطا نام دارد. لازم به یادآوری است که روش آموزش مذکور از نوع آموزش با ناظر است. روش پس‌انتظار خطا در اصل از روش بهینه‌یابی تندترین شیب مشتق می‌شود و یک

الگوریتم تکراری است. بدین معنی که وزن‌های بهینه شبکه پس از چندین مرحله تکرار محاسبات بدست می‌آید. برای آموزش شبکه لازم است ابتدا مجموعه‌ای از داده‌های ورودی-خروجی را بدست آوریم. این مجموعه تنها جزیی از دامنه و برد تابع مورد نظر که قصد تقریب آنرا داریم، تشکیل می‌دهد. الگوریتم آموزش سپس با استفاده از این اطلاعات چنان وزن‌های شبکه را تنظیم می‌کند که برای ورودی‌های موجود در مجموعه آموزش، اختلاف خروجی‌های شبکه و خروجی‌های مورد نظر به حداقل ممکن برسد.

با توجه به مطالبی که عنوان شد، باید اضافه کرد شبکه پرسپترون که دارای سه لایه باشد و نرونهای آن از تابع غیر خطی استفاده کنند، تقریباً در تمام کاربردهایی که به طبقه‌بندی داده‌ها و یا تقریب توابع غیر خطی مربوط می‌شود، به عنوان یک شبکه عصبی فوق‌العاده قوی، با توانایی‌های بسیار زیاد شناخته می‌شود. روش رایج برای آموزش این شبکه روش پس‌انتشار خطا می‌باشد که تا به امروز قوی‌ترین و کارآمدترین روش آموزش شبکه‌های پرسپترون چند لایه بوده است. این روش در ادامه شرح داده می‌شود.

۱-۸-۲-۲- الگوریتم پس انتشار خطا

در این بخش مراحل اجرای الگوریتم پس انتشار خطا با توجه به ساختار شبکه عصبی چند لایه نشان داده شده در شکل (۱-۱۷) توضیح داده می‌شود:

- مرحله صفر: تعیین وزنه‌ها (وزنه‌ها را برابر مقادیر تصادفی کوچکی قرار دهید)
- مرحله اول: تا زمانی که شرایط توقف حاصل نشده مراحل ۲ تا ۹ را دنبال کنید.
- مرحله دوم: برای هر زوج آموزشی (ورودی و خروجی متناظر با آن) مراحل ۳ تا ۸ را دنبال کنید. مرحله پیشرو (تزریق ورودیها به شبکه و یافتن پاسخ آنها):
- مرحله سوم: هر واحد ورودی $(X_i, i=1, \dots, n)$ (در اینجا توابع عملکرد واحدهای ورودی همانی هستند) سیگنال ورودی خود را دریافت کرده و آن را به تمام واحدهای لایه بالاتر منتشر می‌کند (لایه پنهان).
- مرحله چهارم: هر کدام از واحدهای لایه پنهان مجموع ورودیهای وزن دار شده خویش را معین می‌کند. $(Z_j, j=1, \dots, p)$.

$$z_m = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

سپس تابع عملکرد خویش را برای تعیین خروج بکار می‌گیرد:

$$z_j = f(z_m)$$

و این سیگنال را به تمام واحدهای لایه بالاتر منتشر می‌کند.

- مرحله پنجم: هر کدام از واحدهای لایه خروجی $(Y_k, k = 1, \dots, m)$ ، مجموع ورودی وزن دار شده خویش را محاسبه می‌کند:

$$Y_{m_k} = w_{0k} + \sum_{j=1}^n z_j w_{jk}$$

سپس تابع عملکرد خویش را برای تعیین خروجی خود بکار می‌گیرد:

$$y_k = f(Y_{m_k})$$

پس انتشار خطا:

- مرحله ششم: هر کدام از واحدهای لایه خروجی مقدار هدف خود را که متناظر با الگوی ورودی است دریافت می‌کند و عبارت خطا را محاسبه می‌کند:

$$\delta_k = (t_k - y_k) f'(Y_{m_k})$$

سپس مقدار تغییرات وزنه‌ها و بایاس منتهی به خود را به منظور تعدیل وزنه‌ها و بایاس در پایان محاسبه می‌کند:

$$\Delta w_{jk}(t+1) = \alpha \delta_k Z_j + \mu [w_{jk}(t) - w_{jk}(t-1)] = \alpha \delta_k Z_j + \mu \Delta w_{jk}(t)$$

که در رابطه بالا، سرعت آموزش است $(\alpha \in [0, 1])$ که وظیفه آن کنترل میزان تأثیر مقدار خطا در تعدیل وزنه‌هاست و (عبارت است از ضریب تغییرات جزئی که میزان تأثیر تغییر وزنه در مرحله قبلی، در تغییر وزنه مرحله حاضر را، کنترل می‌کند که $\mu \in (0, 1)$ می‌باشد. δ_k عبارت است از خطا در خروجی نرون k . سپس ارسال δ_k به تمام واحدهای لایه پایین‌تر.

- مرحله هفتم: هر کدام از واحدهای لایه پنهان $(Z_j, j = 1, \dots, p)$ ، مجموع δ_k های ورودی از واحدهای لایه بالا را محاسبه می‌کند:

$$\delta_{m_j} = \sum_{k=1}^m \delta_k w_{jk}$$

سپس این مقدار را در مشتق تابع عملکردش در نقطه Z_{mj} ضرب نموده تا عبارت خطا بدست آید:

$$\delta_i = \delta_{m_j} f'(Z_{m_j})$$

سپس مقدار تغییرات وزنه‌ها و بایاس منتهی به خود را به منظور تعدیل وزنه‌ها و بایاس در پایان محاسبه می‌کند:

$$\Delta v_{ij}(t+1) = \alpha \delta_j x_i + \mu \Delta v_{ij}(t)$$

$$\Delta v_{i,j}(t+1) = \alpha \delta_j + \mu \Delta v_{i,j}(t)$$

- مرحله هشتم: هر کدام از واحدهای لایه خروجی وزنه‌ها و بایاس خود را تعدیل می‌کند. ($j = 0, \dots, p$)

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk}(t+1)$$

هر کدام از واحدهای لایه پنهان وزنه‌ها و بایاس خود را تعدیل می‌کند: ($i = 0, \dots, n$)

$$v_{ij}(t+1) = v_{ij}(t) + \Delta v_{ij}(t+1)$$

- مرحله نهم: در صورتیکه شرایط توقف حاصل گردید (مجموع خطای مربعی از حد معینی کمتر گردید) آموزش را متوقف کنید. در غیر اینصورت به گام اول برگردید.

توجه: جهت توقف تکرار الگوریتم پس انتشار از دو شاخص زیر بطور همزمان می‌توان استفاده نمود: الف) میانگین مربعات خطا در هر سیکل با Epoch (جمع مربعات خطا برای تمامی الگوهای یادگیری) کمتر از مقدار از پیش تعیین شده‌ای باشد و یا اینکه فرم تغییرات در پارامترهای شبکه پس از هر سیکل خیلی کوچک باشد. بایستی توجه داشت که هر سیکل برابر با تعداد تکرار، به اندازه تعداد نمونه‌های یادگیری می‌باشد. مثلاً اگر ۱۰۰ تا داده‌های نمونه یادگیری موجود است، سیکل برابر با ۱۰۰ مرحله تکرار می‌گردد.

ب) نرم گرایان خطا، خیلی کوچک باشد.

۱-۸-۲-۳- محدودیت الگوریتم پس انتشار خطا

جهت محاسبه حساسیتها برای نرونهای لایه‌های مختلف در شبکه چند لایه پرسپترون احتیاج به مشتق‌گیری از توابع تبدیل نرونها می‌باشد. از اینرو بایستی توابعی مورد استفاده قرار بگیرند که دارای مشتق باشند، به عبارتی دیگر مشتق‌پذیری توابع تبدیل تنها محدودیتی است که در الگوریتم BP موجود می‌باشد. یکی از انواع توابع تبدیل متداول که شرط فوق را دارا می‌باشند، توابع زیگموئید می‌باشد. یکی از ویژگیهای این توابع آن است که برای مشتق‌گیری احتیاج به

روشهای عددی که حامل خطای محاسباتی خواهند بود، نیست. به عبارت دیگر مشتق این توابع بر حسب مقادیر خود توابع قابل بیان می‌باشند.

۱-۸-۲-۴ - ملاحظاتی در مورد الگوریتم BP

در این بخش بعضی از ملاحظات مهم در مورد متد یادگیری پس انتشار خطا که در بخشهای قبلی ارائه شد را بر می‌شماریم:

نحوه ارائه داده‌های یادگیری

نظر به اینکه یادگیری یک نگاشت مفروض. از روی ارائه مجموعه داده‌های یادگیری ورودی-خروجی صورت می‌گیرد، در انتخاب و نحوه ارائه داده به شبکه بایستی دقت نمود. چونکه در حقیقت در پروسه یادگیری از اطلاعات موجود در نمونه‌های یادگیری جهت فهم و استنتاج نگاشت مفروض و پیاده‌سازی تقریبی آن استفاده می‌شود. از اینرو ترتیب ارائه نمونه‌های یادگیری به شبکه بایستی طوری باشد که شبکه از امکان برابر برای یادگیری همگی نمونه‌های یادگیری برخوردار باشد. به عبارتی روشنتر، نمونه‌ها در هر سیکل یادگیری، بطور تصادفی با امکان انتخاب برابر به شبکه اعمال شوند.

انتخاب مقدار اولیه

نخستین مرحله در بکارگیری الگوریتم BP، تعیین مقادیر اولیه پارامترهای شبکه عصبی MLP می‌باشد. ناگفته پیداست که یک انتخاب خوب می‌تواند کمک بزرگی در همگرایی سریعتر الگوریتم فراهم آورد. از اینرو در مواقعی که اطلاعات اولیه در مورد فضای ورودیهای شبکه موجود است، بهتر است جهت انتخاب بهتر، مقادیر اولیه شبکه مورد استفاده قرار گیرد و اما در حالتی که هیچ اطلاعی در مورد فضای برداری ورودی موجود نیست، معمول این است که مقادیر کوچکی را بطور تصادفی انتخاب نمود.

انتخاب اولیه نادرست پارامترهای شبکه MLP، منجر به گیرافتادن در آغاز راه شبکه در نقاط مینیمم محلی در فضای برداری پارامترهای شبکه می‌گردد که این خود منجر به این می‌شود که شبکه خیلی زودتر از معمول به موضعی بیافتد که منحنی یادگیری شبکه برای تعداد بزرگی از

دفعات تکرار تغییر نکند. پدیده دیگری که اتفاق می افتد این است که پدیده یادگیری بسیار کند صورت می گیرد، طوری که پس از رسیدن به نقطه‌ای از زمان، دیگر تغییر قابل ملاحظه‌ای در میزان میانگین مربعات خطا رخ نمی‌دهد. یک راه ساده و انتخاب سرانگشتی برای مقادیر اولیه پارامترهای شبکه MLP فرمول زیر است:

مقادیر اولیه پارامترهای نرون α را بطور تصادفی و غیر برابر با توزیع یکنواخت در بازه $[-3/N_i^{0.5}, 3/N_i^{0.5}]$ انتخاب کنید، جاییکه N_i تعداد ورودیهای نرون α می‌باشد.

سرعت پایین همگرایی

هر چقدر طول قدم یادگیری، a ، کوچکتر انتخاب گردد، تغییرات ایجاد شده در پارامترهای شبکه پس از هر مرحله تکرار الگوریتم BP کوچکتر خواهد بود که این خود منجر به هموارتر گشتن مسیر حرکت پارامترها می‌گردد. این مسأله موجب کندتر شدن الگوریتم BP می‌گردد. برعکس با افزایش طول قدم a ، اگر چه نرخ یادگیری و سرعت یادگیری الگوریتم BP افزایش می‌یابد، لیکن تغییرات فاحشی در پارامترهای شبکه از هر تکرار به تکرار بعد ایجاد می‌گردد که گاهاً موجب ناپایداری و نوسانی شدن شبکه می‌گردد که به اصطلاح گویند پارامترهای شبکه واگرا شده است.

۱-۲-۵- یادگیری به روش پس انتشار خطا و ارتباطش با یادگیری بیولوژیکی

در مورد یادگیری عملکردی (الگوریتم پس انتشار خطا) و ارتباطش با یادگیری بیولوژیکی نکات زیر ضروریست:

الف) در شبکه‌های عصبی مصنوعی چند لایه، از ارتباط هورمونی و غیر منطقه‌ای صرفنظر می‌شود، در حالیکه در سیستمهای واقعی، این نوع ارتباطات نقش کلیدی و مهمی در تنظیم حالت توابعی از قبیل یادگیری، توجه و تحریکهای سریع دارد.

ب) ارتباطات سیناپسی بین نرونها در شبکه چند لایه می‌تواند تشدید یا ممانعتی باشند، در حالیکه در سیستمهای عصبی واقعی نرونها معمولاً یا تشدید یا ممانعتی هستند.

ج) در یادگیری پس انتشار خطا، تنظیم پارامترها توسط فعالیت پیش سیناپسی و سیگنال یادگیری خطا که مستقل از فعالیت پس سیناپسی است، انجام می‌گیرد، در حالیکه دانش بیولوژی نرونها بحث فوق را در مورد فعل و انفعالات سیناپسی تأیید نمی‌کند.

د) توجه داریم که یادگیری پس انتشار خطا، با ناظر است. حضور ناظر، در مفهوم مغز قاعداً پیش فرض مجموعه‌ای دیگر از نرونها با خواص بدیع را می‌طلبد، وجود چنین نرونهایی، از لحاظ بیولوژیکی قابل دفاع نیست.

ه) اگر بخواهیم تغییر بیولوژیکی از پیاده‌سازی پس انتشار خطا داشته باشیم، به این جا می‌رسیم که قاعداً بایستی انتشار برگشت سیگنالهای اطلاعاتی در مسیر اکسون داشته باشیم. چنین فعل و انفعالاتی در مغز بعید است که رخ دهد.

۱-۸-۲-۶- روش پس انتشار خطا

چنانچه وزن‌های هر لایه از شبکه پرسپترون را بصورت ماتریسی در نظر بگیریم، می‌توان شبکه پرسپترون سه لایه را به شکل بلوک دیاگرام مطابق شکل ۱-۱۸ نشان داد.

در این شکل W_i ماتریس وزن لایه i ام و b_i بردار بایاس نرونهای لایه i ام و b_i تابع برداری و غیر خطی مربوط به نرونهای لایه i ام، R_i و S_i به ترتیب طول بردار ورودی و تعداد نرونهای لایه i ام و O_i بردار خروجی لایه i ام و X بردار ورودی شبکه می‌باشد. با چنین نمایشی می‌توان رابطه ورودی و خروجی این شبکه را به فرم زیر بیان کرد.

$$o = o_3 = f_3[w_3 \cdot f_2[w_2 \cdot f_1[w_1 \cdot x + b_1] + b_2] + b_3]$$

چون شبکه پرسپترون، دارای یک ساختار لایه‌ای است می‌توان برای بدست آوردن خروجی، ورودی را از لایه اول به طرف لایه آخر منتشر کرد. بدین ترتیب می‌توان در حالت کلی برای یک شبکه با M لایه، روابط زیر را برای بدست آوردن خروجی شبکه بیان کرد. این معادلات به روابط پیش خود موسوم می‌باشند.

$$\begin{aligned} O_0 &= X \\ R_k &= W_k \cdot O_{k-1} + b_k \\ O_k &= f_k(R_k) \\ O &= O_M \end{aligned} \quad k = 1, 2, \dots, M$$

برای آموزش شبکه به مجموعه‌ای از داده‌های ورودی- خروجی نیاز داریم. این داده‌ها را بصورت زوجهای (Xq, Odq) $q=(1,2,\dots,Q)$ نمایش می‌دهیم که Xd ورودی و Odq خروجی مورد نظر می‌باشد. در این صورت تابع خطای شبکه بصورت زیر تعریف می‌شود:

$$E(W, b) = \frac{1}{2} \sum_{q=1}^Q [(Od_q - O_q)^2]$$

با استفاده از روش پهنه‌یابی تندترین شیب می‌توان گفت که بهترین جهت برای تغییر مقادیر وزنهای شبکه در امداد گرادیان تابع E می‌باشد. بدین ترتیب برای تنظیم وزنهای شبکه باید از روابط زیر استفاده کرد:

$$\Delta W_k(i, j) = -\eta \cdot \frac{\partial E}{\partial W_k(i, j)}$$

$$\Delta b_k(i) = -\eta \cdot \frac{\partial E}{\partial b_k(i)}$$

در روش پس‌انتشار خطا برای محاسبه مقدار گرادیان تابع خطا در دو رابطه اخیر، الگوریتمی معرفی شده است که در آن برای هر لایه شبکه یک خطای معادل بدست می‌آید. از خطای معادل هر لایه می‌توان برای محاسبه مشتق تابع خطا نسبت به وزنهای آن لایه استفاده کرد. بدست آوردن رابطه خطای معادل هر لایه در شبکه پرسپترون چند لایه مستلزم اندکی محاسبات جبری می‌باشد. نتیجه این محاسبات را می‌توان بصورت ساده شده زیر نمایش داد.

$$\delta_{l+1} = -f'(R_{l+1}) \cdot (Od_{l+1} - O_{l+1})$$

$$K = M - 1, M - 2, \dots, 1$$

$$\delta_k = -f'(R_k) \cdot W_{k+1}^T \cdot \delta_{k+1}$$

در این روابط k خطای معادل لایه k ام می‌باشد. ماتریس $f'(R_{l+1})$ و $f'(R_k)$ بصورت زیر محاسبه می‌شوند:

$$f'(R_k) = \begin{pmatrix} f'[R_k(I)] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & f'[R_k(S_k)] \end{pmatrix}$$

همانطور که از معادلات فوق دیده می‌شود محاسبه خطای معادل از لایه آخر شروع و در هر لایه به خطای معادل در لایه بعدی نیاز است. بدین ترتیب خطای خروجی شبکه از لایه آخر به لایه اول منتشر می‌شود. این بخش از محاسبات به محاسبات پس‌انتشار موسوم است، که نام‌گذاری این روش نیز به دلیل همین بخش از محاسبات می‌باشد. با محاسبه خطای معادل در هر لایه از

شبکه می‌توان پارامترهای وزن شبکه را بگونه‌ای تنظیم کرد که با وزنه‌های جدید، خطای شبکه برای ورودیهای آموزشی کاهش یابد:

$$\begin{aligned}\Delta W_k &= -\eta \cdot \delta_k \cdot O_{k-1}^f \\ k &= 1, 2, \dots, M \\ \Delta b_k &= -\eta \cdot \delta_k\end{aligned}$$

روابطی که تا کنون ارائه شدند روابطی بودند که در حالت کلی و بصورت ماتریسی نوشته شده بودند. اکنون روابط نهایی را بصورت مفصل‌تر برای تک‌تک گره‌ها می‌نویسیم. از طرف دیگر به منظور ساده‌تر شدن محاسبات فرض می‌کنیم که تابع فعالیت استفاده شده در گره‌ها تابع معروف و متداول سیگموئید باشد. در این صورت می‌توان ثابت کرد که عبارت (برای لایه‌های خروجی و پنهان بصورت زیر محاسبه می‌شود:

$$\delta_{qj} = (Od_{qj} - O_{qj}) O_{qj} (1 - O_{qj}) \quad \text{در لایه خروجی}$$

$$\delta_{qj} = O_{qj} (1 - O_{qj}) \sum_i \delta_{qi} W_{ij} \quad \text{در لایه پنهان}$$

در این روابط اندیس i نشان دهنده شماره گره در لایه خروجی، اندیس j نشان دهنده شماره گره در لایه پنهان، اندیس h نشان دهنده شماره گره در لایه ورودی، Od نشان دهنده خروجی مورد نظر، O نشان دهنده خروجی واقعی و اندیس q نشان دهنده شماره مثال آموزشی می‌باشد. با توجه به مطلب فوق، روابط زیر برای تصحیح وزنها در لایه‌های خروجی و پنهان ارائه می‌شود.

$$\Delta_j W_{qj} = \eta \delta_{qj} O_{qj}$$

$$\Delta_j W_{jh} = \eta \delta_{qj} O_{jh}$$

ضریب η در روابط فوق نرخ یادگیری نامیده می‌شود و مشخص کننده سرعت یادگیری می‌باشد. البته ممکن است که باعث نوسان هم بشود. Rumelhart در سال ۱۹۸۶ پیشنهاد کرد که جمله‌ای بنام ممنتوم به رابطه تصحیح وزن بصورت زیر اضافه شود:

$$\Delta W_{qj}(n+1) = \eta(\delta_{qj} O_{qj}) + \alpha \Delta W_{qj}(n)$$

که در آن n نشان دهنده شماره مرتبه ارائه کل الگوهای ورودی-خروجی آموزشی به شبکه و α ضریب ممنتوم می‌باشد که ممکن است نوسانات را حذف کند ولی در بعضی موارد نیز ممکن است باعث کاهش سرعت همگرایی شود.

مطلب مهمی که در مورد الگوریتم پس‌انتشار خطا بایستی مورد توجه قرار گیرد احتمال توقف شبکه در مینیممهای موضعی (شکل ۱-۱۹) می‌باشد. به این صورت که ممکن است مقدار تابع خطا در فضای وزنها به یک مینیمم محلی برسد که در این صورت این قاعده یادگیری قادر نیست که تابع خطا را از این حالت خارج کرده و به مینیمم مطلق برساند.

۱-۸-۳- شبکه هاپفیلد

شبکه هاپفیلد از دو لایه تشکیل شده است، یکی لایه ورودی و یکی لایه هاپفیلد (شکل ۱-۲۰). هر گره که در لایه ورودی قرار دارد مستقیماً تنها به یک گره در لایه هاپفیلد وصل می‌شود. گره‌های لایه هاپفیلد نرونها می‌باشند که دارای تابع فعالیت پله‌ای یا سیگموئیدی می‌باشند. خروجی این گره‌ها وزن دار و سپس به تمام گره‌های دیگر برگشت داده می‌شود.

کارکرد شبکه بدین ترتیب است که در طی فرایند آموزش، اغلب لازم می‌شود که خروجی شبکه با ورودی آن مساوی باشد. در این صورت اگر خروجی یک نرون متفاوت از ورودی اش باشد، وزنهاى اتصالات ضعیف می‌شوند بدین ترتیب که مقادیر وزنهاى اتصالات مربوط به آن نرون کاهش داده می‌شوند، و اگر خروجی نرون با ورودی آن مساوی باشد وزنهاى اتصالات تقویت می‌شوند. شبکه آموزش داده شده سپس مورد استفاده قرار می‌گیرد، بدین صورت که یک الگوی ورودی به شبکه ارائه می‌شود. خروجی‌های شبکه سپس بطور پیوسته و از طریق وزنهاى به درون شبکه برگشت داده می‌شوند. این کار تا آنجا ادامه می‌یابد که یک شرط همگرایی برقرار شود، مثلاً خروجی‌های شبکه پس از تکرارهای متوالی به یک عدد همگرا شوند. الگوی خروجی که پس از برقراری شرط همگرایی بدست می‌آید، الگوی خروجی نهایی شبکه می‌باشد.

خروجی‌ها و ورودی‌های شبکه هاپفیلد معمولاً از نوع باینری (دودویی +۱ و -۱) می‌باشند. بنابر این کاربردهای این شبکه مربوط به جاهایی است که داده‌های باینری بطور مرتب استفاده می‌شوند. از جمله کاربردهای شبکه هاپفیلد در پردازش تصاویر از روی داده‌ها نویزدار و ناکامل تشخیص الگو در تصاویر بطور مستقل از اندازه، جهت‌یابی و موقعیت آنها می‌باشد. حضور پس‌خور در این شبکه‌ها همچنین باعث شده است که این شبکه‌ها کاربردهایی در زمینه مدلسازی و کنترل سیستمهای دینامیکی داشته باشند.

۱-۸-۴- شبکه کوهن

ویژگی بارز این شبکه در مقایسه با شبکه‌های پرسپترون چند لایه و شبکه‌های هاپفیلد این است که برای آموزش این شبکه نیاز به داده خروجی نمی‌باشد. یک شبکه کوهن از یکسری نرون تشکیل شده است که هر نرون به تنهایی الگوی ورودی شبکه را دریافت می‌دارد، همچنین مابین خود نرونها نیز اتصالات بسیار زیادی وجود دارد. (شکل ۱-۲۱)

دو سری وزن برای نرونها وجود دارد: یکسری وزنها قابل تغییر (انطباق‌پذیر) و یکسری وزنها ثابت بین خود نرونهای شبکه از وزنها سری اول بمنظور محاسبه مجموع وزن دار شده ورودیهای خارجی (ورودیهای که از خارج به شبکه اتصال می‌شوند) و از وزنها سری دوم به منظور کنترل بر هم کنش نرونها در شبکه استفاده می‌شود.

در طی فرایند آموزش، یک الگوی ورودی به شبکه ارائه می‌شود و خروجی هر نرون محاسبه می‌شود. سپس به نرونها اجازه داده می‌شود تا با هم بر هم کنش داشته باشند و در این میان نرونی که بیشترین تحریک را دریافت می‌کند (مثلاً نرونی که دارای بزرگترین خروجی باشد) شناسایی می‌شود. تنها به این نرون و نرونهایی که با یک فاصله معین در اطراف آن قرار دارند اجازه داده می‌شود، تا به منظور حساس‌تر شدن در برابر الگوی خاص آموزش داده شده، وزنها اتصالاتشان را تصحیح کنند. این نحوه آموزش باعث سازمان یافتن نقشه گره‌های خروجی شبکه می‌شود. بدین ترتیب که بخشهای مختلف شبکه (نقشه شبکه) هر کدام در برابر الگوهای ورودی خاصی حساس می‌شود. بنابراین شبکه کوهن دارای ویژگی خود سازماندهی بوده و دارای قابلیت شناسایی (تشخیص) می‌باشد. نمونه‌هایی از کاربرد شبکه کوهن تشخیص تصاویر و سیگنالهای صوتی می‌باشد.

۱-۸-۵- شبکه‌های عصبی مبتنی بر توابع پایه شعاعی (RBF)

شبکه‌های RBF علاوه بر لایه ورودی که متشکل از گره‌های منبع است. شامل دو لایه دیگر نیز هستند. یک لایه پنهان که شامل S1 سلول است و از توابع غیر خطی استفاده می‌کند و لایه دیگر که لایه خروجی خطی شامل S2 سلول (نرون) است که S2 بیان کننده تعداد خروجی‌های شبکه آموزش یافته می‌باشد.

ساختار یک نرون براساس توابع پایه شعاعی مطابق شکل ۱-۲۲ است.

تابع تبدیل استفاده شده جهت سلول بصورت زیر می باشد:

$$\text{rad bas}(n)=e^{-n^2}$$

که این تابع در شکل (۱-۲۳) نشان داده شده است.

اگر ورودی آن یعنی تفاضل (w-p) در حداقل مقدار خود و صفر باشد، برابر حداکثر مقدار خود یعنی ۱ می شود و با افزایش این تفاضل میزان تابع کاهش می یابد.

این نوع شبکه ممکن است در لایه مخفی به تعداد بیشتری نرون (سلول) نسبت به شبکه های پیشخورد نیاز داشته باشد ولی معمولاً در مدت زمان بسیار کمتری می تواند به خطای دلخواه رسیده و همگرا شود.

در شکل (۱-۲۴) ساختار کلی این شبکه نشان داده شده است. با استفاده از خروجیهای محاسبه شده در لایه مخفی در پاسخ به یک بردار ورودی و با استفاده از یک پاسخ مطلوب در خروجی، وزنها با روش آموزش با نظارت معین می شوند.

فرآیند نگاشت ورودی-خروجی در شبکه RBF در دو مرحله انجام می شود:

الف- یک تبدیل غیر خطی که فضای ورودی مختلط را به فضای واسطه حقیقی نگاشت می کند.

ب- یک تبدیل خطی که فضای واسطه حقیقی را به فضای خروجی مختلط نگاشت می کند.

۱-۸-۶- مقایسه شبکه های عصبی پیشخور (پرسپترون) و شبکه های عصبی

مبتنی بر توابع پایه شعاعی:

- شبکه های عصبی مبتنی بر توابع پایه شعاعی (RBF) تنها یک لایه مخفی دارد، در حالیکه شبکه های پیشخورد می تواند دارای یک یا چند لایه مخفی باشد.

- در شبکه های RBF توابع انتقال بین لایه ورودی و لایه مخفی غیر خطی اند و توابع انتقال بین لایه مخفی و لایه خروجی خطی هستند ولی در شبکه های پرسپترون چند لایه، نگاشت بین هر لایه مخفی و لایه قبلی اش غیر خطی است و توابع انتقال لایه خروجی ممکن است خطی یا غیر خطی باشند.

- سلولهای مخفی در شبکه‌های RBF در واقع فاصله (اندازه هندسی یا اندازه اقلیدسی) مابین بردار ورودی و مذکور آن را پیدا می‌کند، در حالیکه در شبکه‌های پیشخورد سلول مخفی، حاصلضرب داخلی بردار ورودی و بردار وزن آن سلول را مشخص کند.

- شبکه پیشخور تقریب جهانی (عمومی) برای نگاشت غیر خطی ورودی-خروجی پیدا می‌کند، این تقریب حتی برای قسمتهایی از فضای ورودی صادق است که هیچگونه اطلاعاتی از آن قسمت شبکه داده نشده باشد. شبکه RBF، تقریب‌های غیر خطی مکانی برای نگاشت ورودی-خروجی پیدا می‌کند، در نتیجه سرعت یادگیری شبکه نسبت به شبکه پیشخورد بیشتر خواهد بود و همچنین شبکه RBF نسبت به ترتیب ارائه ورودیها کمتر حساس است ولی برای بدست آوردن یک توابع هموار ممکن است تعداد سلولهای لایه مخفی زیادتر شود.

- شبکه‌های عصبی RBF و شبکه‌های پرسپترون چند لایه این خاصیت مشترک هستند که هر دو یک تقریب زن پیشخورد هستند و هر دو نگاشت ورودی-خروجی ارائه می‌دهند.

- شبکه‌های RBF در حل مسائل کلاسه‌بندی الگوهای پیچیده استفاده و کاربرد بسیاری دارند.

- گروههایی که بطور تحقیقاتی بر روی شبکه‌های RBF مطالعاتی انجام داده‌اند، معمولاً روند یادگیری شبکه را به دو مرحله تقسیم کرده‌اند:

الف- یادگیری مراکز و پهنه توابع شعاعی در لایه مخفی

ب- یادگیری وزنها متصل از لایه میانی به لایه خروجی

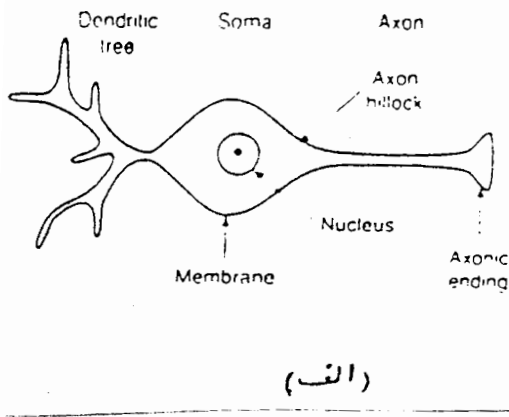
نحوه انتخاب نرونها در این شبکه نیز معمولاً بصورت آزمایش و خطاست. بدین ترتیب که ابتدا الگوریتم یادگیری با تعداد زیاد نرون در لایه مخفی شروع شده و سپس تا جایی که بتوان از این مقدار نرونها کاسته می‌شود، این کاستن تعداد نرونها با کمینه کردن میزان خطای محاسباتی همراه است و در خطای مطلوب که با تست کردن ورودیهای آموزش نیافته بدست می‌آید، الگوریتم آموزش به پایان می‌رسد.

۹-۱- خلاصه

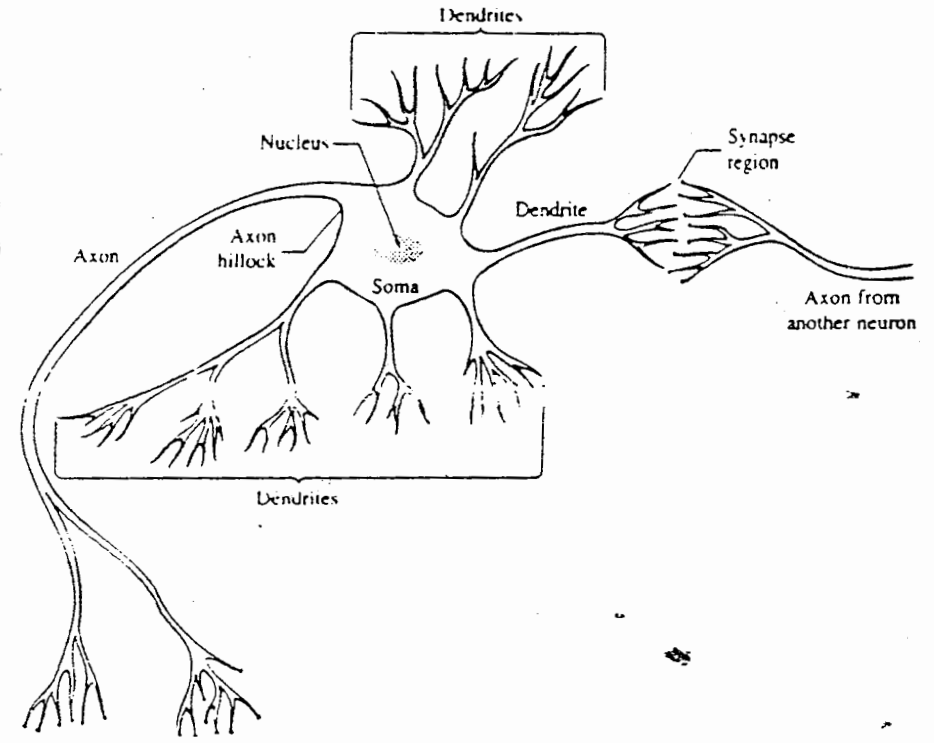
در این فصل مبانی فیزیولوژیکی شبکه‌های عصبی و نحوه مدل نمودن آنها توسط مدل‌های مصنوعی بیان شد. نوعی از شبکه‌های عصبی مصنوعی به نام پرسپترون چند لایه و نحوه آموزش آن

توسط قاعدهٔ پس انتشار خطا بیان شد. همچنین ملاحظات عملی احاطه کننده در طراحی، آموزش و تست شبکه‌های عصبی به اختصار بیان شد.

مطالب این فصل از منابع [۱]، [۸]، [۱۰] و [۱۵] آورده شده است.



(الف)

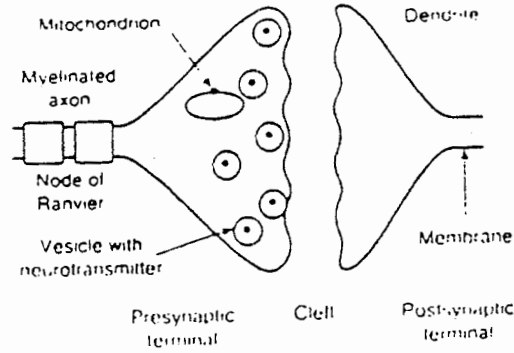


(ب)

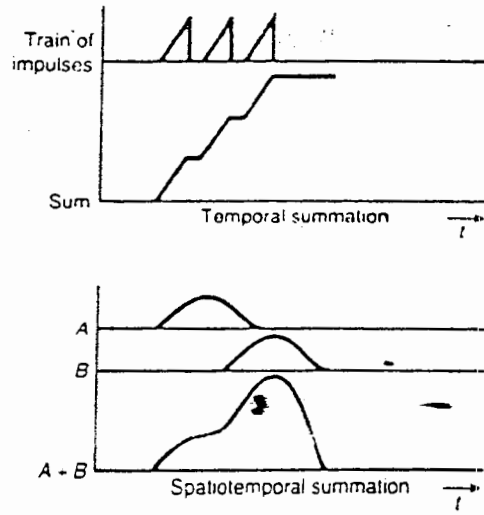
شکل ۱-۱: نرون بیولوژیکی

الف- مدل ساده تر

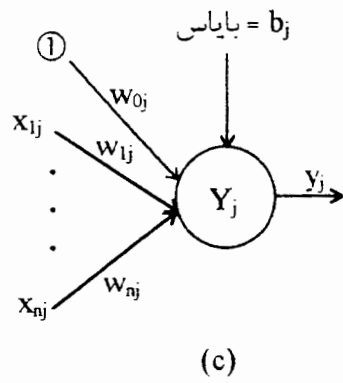
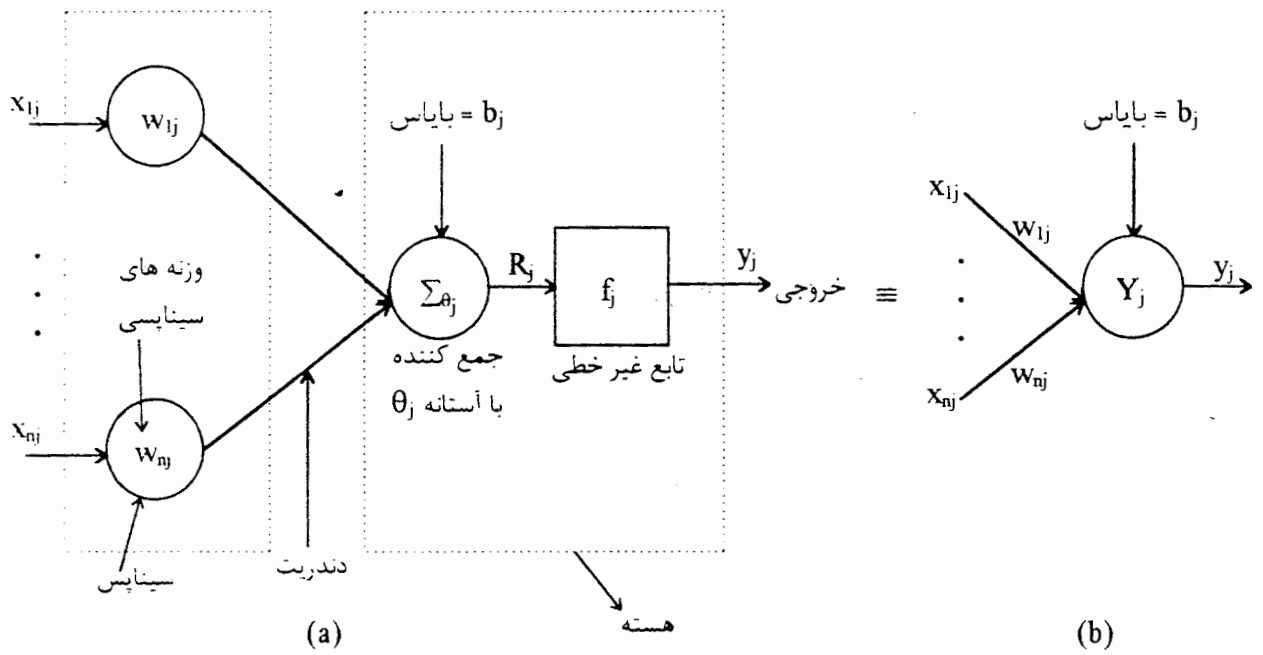
ب- مدل واقعی تر



شکل ۲-۱: جزئیات سیناپس



شکل ۱-۳: جمع‌زنی زمانی و مکانی

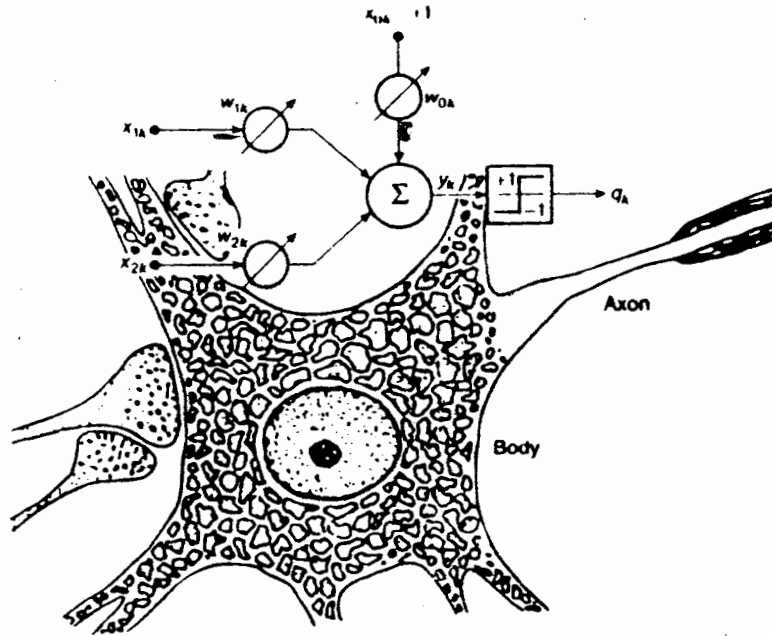


شکل ۴-۱: مدل پایه نرون

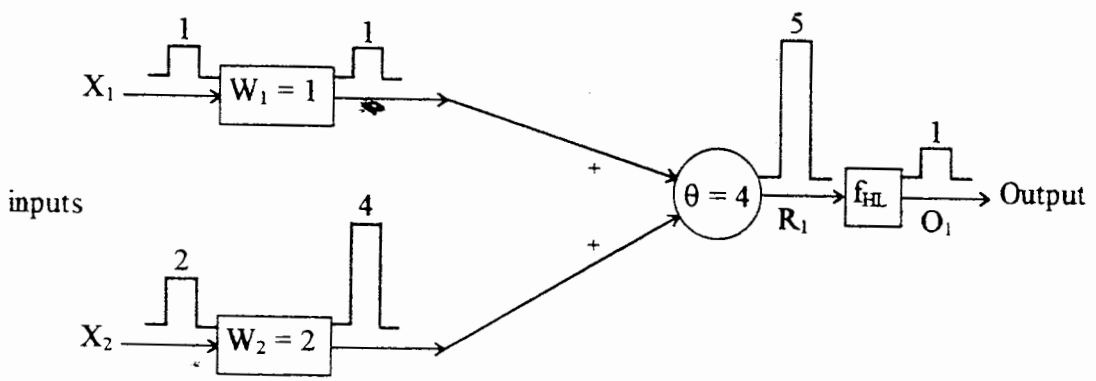
(a) شکل تحلیلی یک نرون مصنوعی

(b) علامت اختصاری نرون به همراه ورودیهای آن

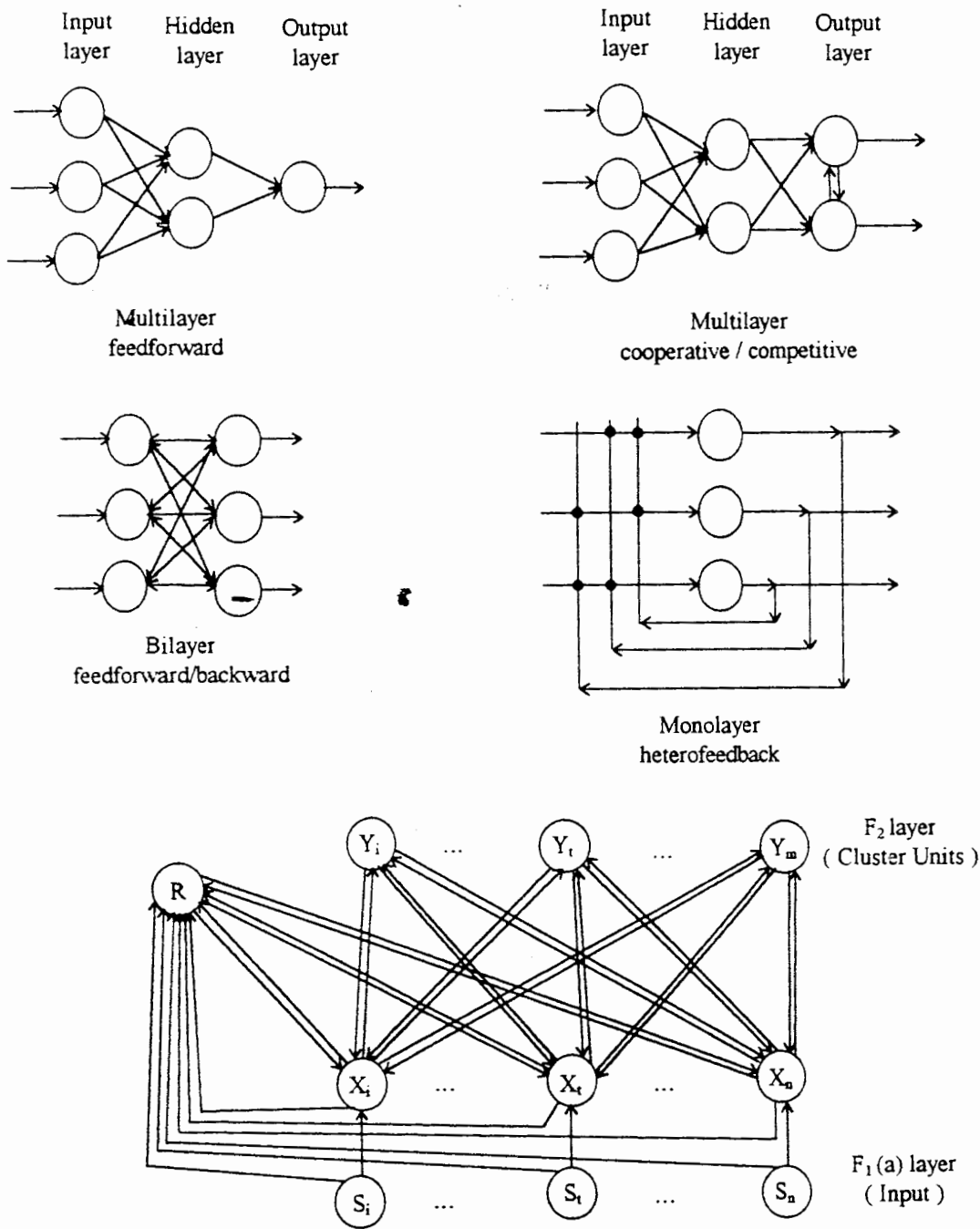
(c) نمایش بایاس در یک نرون



شکل ۵-۱: مقایسه نرون مصنوعی با نرون بیولوژیکی



شکل ۶-۱: شیوه عملکرد یک نرون مصنوعی



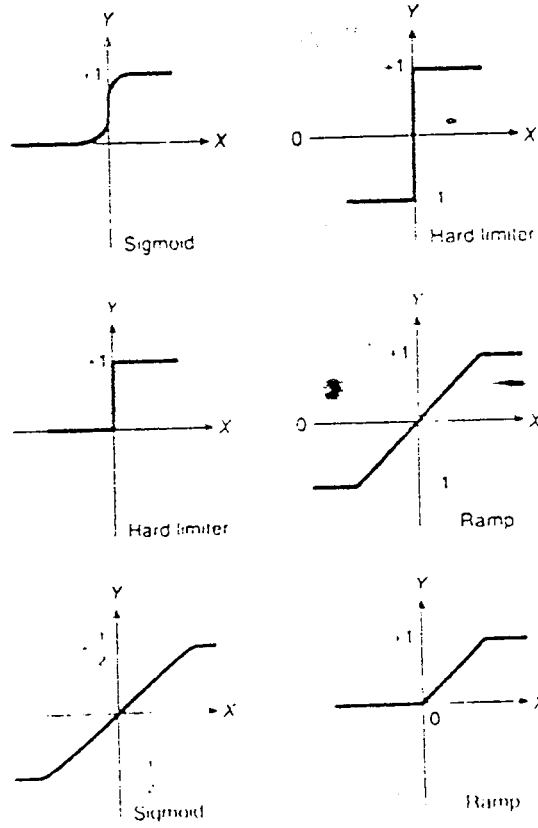
شکل ۸-۱: نمونه‌های مختلف شبکه‌های عصبی مصنوعی

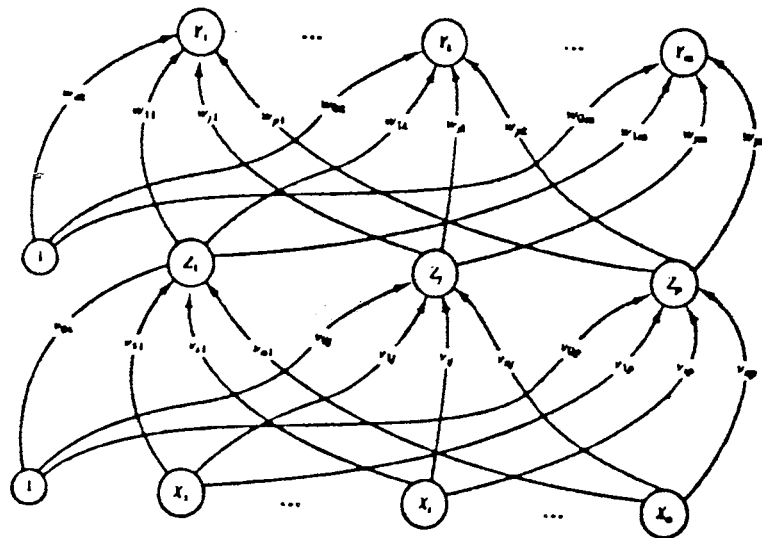
بالا و از چپ به راست: شبکه چند لایه پیشرو، چند لایه یارانه‌ای-رقابتی

وسط و از چپ به راست: شبکه دو لایه (پیشرو و پسرو)، شبکه تک لایه دیگر پس‌خورد

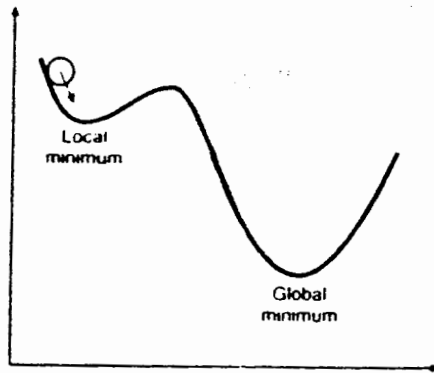
پایین: شبکه ART 1

شکل ۱-۷: انواع توابع غیرخطی مورد استفاده در نرونها

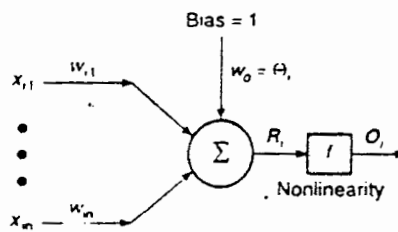




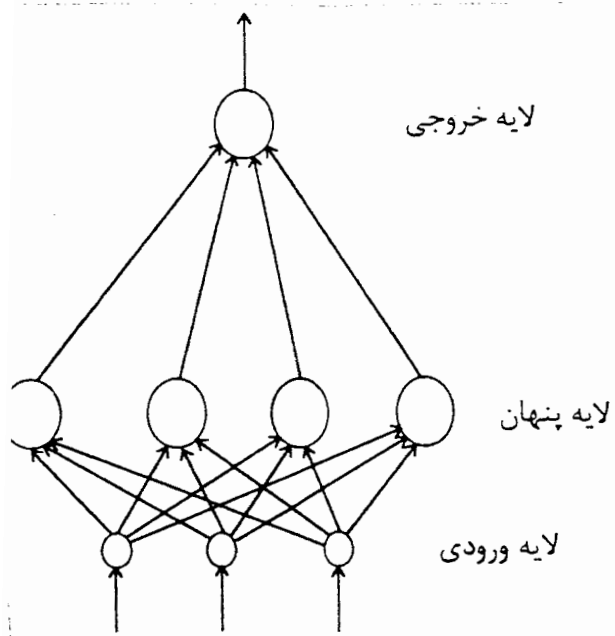
شکل ۹-۱: شبکه پیشخورد، دو لایه با لایه پنهان



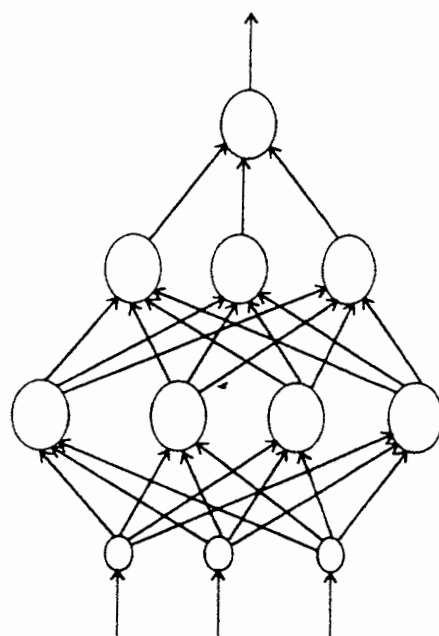
شکل ۱-۱۰: نمایش حداقل موضعی و حداقل مطلق در روند آموزش شبکه



شکل ۱-۱۱: مدل نرون مصنوعی McCulloch و Pitts

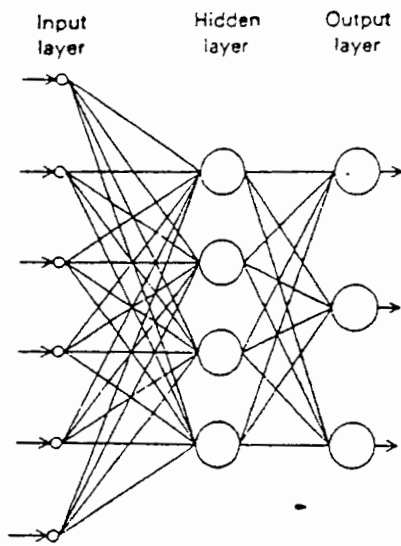
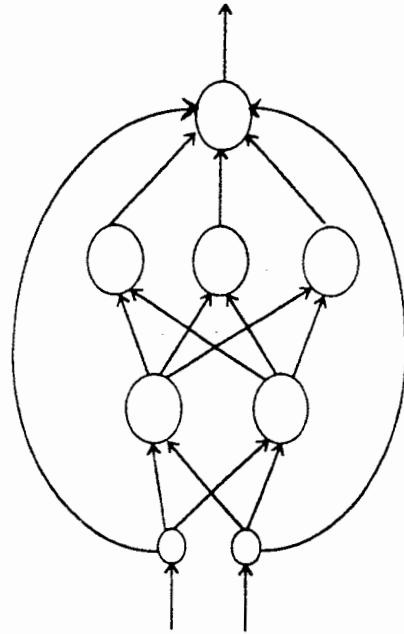


شکل ۱-۱۲: پرسپترون دو لایه

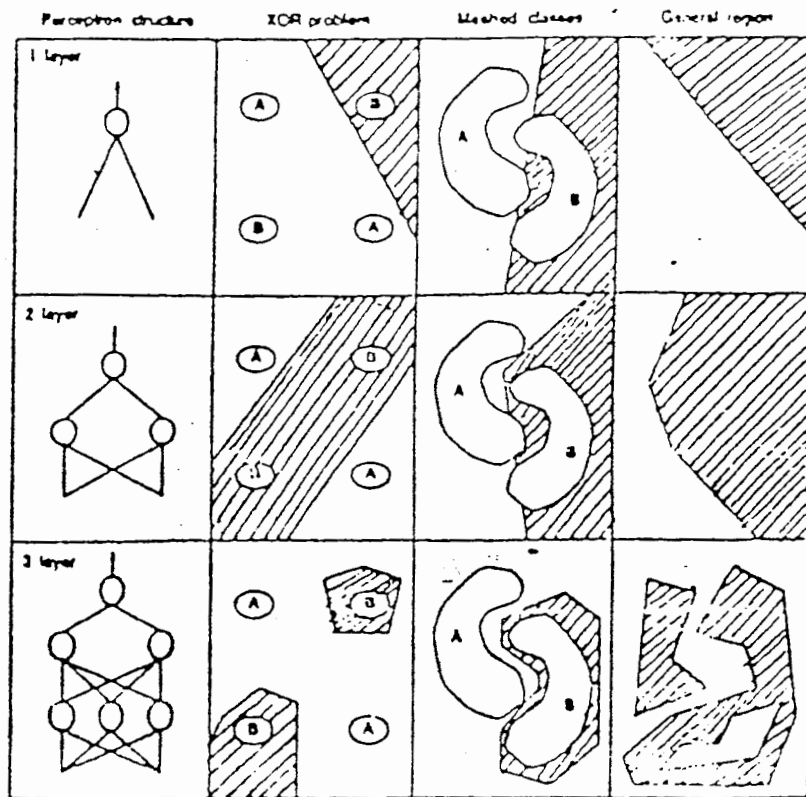


شکل ۱-۱۳: پرسپترون چند لایه (سه لایه)

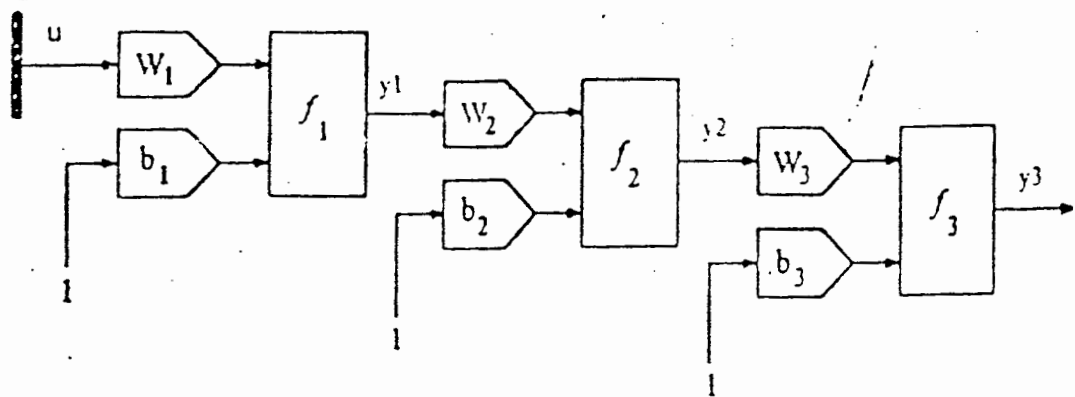
شکل ۱-۱۴: ساختار یک پرسپترون بدون لایه با اتصالات میان بر



شکل ۱-۱۵: شبکه پرسپترون سه لایه با یک لایه مخفی



شکل ۱-۱۶: دسته‌بندی الگوها بوسیله شبکه‌های عصبی چند لایه



$$u: R \times 1$$

$$y_1, \underline{b}_1: s_1 \times 1$$

$$W_1: s_1 \times R$$

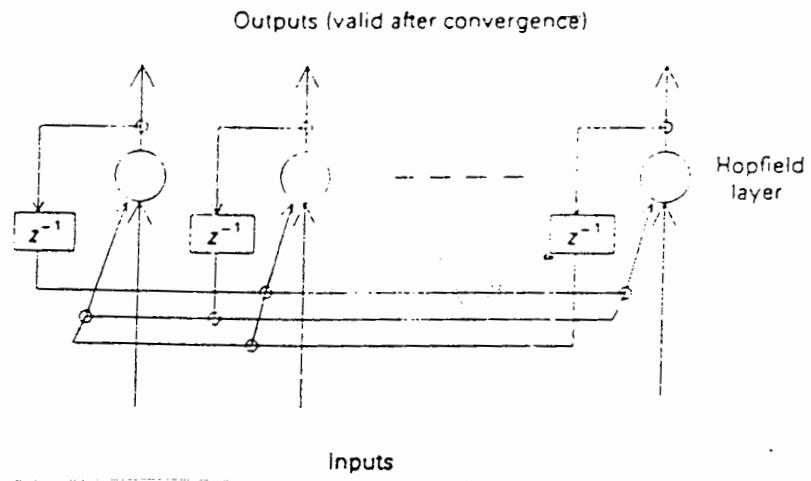
$$y_2, \underline{b}_2: s_2 \times 1$$

$$W_2: s_2 \times s_1$$

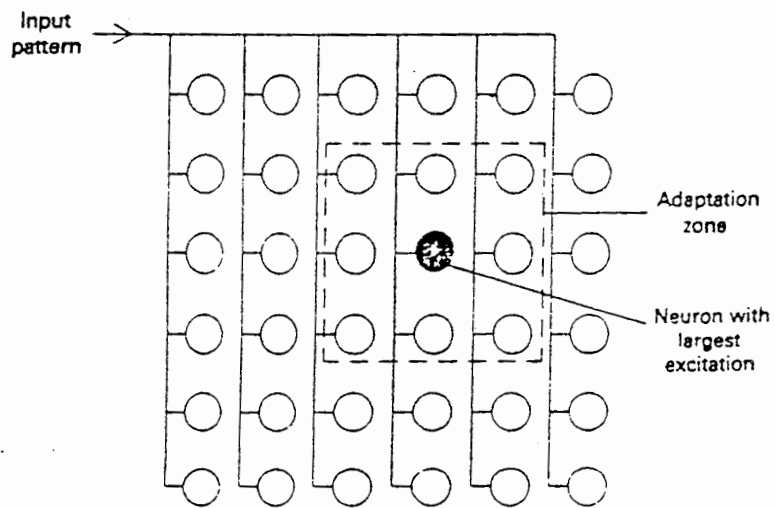
$$y_3, \underline{b}_3: s_3 \times 1$$

$$W_3: s_3 \times s_2$$

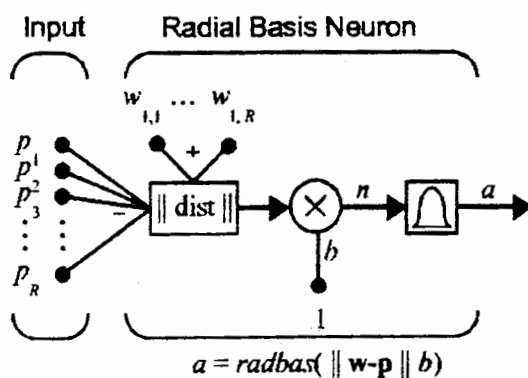
شکل ۱-۱۸: بلوک دیاگرام شبکه پرسپترون سه لایه



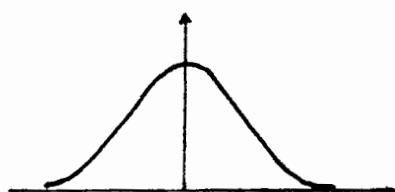
شکل ۱-۲۰: شبکه هاپفیلد



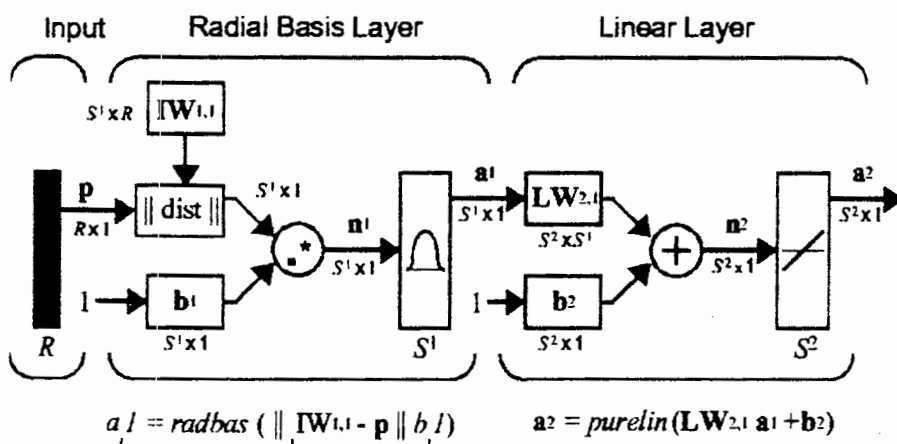
شکل ۱-۲۱: شبکه کوهنن



شکل ۱-۲۲: نرون بر اساس تابع پایه شعالی



شکل ۱-۲۳: تابع پایه شعالی



Where...
 R = # elements in input vector
 S_1 = # Neurons in layer 1
 S_2 = # Neurons in layer 2

شکل ۱-۲۴: ساختار کلی شبکه RBF

فصل دوم

کاربرد شبکه‌های عصبی در ژئوفیزیک

۱-۲- مقدمه

شبکه‌های عصبی در دهه گذشته هر چه بیشتر در ژئوفیزیک فراگیر شده‌اند. آنها برای انواع مختلفی از مسائل بطور موفق‌تری بکار برده شده‌اند. در محدوده ژئوفیزیکی شبکه‌های عصبی برای تشخیص نوع موج و اولین توقف First-break picking توسط مارت و رادمن (Maurt & Rudman 1992) و مک کورماک و دیگران (McCormac 1993) بکار برده شده‌اند. برای الکترومغناطیس نیز توسط پل‌تون و دیگران (Poulton) در ۱۹۹۲ برای مگنتوتلوریک توسط ژانگ و پاولسون (Zhang & Paulson) در ۱۹۹۷ و برای مقاصد مدل‌سازی معکوس در لرزه‌شناسی توسط رات و تارنتولا در ۱۹۹۴ (Roth & Tarantola) و لانگر و دیگران (Langer) در ۱۹۹۶ کالدرون و ماسیاس و دیگران (Calderon - Macias) در ۱۹۹۸، برای تجزیه امواج برشی توسط دایی و مک‌بث (Dai & McBeth) در ۱۹۹۴. برای تحلیل‌های چاه‌نگاری توسط هوانگ و دیگران (Huang) در ۱۹۹۶، برای ویرایش اثرهای جزئی توسط مک‌کورماک و دیگران در ۱۹۹۳ برای واپیچش لرزه‌شناسی توسط وانگ و مندل (Wang & Mendal) در ۱۹۹۲، کالدرون و ماسیاس در ۱۹۹۷ و برای دسته‌بندی رویدادها (دوولا و دیگران (Dowla) در ۱۹۹۰ و رومئو (Romeo) در ۱۹۹۴ و برای تعداد دیگری از مسائل بکار رفته است.

با وجود این بیشتر این کاربردها توسط روشهایی که امروزه رایج است استفاده نشده است که برای استفاده آسان بکار رود. از اینرو عباراتی مانند "طراحی و پیاده‌سازی شبکه هنوز بیشتر هنر است تا علم" هنوز نادر نمی‌باشند. هدف این مقاله فراهم کردن مقدمه‌ای کوتاه بر این تکنیکها و روشها است. برای گرفتن اطلاعات کامل که تمام محدوده انواع شبکه‌های عصبی را پوشاند به تحقیقات بسیار عالی لیپمن (Lippmann) ۱۹۸۷ و هاش و هورن (Hash & Horne) ۱۹۹۳ هرات و جوتن (Herant & Hutten) ۱۹۹۴ و چنتف (Chentouf 1997) مراجعه کنید.

عبارت "طراحی و پیاده‌سازی شبکه هنوز بیشتر هنر است تا علم" اساساً قابل استناد به مشکلات شناخته شده در رابطه با شبکه‌های عصبی است. در میان اینها، مسأله تعیین شکل شبکه بهینه (به طور مثال، ساختار آن)، توزیع وزن بهینه یک شبکه خاص، و تضمین یک کارایی رویهمرفته خوب (بطور مثال، عمومیت دادن خوب) بسیار برجسته هستند. در اینجا تکنیکهایی برای دسته‌بندی این مسائل و مشکلات شناخته شده عنوان شده است.

شبکه‌های عصبی انواع مختلفی دارند. بعضی از آنها در مسائل ژئوفیزیکی بکار می‌روند. بهر حال ما در اینجا به شناخت شبکه‌های بدون پسخور و ایستا می‌پردازیم. کلمه ایستا این مسئله را می‌رساند که وزنها، هنگامی که یکبار تعریف شدند، ایستا باقی مانده و با زمان تغییر نمی‌کنند، کلمه بدون پسخور نیز نشان می‌دهد که خروجی دوباره مورد استفاده قرار نمی‌گیرد و دوباره به عنوان ورودی به شبکه داده نمی‌شود. همچنین این نوع از شبکه به یک راه حل نهایی تکرار نمی‌شود اما مستقیماً سیگنال ورودی را به یک خروجی مستقل از ورودی قبلی تبدیل می‌کند.

بعلاوه، تنها شبکه‌های عصبی با ناظر مدنظر گرفته شده اند، به ویژه آنهایی که برای دسته بندی مسائل مناسب‌ترند. علی‌رغم این، انواع مشابهی از شبکه‌های عصبی می‌توانند برای تقریب توابع و مسائل معکوس مورد استفاده قرار گیرند (پلتون و دیگران (Poulton 1992) و روت و تارانتولا (Roth & Tarantola 1994)). تقسیم‌بندی با ناظر عمدتاً شامل سه مرحله متفاوت می‌باشد (ریچارد ۱۹۹۳ (Richard)). انتخاب، یادگیری یا آموزش، و دسته‌بندی. در اولین مرحله، عدد و طبیعت رده‌های مختلف معرفی شده و مثالهای نماینده برای هر رده انتخاب می‌شوند. در مرحله یادگیری، ویژگی‌های هر رده منحصر به فرد بایستی از نمونه‌های آموزشی استخراج شود. در نهایت همه داده‌ها می‌توانند با استفاده از این ویژگیها دسته‌بندی شوند.

با وجود این، چند شبکه جالب دیگر وجود دارند، که متاسفانه فراتر از افق این مقاله قرار دارند. که آنها شامل نگاشت خود-سازمانی کوهونن (Kohonen) ۱۹۸۹ و تئوری رزونانس (تشدید) تطبیقی کارپنتر و گروسبرگ (Carpenter & Grossberg) ۱۹۸۷ و شبکه هوفیلد (Hopfield) ۱۹۸۴ و شبکه‌های دیگر رایج می‌شوند. برای مطالعه بیشتر به لیپمان ۱۹۸۷ (Lippmann) و هوش و هورن ۱۹۹۳ (Hush & Horne) "برای طبقه بندی جزئی" مراجعه شود.

۲-۲- شبکه‌های عصبی: ساختار و رفتار

شکل ۲-۱ یک لایه نرون منفرد را نشان می‌دهند. چنین شبکه‌ای می‌تواند داده‌ها را به دو رده تقسیم‌بندی کند. برای یک ورودی دوبعدی، دو توزیع با یک خط جدا شده‌اند (شکل ۲-۲b). بطور عمومی، دو دسته بوسیله یک صفحه مافوق $(n-1)$ بعدی برای یک ورودی n بعدی جدا می‌شوند.

توزیع‌های پیچیده بیشتری می‌تواند بکار رود اگر یک لایه میانی از نرونها اضافه شود. چنین لایه‌هایی در بین لایه‌های ورودی و خروجی قرار می‌گیرند که بطور غیر مستقیم آنها را وصل می‌کند. به هر حال، راه عمومی برای کار کردن آنها بطور کلی تغییر نمی‌کند همانطوری که در شکل (۲-۳ a و b) نشان داده شده است. دوباره هر نرون در لایه میانی، فضای ورودی را به دو نیم فضا تقسیم می‌کند. در نهایت آخرین نرون تمام اینها را به فرم یک شکل یا یک زیرفضای بسته ترکیب می‌کند. با افزودن دومین لایه میانی اشکال کاملاً پیچیده‌ای می‌تواند تشکیل شود.

استفاده کردن از یک تابع سیگموئید بجای یک تابع محدود سخت، تغییری در شکل کلی بدست نمی‌دهد. تغییرات بین رده‌ها هموارتر می‌شود. یا به عبارت دیگر، استفاده از تابع فعال ساز گاوسی تغییرات عمده را نمایش می‌دهد، زمانی که یک پاسخ محلی وجود دارد. به هر حال فضای نمونه به دو بخش تقسیم می‌شود. بخش نزدیک به میانه گاوسی با خروجی‌های بزرگ، توسط زیرفضا در دنباله‌های آن برای نشان دادن مقادیر خروجی کوچک احاطه شده است. بنابراین یک نرون منفرد با تابع فعال ساز گاوسی و واریانس ثابت برای توصیف رده خاکستری در شکل (۲-۳) بجای بکارگیری سه نرون با تابع پالس یا توابع فعال ساز سیگموئیدی نیاز است. بعلاوه تابع گاوسی

یک دایره کامل را دور رده در وسط ایجاد خواهد کرد (اگر یک واریانس عمومی برای تمام پارامترهای ورودی استفاده شود).

این بینش به روش عمومی شبکه‌های عصبی، مسائل مهم دسته‌بندی را حل کرده و کاربرد را برای بدست آوردن یک تصور اولیه از ساختار لازم برای یک کاربرد خاص توانا می‌سازد. در حالی که مسائل بسیار پیچیده باشند، که گفته می‌شوند، نافرما، توزیع‌های چندگانه، احتمالاً با انتخاب ساختار شبکه عصبی با دو لایه میانی قابل حل خواهند بود. به هر حال، سینکو ۱۹۸۹ (C ybenko) نشان داد که شبکه‌های عصبی که از سیگموئیدها استفاده می‌کنند می‌توانند هر تابع پیوسته‌ای را با دقت زیادی تقریب بزنند و فقط یک لایه میانی یک لایه غیرخطی و دو واحد خروجی داشته باشند. بطور مشابه پارک و سندبرگ ۱۹۹۱ (Park & Sandberg) نشان دادند که تحت شرایط معتدل، شبکه‌های عصبی با توابع فعال ساز محلی شده (مانند تابع گاوس) همچنین تقریب‌زندهای عمومی هستند. متأسفانه هیچکدام از این فرضیات قادر به پیش‌بینی دقیق تعداد نرونها مورد نیاز تا هنگامی که نتایج خوبی به دست بیاید نیستند. بعلاوه، کاربردهایی وجود دارند که شبکه‌های عصبی با دو لایه میانی نتایج مشابهی را به عنوان یک شبکه عصبی با یک لایه میانی ایجاد می‌کند با کاهش چشمگیری در تعداد اتصالات و، بنابراین، یک مسأله بهینه‌سازی وزنی پیچیده کمتر، بطور مثال، ایجاد آموزش خیلی ساده‌تر (چنتف ۱۹۹۷ (Chentouf)).

دو نوع از توابع فعال ساز در شکل (۲-۱) استفاده شده‌اند. تابع پالس و سیگموئید بطور یکنواخت توابع بالا رونده هستند، در حالیکه تابع گاوس یک عملگر محلی دارد. هر دو نوع در کارهای شبکه‌های عصبی کاربرد دارند. بطور عمومی، شبکه‌های عصبی با افزایش یکنواخت توابع فعال‌ساز را MLP یا Multilayer Perceptrons (پرسپترون چند لایه) می‌نامند و شبکه‌های عصبی با توابع فعال‌ساز محلی را توابع پایه شعاعی یا RBF یا Radial Basis Functions می‌نامند (جدول ۲-۱).

از اینرو، شبکه‌های MLP با یک پرسپترون خروجی و یک لایه میانی منفرد با فرمول زیر بیان می‌شوند

$$f_{MLP}(X) = \sigma \sum_{k=1}^{n_{h1}} w_k \sigma(W^{(k)} \cdot X - \theta^{(k)}) - \theta \quad (۲-۱)$$

$\sigma(\cdot)$ تابع فعال ساز سیگموئیدی است، X ورودی، w_k وزن اتصال k به ند خروجی، n_{h1} شماره نودها در لایه میانی، w^k وزن تمام اتصالات به گره k در لایه میانی، و θ جمله بایاس است. عناصری که با حروف درشت چاپ شده اند علامت بردار را نشان می دهند. فرمول (۱-۲) می تواند به راحتی به چندین ند خروجی و لایه های میانی بیشتری بسط داده شود. همچنین شبکه های RBF با یک لایه میانی منفرد و یک پرسپترون خروجی توسط فرمول زیر بیان می شوند:

$$f_{RBF}(X) = \sigma \left(\sum_{k=1}^{n_{h1}} w_k K \left(s_k \|X - C^{(k)}\| \right) - \theta \right) \quad \text{فرمول (۲-۲)}$$

$K(\cdot)$ تابع محرک (فعال ساز) محلی، $\| \cdot \|$ یک نرم (فاصله) می باشد، $C^{(k)}$ مرکز تابع محرک محلی در گره لایه میانی k می باشد، و S_k پهنای وابسته آن است.

پرهیز از تعداد کل (n_{tot}) متغیرهای درونی که رفتار ساختار شبکه های عصبی استفاده شده را تعیین می کنند مهم است، همانطوری که بعداً نشان می دهیم. خوشبختانه این عدد به راحتی از این معادلات (۱-۲) و (۲-۲) بدست می آید. برای شبکه های MLP، آن ترکیبی است از تعداد اتصالات بعلاوه تعداد پرسپترونها برای ترکیب تعداد بایاس ها. اگر n_i نشانگر تعداد متغیرهای ورودی باشد، n_{in} تعداد پرسپترونها در (i) امین لایه میانی و n_0 تعداد پرسپترونهای خروجی باشد، پس n_{tot} بوسیله فرمول زیر حساب می شود

$$n_{tot} = (n_i + 1) * n_{in} + (n_{in} + 1) * n_0 \quad \text{فرمول (۳-۲)}$$

برای یک MLP با یک لایه میانی منفرد و

$$n_{tot} = (n_i + 1) * n_{h1} + (n_{h1} + 1) * n_{h2} + (n_{h2} + 1) * n_0 \quad \text{فرمول (۴-۲)}$$

برای یک MLP با دو لایه میانی می باشد. تعداد متغیرهای داخلی دقیقاً برابر با شبکه های RBF همسانگرد (یکنواخت) می باشد، زمانی که هر تابع گاوسی بوسیله متغیرهای $n_i + 1$ برای موقعیت و واریانس خودش، مانند عرض، تعریف شود. در اینجا تنها شبکه های RBF با یک لایه میانی منفرد فرض شده اند. بعلاوه، نرون های RBF تنها در لایه میانی توابع محرک گاوسی دارند. نرونهای خروجی، سیگموئیدها را بعنوان توابع محرک دارند. از اینرو n_{tot} توسط فرمول (۳-۲) بدست می آید.

چنانکه خواهیم دید، نسبت $\frac{n_{tot}}{m}$ مقدار مناسب بهینه شدن یک شبکه را تعیین می‌کند و در آن m تعداد نمونه‌های مورد آموزش می‌باشد.

۲-۳- بهینه‌سازی شبکه

۲-۳-۱- مسائل شناخته شده

دو مرحله مهم در بکاربردن شبکه‌های عصبی برای شناختن مسائل، انتخاب و مراحل آموزشی هستند، از اینرو اینها به طور مستقیم بر کل اجرا و بنابراین نتایج بدست آمده تاثیر می‌گذارد. سه دلیل باعث بازده بد خواهند شد (رومئو، ۱۹۹۴ (Romeo)): یک شبکه‌بندی نامناسب، الگوریتم آموزشی که درون یک مینیمم محلی گیر افتد، یا یک مجموعه آموزشی نامناسب.

حال در اینجا با نحوه شبکه‌بندی آشنا می‌شویم. همانطوری که در شکل‌های (۲-۲) و (۳-۲) نشان داده شده است، شیوه شبکه‌بندی باید به ما امکان یک تعریف مناسب از توزیع آماری از گسترش در داده‌ها را بدهد. وقتی که تعداد ورودی و خروجی‌های نرون‌ها (سلولهای عصبی) در چندین کاربرد ثابت شده‌اند، توجه ما به تعداد لایه‌های پنهان و تعداد سلولهای درون آن می‌باشد. هیچ قانونی وجود ندارد که بتواند تعداد دقیق سلولهای عصبی را در لایه پنهان معین کند. به هر حال هوانگ و هوانگ (1991) (huang & huange) نشان دادند که حد بالای تعداد سلولهای عصبی مورد نیاز برای بازسازی دقیق خروجی دلخواه از نمونه‌های آموزشی به تعداد m می‌باشد، تعداد نمونه‌های آموزشی. بنابراین تعداد سلولهای عصبی در لایه پنهان هرگز نباید بیش از تعداد نمونه‌های آموزشی شوند. بعلاوه، برای نگهداری مشکل آموزشی تحمیل شده، تعداد نمونه‌های آموزشی همیشه باید بزرگتر از تعداد وزن‌های داخلی باشد. در عمل، $m \approx 10n_{tot}$ یک انتخاب خوب فرض می‌شود. همچنین تعداد نرون‌ها بایستی محدود باشد؛ وگرنه این خطر وجود دارد که مجموعه مورد آموزشی به سادگی بوسیله شبکه به خاطر سپرده شود (Overfitting). به طور کلاسیک، بهترین آرایش بوسیله آزمون و خطا یافت می‌شود، که با یک عدد کوچک از ندها شروع می‌شود.

دلیل دوم اینکه ممکن است شبکه جواب مورد نظر را ندهد این است که ممکن است پاسخ در یک مینیمم محلی به دام افتاده باشد. تابع غیرمتجانس (misfit) اغلب بسیار پیچیده می‌باشد

(هاش و دیگران ۱۹۹۲ (Hush)). بنابراین شبکه به سادگی می‌تواند بجای آنکه به یک مینیمم مطلق دست یابد، در یک مینیمم محلی اسیر شود. که در آن مورد حتی مجموعه آموزشی نمی‌تواند به طور مناسب جور باشد.

راه حل‌ها ساده هستند. هر یک از چندین تلاشهای نقطه‌های کمینه (مینیمم) بایستی انجام شود، هر بار با استفاده از یک (تصادفی یا غیر تصادفی) مقادیر وزن اولیه، یا الگوریتم‌های معکوس دیگری بایستی فرض شوند، مانند جستجوی عمومی.

در نهایت، مسأله می‌تواند با مجموعه آموزشی انتخاب شده انجام شود. دو مسأله‌ای که غالباً پیش می‌آیند یکی آموزش بیش از حد و یک آموزش بد است، بطور مثال نمایش ندادن مجموعه یادگیری. در مورد آخر، هر یک از، تعداد بدی از نمونه‌ها انتخاب شوند (بطور مثال، الگوهای نسبت داده شده به دسته غلط) یا مجموعه آموزشی اجازه یک تعمیم‌سازی خوب را ندهد، برای مثال، فضای عمومی کامل نباشد، بطور مثل، نمونه‌هایی که برای یک آموزش کافی شبکه نیاز است به سادگی ناپیدا باشد.

آموزش بیش از حد مجموعه یادگیری مورد مطالعه نیز ممکن است مشکلاتی به همراه داشته باشد. آزمایش بیش از حد یعنی آنکه فرض کنیم مجموعه‌های آموزشی که در حافظه بارگذاری شده است برای این مجموعه عالی‌ترین می‌باشد و هیچ داده دیگری قابل قبول نمی‌باشد. برای حل این مشکل، مجموعه انتخاب شده از مثالها اغلب به یک مجموعه آموزشی و یک مجموعه معتبر تفکیک می‌شوند. وزن‌ها با استفاده از مجموعه آموزشی بهینه می‌شوند. به هر حال، اعتبار متقابل با مجموعه دوم، یک اجرای خوب را تضمین می‌کند.

در بخشهای بعدی تمام این مسائل که عنوان شد با جزئیات آنها مورد توجه قرار می‌گیرد و چند روش برای استفاده آسان در آنها نیز توضیح داده می‌شود.

۲-۳-۲- تخمین آموزش/وزن شبکه: یک مسأله بهینه‌سازی

اگر یک پیکربندی از شبکه انتخاب شود، یک توزیع وزنی بهینه بایستی تخمین زده شود. این یک مسأله معکوس‌سازی یا بهینه‌سازی است. یکی از روشهای بسیار متداول معکوس‌سازی متمرکز است. در چنین روشی، ما ابتدا فرض می‌کنیم که خروجی y می‌تواند از ورودی x با استفاده

از برخی انواع تابع f محاسبه شود، بطور مثال $y=f(s)$. خروجی ممکن است با نویز مخلوط شده باشد، که ناهمبسته با داده‌ها و میانگین صفر فرض شده است. سپس ما فرض می‌کنیم که تابع می‌تواند در حوالی برخی تخمین‌های اولیه x_0 از بردار ورودی x با استفاده از یک بسط تیلور مرتبه یک خطی باشد. به طور مثال،

$$\text{فرمول (۵-۲)} \quad y = f(x_0) + \frac{\partial f(x_0)}{\partial x} \Delta x$$

اگر بنویسیم $y_0=f(x_0)$ ، $\Delta y=y - y_0$ و $A^{(x)}=\partial f/\partial x$ ، معادله ۵ بصورت زیر در خواهد آمد:

$$\text{فرمول (۶-۲)} \quad \Delta y = A^{(x)} \Delta x$$

که ژاکوبین $A^{(x)} = \nabla_x f$ شامل جزء اول مشتقات نسبت به x می‌شود. برای ترسیم یک تناسب با یک مسأله وارونه سازی، در یک کاربرد توموگرافی (Tomography) y شامل زمان سیر مشاهده شده، x مدل آرام مطلوب، و A_{ij} طول مسیر پرتو i ام در سلول j ام است.

به هر حال، یک تفاوت زیربنایی در مسائل توموگرافی وجود دارد. در یک کاربرد شبکه‌های عصبی هم خروجی y و هم ورودی x شناخته شده هستند، بصورتی که $y^{(i)}$ یک خروجی مطلوب برای مجموعه آموزشی $x^{(i)}$ است. از اینرو، مسأله ساختاری از یک مدل x نمی‌سازد تا مشاهدات را توضیح دهد، بلکه ساختاری از تابع تقریب f می‌سازد. وقتی که این تابع بوسیله متغیرهای داخلی‌اش تعریف می‌شود، یک سیستم خطی دیگر خواهد بود که بایستی حل شود از قرار زیر:

$$\text{معادله (۷-۲)} \quad \Delta y = A^{(w)} \Delta w$$

که ژاکوبین $A^{(w)} = \nabla_w f$ شامل جزء اول مشتقات نسبت به مقادیر داخلی w می‌شود. بردار w شامل پایه‌ها و وزن‌ها برای شبکه‌های MLP می‌باشد و شامل وزن‌ها، واریانس‌ها، و مراکز برای شبکه‌های RBF می‌باشد. برای بیان دقیق $A^{(w)}$ ، بایستی به هاش و هورن (Hush & Horne 1993) و هرالت و جونت (Heranuh & Junnten 1994) مراجعه نمود. علیرغم آنکه تمام عبارات می‌توانند بصورت تحلیلی محاسبه شوند. بعلاوه، هم سیگموئید و هم تابع گوسی به طور پیوسته مشتق‌پذیر هستند، که آخرین دلیل برای استفاده از آنها می‌باشد. بنابراین، هیچ تئوری انحراف یا آشفتگی مرتبه اولی نباید برای بدست آوردن تخمینات مشتق جزئی مطلوب بکار برده شود، و به این خاطر بهره زیادی را در زمان محاسبه برای شبکه‌های عصبی بزرگ خواهد داشت.

بطور عمومی مسائل بهینه‌سازی دارای حالت ill-posed می‌باشند به طوری که $A^{(w)}$ از یک مرتبه ناقص مشکل دارد. بطور مثال مرتبه $n_{tot} \leq A^{(w)}$. بنابراین، سیستم معادله (۷-۲) مورد شک می‌باشد. به هر حال، در یک زمان مشابه، هر مسأله معکوس که بطور صحیح فرمول‌بندی شده باشد مورد مطالعه قرار خواهد گرفت زیرا $n_{tot} \gg m$ ، با توجه به آنکه نمونه‌های آموزشی بیشتر از متغیرهای داخلی هستند.

از آنجایی که سیستم معادله (۷-۲) دارای حالت ill posed است، یک فضای تهی وجود خواهد داشت. در هر صورت متغیرهای داخلی بنا شیوه واحدی نمی‌توانند بدست بیایند. اگر بعلاوه $m \ll n_{tot}$ باشد، خطر آموزش بیش از حد وجود دارد، بطور مثال معرف حافظه بالا رفته و بازده نتایج پایین‌تر بدست خواهد آمد. دو دلیل باعث کاهش رتبه A می‌باشند. اول آنکه فضای نمونه کافی نباشد، بطور مثال بعضی نمونه‌ها که نیاز به یک دقت بالا دارند، گم باشند و بعضی نمونه‌های آموزشی بطور غلط برای یک کلاس دیگر به کار روند. دومین دلیل آلودگی به پارازیت است که از یک تطبیق مناسب داده‌های تهیه شده و تهیه نشده جلوگیری می‌کند. برای مثل، در یک مسأله توموگرافی کاهش رتبه اتفاق می‌افتد اگر سلولهای مشاهده نشده ارائه شوند، برای این سلولها یک تخمین دقیق از سرعت غیر ممکن خواهد بود.

برای بدست آوردن یک ذهنیت کلی برای تعداد نمونه‌های لازمه یک مقاله با مطالعه تئوری توسط باوم و هوسلر (۱۹۸۹) (Baum & Haussler) نشان می‌دهد که برای یک سطح دقیق مطلوب از $(1-\epsilon)$ ، حداقل n_{tot}/ϵ نمونه نیاز داریم که باید تهیه شود، بطور مثال، $m \geq n_{tot}/\epsilon$ همچنین برای رده‌بندی ۹۰ درصد داده‌ها بطور صحیح، حداقل ۱۰ برابر نمونه‌های بیشتر از متغیرهای داخلی که ارائه شده‌اند باید تهیه شوند، بطور مثال، $m \geq 10 n_{tot}$.

حال باید پرسید که چگونه می‌توان معادله (۷-۲) را حل نمود؟ یکی از راههای ممکن برای بهینه کردن w کمینه کردن مجموع مربعات تفاضلات بین خروجی واقعی و مطلوب شبکه می‌باشد. این عمل ما را به راه حل کمترین مجموع مربعات می‌رساند و بردار وزنی بوسیله معادله عادی زیر بدست خواهد آمد.

$$\Delta w = (A' A)^{-1} A' \Delta y \quad \text{معادله (۸-۲)}$$

در اینجا w حذف شده است.

به هر حال این روش مسائل مبهم شناخته شده زیادی دارد که یکنواختی در ΔA باعث واگرایی نرم $|\Delta w|$ وزنی می‌شود، وقتی که این فرم معکوس کمترین مقدار واحد A باشد. بعلاوه اگر A دارای کاهش رتبه باشد، مقدار واحد برابر یا نزدیک به صفر خواهد بود به خاطر وقت رقمی ماشینی. نرم توان دوم $|\Delta w|^2$ واریانس حل نامیده می‌شود.

برای پیشگیری از واگرایی جوابها یک راه ساخته شده از معادله (۲-۸) استفاده می‌شود که یک متغیر به نام β که به آن متغیر مثبت کاهنده می‌گویند، استفاده می‌شود. این متد به نام روش لونبرگ مارگوارت (Levenberg & Marquart) به تصحیح تیخونوف (Tikhonov) سیستم جایگزین برابر خواهد بود با:

$$\Delta w = (A' A + \beta I)^{-1} A' \Delta y \quad \text{معادله (۲-۹)}$$

با ماتریس شناسایی I (لینس و ترتیل ۱۹۸۴ Lines & Treitel وان در اسلویز و واندروریت ۱۹۸۷ (Van & der sluis & Van der vorst).

در مقایسه با ماتریس ΔA ماتریس $\Delta A + \beta I$ دارای کاهش رتبه نمی‌باشد و در هر صورت واریانس حل واگرا نمی‌شود اما ممکن است پاسخ بدتری بدهد. برخلاف روشی که در اینجا می‌آید، راه حل ممکن است به یک طرف متمایل شود، به علت پارامتر مرتب سازی β . به هر حال این راه حل در روش کمترین مجموع مربعات بهینه‌ای را ارائه نمی‌کند. مقدار دقیق β بایستی به دقت برای بهینه‌کردن بین واریانس و bias بهینه شود (به مقالات وان در اسلویز و واندروریت ۱۹۸۷ Van der sluis & van deer vorst و گمان و دیگران ۱۹۹۲ Geman نگاه کنید).

تعدیل سازی پیچیده‌تری می‌تواند هم بر Δw و هم بر Δy اعمال شود. برای مثال، اگر محدوده خروجی نامطمئن باشد (و یا مثلاً واریانس آنها)، آنها برای دوباره مقیاس کردن خروجی بکار می‌روند. مقیاس کردن مشابهی برای ورودی و یا بردار وزنی می‌تواند بکار رود. این روش به ما اجازه برای ترکیب تمام اطلاعات قبلی را می‌دهد. به هر صورت یک بایسان پیچیده در مسأله متن می‌تواند فرموله شود. برای جزئیات این روش به تارانتولا ۱۹۸۷ (Tarantola) مراجعه کنید.

مسائل توموگرافی (Tomography) مانند معادلات (۲-۸) و (۲-۹) به ندرت مستقیماً حل می‌شوند. بیشتر از روشهای تکرار برای حل آنها استفاده می‌شود. بهترین روش شناخته شده در کاربردهای شبکه‌های عصبی روش بازگشت تدریجی رامل هارت و دیگران ۱۹۸۶ (Ramerlhart et al) می‌باشد یا بدون مومنتوم، یک واژه که در مورد توابعی با فاکتور هموارسازی β بکار می‌رود.

همچنین این روش به عنوان روش درجه اول بهینه‌سازی با تقریب زدن $(\dot{A}A)^{-1}$ در معادله (۲-۸) و (۲-۹) بکار می‌رود با $\alpha=0$ و $\beta=0$.

این متد بصورت پایه یک الگوریتم تدریجی افت می‌باشد. به هر صورت تمام نتایج ناخوش‌آیند روشهای الگوریتم تدریجی افت بکار برده شده‌اند. برای مثال در حالتی که یک نمودار به غلط تطبیق داده شده است، گرادیان همواره به نقطهٔ مینیمم مطلق اشاره نمی‌کند، همچنین همگرایی ممکن است کند باشد (رجوع کنید به لینس و ترتیل ۱۸۹۴ Leines & Treitel). برای سرعت بخشیدن به همگرایی گرادیان محاسبه شده بایستی در یک فاکتور ثابت α ضرب شود که $0 < \alpha < 2$ می‌باشد. به هر حال یک انتخاب درست از α مورد نیاز است. وقتی که انتخاب بهینه نباشد اثر آن کاملاً برعکس خواهد شد و همگرایی کندتر می‌شود. برای مثال اگر α زیاد بزرگ باشد، تطبیق غلط در بازهٔ بیشتری بازی خواهد کرد که به مقدار کمینه همگرا نمی‌شود. انتخاب مقدار زیاد کوچک نیز همگرایی را بسیار کند می‌کند و ممکن است ما را در دام یک مقدار کمینهٔ محلی گرفتار کند و همچنین همگرایی با تعداد مشخصی از دوره‌های چرخش ممکن است انجام نشود. بعلاوه بهتر شدن تطبیق در یک حالت ممکن است به چرخش بعدی کاملاً ما را از حقیقت دور کند. همچنین چندین نکتهٔ مهم در رابطه با ضریب تطبیق α (دال ۱۹۸۷ Dahl و ژاکوبز ۱۹۸۸ Jacobs، و رید میلر و براون ۱۹۹۳ Reil miller & Braun) و جمله‌های پیچیدهٔ منظم‌سازی (هانسون و پرات ۱۹۸۹ Hanson & Pratt، ویگند و دیگران ۱۹۹۱ Weigend et al، ویلیامز ۱۹۹۵ Williams) بیان شده است، که در آنها الگوریتم اولیه همچنان منطقی مانده است. خوشبختانه الگوریتم‌های دیگر را می‌توان برای حل مسائل معکوس بکار برد. بعنوان یک حقیقت، هر روشی می‌تواند برای حل معادلات عادی (۲-۸) و (۲-۹) استفاده شود، مانند روش‌های گوس نیوتن. همچنین بطور جزئی نیز برای این روش گرادیان همبستهٔ مقیاس شده (Sealed Conjugate Gradient SCG) استفاده می‌شود که اثبات شده برای همگرا شدن به حداقل مقدار کمینه m یا n_{tot} تکرار نیاز دارد و بدون هیچ محاسبهٔ اضافه‌ای مستقیماً $(\dot{A}A + \beta I)^{-1}$ را حساب می‌کند و نمونهٔ قبلی را نیز در حافظه دارد تا زمانی که این گرادیان با تمام گرادیانهای فعلی محاسبه شده همگرا شود (مولر ۱۹۹۳ Moller مسترز ۱۹۹۵ Masters).

بعلاوه در حالتی که خطا به شدت غیر خطی است، برای مثال، چندین مقدار کمینهٔ محلی وجود دارد، هر دو روش ژنتیک الگوریتم و ذوب شبیه‌سازی شده (قوی گلدبرگ ۱۹۸۹ Goldberg)

هرتز و دیگران (Hertz et al ۱۹۹۱؛ ماسترز ۱۹۹۵ Masters) روشهای جایگزین این دو را پیشنهاد می‌کنند و یک تکنیک ادغام شده از این دو روش می‌تواند مورد توجه قرار گیرد (ماسترز ۱۹۹۵ Masters). برای مثال ذوب شبیه‌سازی شده قوی می‌تواند در توزیع خوب وزن اولیه استفاده شود که می‌توان آنها را بوسیله روش SCG بهینه نمود. مروری بر الگوریتم‌های یادگیری که شامل روشهای درجه دوم هستند نیز می‌تواند در کتاب (باتیتی ۱۹۹۴ Battiti) یافت شوند. رید ۱۹۹۳ Reed نیز یک دید کلی نسبت به روشهای هموار سازی دارد.

در آخرین نکته توجه به مقداردهی اولیه به وزن می‌کنیم. معادله ۵ به روشنی نشان می‌دهد که نیاز به شروع با یک حدس زدن خوب از این وزنها می‌باشیم وگرنه ممکن است روش بسیار کند شود و شانس افتادن در یک مقدار کمینه محلی به طور خاص زیاد شود. علیرغم آنکه بیشتر روشها از مقداردهی اولیه اتفاقی استفاده می‌کنند، بطور مثال $w_i \in [-r, r]$. حتی اگر محدوده بهینه‌ای برای r در نظر گرفته شود (برای مثال رجوع کنید به نگوین و ویدرو ۱۹۹۰ N. Guyen & widrow).

همانطوری که خاطر نشان شد یک روش جایگزین، استفاده از یک تصویر آموزشی مطلق اول برای بدست آوردن چندین حدس اولیه خوب برای شروع یک بهینه‌سازی محلی است. به هر صورت چندین روش تئوری در این زمینه گسترش یافته است که خوانندگان علاقمند می‌توانند به مجموعه‌های نگوین و ویدرو (Ngyyen & widrow) ۱۹۹۰ مراجعه کنند که از یک خطی‌سازی بوسیله بخشهایی از خروجی تولید شده از نرونهای لایه‌های میانی استفاده کرده‌اند یا دانوکس و لنگی (Denaoux & Lengelle) ۱۹۹۳ که از مدل اولیه استفاده کردند (نمونه‌ها را انتخاب کردند) برای یک مقداردهی اولیه مناسب و سستی ۱۹۹۰ و Sethi) ۱۹۹۵ که از درخت‌های تصمیم برای یک شبکه عصبی چهار لایه استفاده کرد. روشهای جالب دیگر بوسیله کاروییا و دیگران ۱۹۹۴ (Kerouia et al) از نتایج گالیناری و دیگران ۱۹۹۱ (Gallinari et al) که نشان می‌دهد بین شبکه‌های عصبی خطی و تفکیک کننده یا تحلیل‌های عامل یک تعادل رسمی وجود دارد. آنها شبکه عصبی خود را مقدار دهی اولیه کردند به آن صورت که یک تحلیل بکار گرفته شد و از آنجا بقیه نمونه‌ها را شروع کردند.

تمام این روشهای مقدار دهی اولیه این مطلب را می‌رساند که روشهای خطی توانایی تمام کاربردهای مورد توجه را ندارند، آنها یک نقطه شروع خوب فقط برای شبکه‌های عصبی می‌باشند. در هر صورت یک مقدار دهی خطی بهتر از مقدار دهی اتفاقی برای وزن می‌باشد.

۲-۳-۳- عمومی سازی

حالا ما می‌توانیم یک شبکه را آموزش دهیم. یک سؤال جدید ایجاد خواهد شد و آن اینکه چه زمانی باید آموزش ما متوقف شود؟ به نظر می‌رسد که یک ایده خوب برای توقف آموزش وقتی است که یک مقدار کمیته محلی بدست آید یا وقتی که میزان نسبت همگرایی خیلی کوچک شود و یا مثلاً در رابطه با تعداد تکرارها تعداد تکرارها صفر یا کمیته باشد. به هر حال گمن و دیگران (Geman et al) ۱۹۹۲ نشان می‌دهد که این مطلب ما را به یک افزایش بیشینه در تعداد دوره تکرار می‌رساند، و برای مثال در حافظه گذاشتن مجموعه مورد آموزش باعث این می‌شود که مجموعه با پارازیت تطبیق داده شود نه با جواب کلی! در هر صورت برای نمونه‌های آموزشی توزیع وزن بصورت بهینه حاصل می‌شود اما بصورت عمومی این مسأله بازده بدی در کل خواهد داشت. پدیده شبیه این مسأله در مسائل توموگرافی اتفاق می‌افتد. وقتی که تطبیق بالاتر از حد معمول شود (اسکالس و اشنايدر ۱۹۹۸ Scales & Snieder).

آموزش بیش از حد بوسیله این واقعیت که معادله ۷ هنگامی که ill-posed می‌شود اتفاق می‌افتد، به عنوان مثال یک فضای تهی بوجود می‌آید. راه حل کمترین مجموع مربعات در مورد سیستم معادلات (۲-۷) و (۲-۸) دارای بهترین بازده خواهد بود. فقط اگر داده‌ها بسیار عالی باشند و سیستم مورد آموزش کوچکترین پارازیتی را نداشته باشد. وگرنه هر راه حلی به خاطر وجود فضای تهی جواب منحصر به فرد خود را خواهد داشت. هموارسازی با معادله (۲-۹) احتمال وجود فضای تهی را کم می‌کند اما همچنان جواب یک راه حل پیش‌بینی شده نخواهد داشت همچنان که در قبل نیز گفته شد.

راه حل کلاسیک این مسأله غامض استفاده از یک مجموعه جداسازی نمونه‌ها است. یک قسمت برای آموزش استفاده می‌شود؛ و قسمت دیگر به عنوان یک مجموعه مرجع برای تعیین حدود اجرای کلی استاده می‌شود (شکل ۲-۴). آموزش متوقف می‌شود وقتی که خطای تطبیق

مجموعه مرجع به یک مینیمم برسد. این روش، نگاهداشتن تقاطع (holdout crossvalidation) نامیده می‌شود.

همچنین این روش عموماً نتایج خوبی دارد، آن در مجموعه‌های آموزشی کاهش یافته که ممکن است یک مشکل بوجود آورد اگر یک تعداد محدودی از نمونه‌ها در دسترس باشد نتیجه می‌دهد. زیرا این روش نیاز به تقسیم‌بندی داخلی برای نمونه‌های حاضر دارد. تعداد نهایی نمونه‌های مورد آموزش حتی اگر زیاد باشند کم می‌شوند. در هر صورت اطلاعات شامل مثالهای انتخاب شده بطور بهینه‌ای استفاده نشده‌اند و شانس آموزش بازسازی داده‌ها بالا می‌باشد.

همچنین افزایش هوشمندانه تعداد نمونه‌های آموزشی m بوسیله وارد کردن پارازیت یا داده‌های ساخته شده از مدل که بدون پارازیت هستند نیز ممکن است. به هر صورت هنگامی که روش‌های هوشمند استفاده می‌شوند بایستی بسیار احتیاط نمود. در این حالت کوچکترین تغییر اتفاقی در ورودی نتایج داده‌ها را به اشتباه خواهد کشاند. از نظر ریاضی این روش هموارسازی وزنی را انجام می‌دهد، ماتسوکلا (Matsouka) (۱۹۹۲) (Bishop) (۱۹۹۵) گراندوال و کانو (۱۹۹۵) (Grabd valet & Ganu)، به وسیله تنها کاهش بعضی از وزنها مؤثر. بعلاوه پارامترهای پارازیت بایستی انتخاب شده و دوباره برای بهینه شدن مورد قضاوت قرار گیرند و یک مدل از پارازیت‌های بد بایستی برای خطاهای سیستماتیک در نظر گرفت. در حالت بعدی مدل ممکن است در حالت واقعی در حالت نامناسبی باشد که باعث حذف و یا تفسیر اشتباه از مکانیسم مهم مدل خواهد شد.

به غیر از حل مسائل با جداسازی مجموعه داده‌ها چند راه حل دیگر نیز وجود داد: روش عمومی‌سازی نقاط تقاطع متغیر، تحلیل باقیمانده و اندازه‌گیری‌های تئوری دیگر که هم خروجی را و هم پیچیدگی شبکه را بررسی می‌کند.

مسئله‌ای که روش تقاطع معتبر را کم بها می‌کند داده‌هایی در بعضی از مثالها هست که از چرخه آموزش بیرون می‌روند. از اینرو، قسمتی از این اطلاعات از دست خواهد رفت، اما نه برای بدست آوردن مسائل مبنایی در این روند مورد توجه. بعنوان یک جایگزین تقاطع متعبر ۷ برابر می‌تواند مد نظر قرار گیرد (مودی ۱۹۹۴؛ چنتف ۱۹۹۷). (Chentouf).

در این روش مثالها به دو مجموعه ۷ (تقریباً) با اندازه مساوی تقسیم می‌شود. آموزش سپس ۷ بار در ۷-۱ مجموعه انجام می‌شود، در هر بار یک عضو از مجموعه خارج می‌شود. خطای تطبیق جزئی شخصی برای هر عضو خارج شده مشخص می‌شود؛ تا وقتی که تطبیق کلی بعنوان

یک میانگین از v تطبیق جزئی شده باشد. آموزش وقتی که مجموع تطبیق‌ها در حالت کمینه خود قرار گیرد یا همگرایی کند شود، متوقف می‌شود. در محدوده $v = m$ این روش به نام «خارج کردن یک عضو» (leave one out) شناخته می‌شود. در این حالت آزمایش بر روی $m-1$ نمونه انجام می‌شود و هر تطبیق جزئی بر روی داده‌های خارج شده محاسبه می‌شود.

نتایج تقاطع معتبر چند v آن است که در نهایت هیچ داده‌ای از چرخه بررسی خارج نمی‌شود. به هر حال تمام اطلاعات قابل دسترسی در نمونه‌های آموزشی استفاده می‌شود. علاوه، آموزش بوسیله قسمت بزرگی از داده‌ها انجام شده که $(m-m/v)$ داده را در برمی‌گرفت. از اینرو، مسائل بهینه‌سازی راحت‌تر بازسازی می‌شوند. به عبارت دیگر آموزش آهسته‌تر است به علت تکرارهای تقاطع معتبر. برای جزئیات بیشتر مراجعه کنید به استون ۱۹۷۴ (Stone) و واهباوولد ۱۹۷۵ (Wahba & Wold). روشهای دیگر آماری نیز می‌توانند مورد توجه قرار گیرند، مانند جک نایف (چاقوی جیبی! Jeddknife) یا (بوت استرپ bootstrap، افرون ۱۹۷۹ Efron، افرون و تیشیرانی Efron & Tibshirani، ماسترز ۱۹۹۵ Masters) دو روش آماری که کوشش می‌کنند تحت شرایط آماری و توزیع‌های آماری از مقدار محدودی داده بدون داشتن پیش زمینه‌ای از توزیع آن اطلاع کسب کند. مودی ۱۹۹۴ (Moody) همچنین یک روش را به نام تقاطع معتبر غیرخطی توضیح داده است.

راه ممکن دیگر برای پرهیز از جداسازی داده‌ها کمینه کردن مقدار تئوری مربوط به پیچیدگی شبکه و بازده عمومی تطبیق می‌باشد (چنتف ۱۹۹۷ Chentouf). بعضی استانداردها براساس توجهات تئوری خاصی باید بررسی شوند. بعضی اندازه‌گیریهای خوب شناخته شده دارای استاندارد AIC و BIC از آکانیک ۱۹۷۰ (Akaike) دارند. و بقیه را می‌توان در جاج و دیگران ۱۹۸۰ (Judge et al) پیدا کرد. برای مثال معیار BIC توسط فرمول زیر داده می‌شود

$$BIC = \ln \left(\frac{\sigma_r^2}{m} \right) + n_{tot} \frac{\ln m}{m} \quad \text{فرمول (۱۰-۲)}$$

که σ_r^2 واریانس باقیمانده خطا (خطای تطبیق) را مشخص می‌کند، اولین عبارت مشخصاً به خطای تطبیق باز می‌گردد و دومین عبارت به پیچیدگی شبکه.

این معیار به هر حال برای سیستم‌های خطی بکار می‌رود و بطور مشخص برای شبکه عصبی بخاطر توزیع فعال سازی غیر خطی آنها کاربرد ندارد. در هر صورت، چندین معیار نیز برای سیستم‌های غیر خطی گسترش یافته‌اند (مک کی ۱۹۹۲؛ MacKay؛ مودی ۱۹۹۲؛ Moody؛ موراتا و دیگران ۱۹۹۴). مانند وضعیت پیشین آنها سعی کردند تا یک رابطه برای تطبیق و یک رابطه که پیچیدگی شبکه را بیان می‌کند بدست بیاورند. به هر صورت این معیارها تلاش برای کمینه کردن خطای تطبیق و پیچیدگی شبکه بطور همزمان بکار می‌رود.

همچنین این معیارها بشدت قدرتمند هستند اگر تحت شرایط فرضیات تئوری باشند که بوسیله تعداد محدودی مجموعه داده مورد بررسی قرار گیرد، بطور مثال، $n_{tot} \gg m$. وگرنه آنها ممکن است یک نقطه خطا در کاهش بازده عمومی شبکه از میزان پیش‌بینی مؤثر باشد. بعلاوه این معیار تنها در شبکه‌های عصبی می‌تواند مورد استفاده قرار گیرد که ساختار آنها برای داده‌های همزمان قابل انطباق باشد.

روش سوم که در شبکه عصبی که بوسیله جوتن و چنفت ۱۹۹۵ (Jutten & Chentouf) تهیه شده است، بوسیله کمینه کردن روشهای بهینه‌سازی آماری صورت گرفته است. که این روش شامل تحلیل آماری خطای باقیمانده است که آن یک تحلیل تطبیقی برای تمام مقادیر خروجی تمام نمونه‌های مورد آموزش انجام می‌شود. و این مطلب بیان می‌کند که یک شبکه بهینه شده بدست می‌آید اگر پارازیت (نوفه) و باقیمانده‌ها رفتار مشابهی را داشته باشند. برای مثال اگر نویز سفید (پاک) فرض شود و اگر باقیمانده صفر باشد یعنی هیچ تطبیقی خارجی وجود ندارد آموزش متوقف خواهد شد (همانطور که بوسیله آموزشهای آماری انجام می‌شود) این روش برای رسیدن به نویز غیرسفید تعمیم داده می‌شود (یعنی جایی که کمی نویز وجود دارد) حسینی و جوتن ۱۹۹۸ (Hosseini & Jutten). پیش زمینه اصلی این روش آن است که فرض قبلی پایه بر اساس خصوصیات و رفتار نویز صورت می‌گیرد.

۲-۳-۴- بهینه‌سازی آرایش: پیش پردازش و هموار سازی وزنی

آخرین مسأله باقیمانده به ساختن یک آرایش شبکه برای نتایج بهینه می‌پردازد. در راه مدل سازی شبکه‌های عصبی ابزارهایی وجود دارند که تعداد مورد نیاز لایه‌های پنهان و نوع

شبکه‌های عصبی را بیان می‌کنند. علیرغم حالت‌های زیاد فقط تعداد سلول‌ها به تعداد لایه‌های پنهان را ارائه می‌دهند.

بطور رسمی این مسأله با روش سعی و خطا حل می‌شود فقط بایستی چند آرایش و بازده آنها آزموده شوند. در نهایت بهترین آرایش انتخاب می‌شود. مشکل اصلی با این روش آن است که نیازمند کار پر زحمت دستی زیادی می‌باشد که ممکن است خیلی هزینه‌بر باشد حتی اگر داده‌های خودکاری برای آموزش و ساختار و آزمایش اجرا داشته باشیم.

به طور که نشان داده شد، نسبت تعداد متغیرهای داخلی کل به تعداد نمونه‌های آموزشی مستقیماً برای ایجاد یک مسألهٔ بینه‌سازی یک شبکه مهم هستند. این مسأله توجه ناگهانی را برای کاربردهایی که مقادیر زیادی بردار ورودی می‌گیرند دارد، برای مثال، n_i بزرگ است. همچنانکه هموارسازی ممکن است به محدودیت تعداد وزن‌های مؤثر کمک کند هاش و هورن (۱۹۹۳) (Hush & Hurne). اغلب تعداد اتصالات مورد نیاز و ندها می‌تواند به آسانی کاهش یابد با استفاده از پیش پردازش تکنیک‌ها برای روشن ساختن اطلاعات مهم که در ورودی می‌باشند یا بوسیلهٔ استفاده از ارتباطات محلی و اشتراک وزنی.

چندین روش پیش پردازش مختلف وجود دارد. به هر حال بهترین روش شناخته شده تحلیل اصول اجزاء می‌باشد یا تبدیل کارهونن لوو (Karhunen Loeve transform). در این روش نمونه‌های آزمایش مانند بردارهای ستونی در ماتریس X جاگذاری می‌شوند. ماتریس کوواریانس XX^T برای بدست آوردن مقادیر ویژه و بردار ویژه، عملیاتی روی آن انجام می‌شود. در نهایت نمونه‌های آموزشی و بعد داده‌ها براساس این بردارهای ویژه تحلیل می‌شوند. بردارهای ویژه P که بزرگترین مقدار ویژه را دارد ($m > P$). این بردارهای ویژه یک مجموعهٔ جدید از محورهایی را بدست می‌دهد که از رتبهٔ معادلهٔ اصلی اولیه کاسته شده و برای تطبیق خطی بین نمونه‌های آموزشی بکار می‌رود. در این روش هر فراوانی در ورودی ممکن است دوباره سازی شود. و بعلاوه آنها شباهتها از داده‌هایی که بوسیلهٔ پارازیت مخدوش شده‌اند از داده‌های اصلی بیرون کشیده می‌شوند. نسبت مجموع بزرگترین مقدار ویژه P (مربع) به مجموع مربعات مقادیر ویژه یک برآورد دقیق برای اطلاعات دخیل در داده‌های هر پروژه می‌باشد. برای پیش زمینهٔ بیشتر می‌توانید به ریچارد (۱۹۹۳) (Richards) و وان دربان و پاول (۲۰۰۰) (Van der Baan & Paul) مراجعه کنید.

ماتریس X ممکن است شامل تمام نمونه‌های آموزشی باشد. نمونه‌هایی فقط از یک کلاس منفرد و یا ماتریس‌های خصوصی برای کلاسهای حاضر. در حالت اخیر هر کلاس باید شبکه مخصوص به خود را داشته باشد و همچنین روال پیش پردازشی داده‌های مربوط به خود را. شبکه‌های خصوصی معمولاً سیستم‌های متخصص نامیده می‌شوند و تنها قادر به شناخت یک کلاس منفرد و همچنین نیاز به پردازش تکراری داده‌ها برای گرفتن تمام کلاسها دارند.

استفاده از تبدیل کارهونن لوو (Karhunen Loeve) اگر چند کلاس مختلف وجود داشته باشد با مشکلاتی همراه می‌باشد. چون این روش بسیار سخت می‌تواند بین کلاسهایی که از عناصر مشترکی استفاده می‌کنند تفاوت قائل شود بعنوان یک جایگزین، یک فاکتور یا تحلیل کانونی می‌تواند مورد توجه قرار گیرد. این روش ماتریس کوواریانس مربوط به تمام نمونه‌ها را به دو ماتریس کوواریانس با کلاس و بین کلاسهای مختلف مجزا می‌کند. سپس یک عملیات انجام می‌شود که نقاط همزمان کمینه را در فاصله‌های بین کلاس و فاصله بیشینه بین کلاسها را پیدا کند. در هر صورت یک عملیات تکی مورد نیاز می‌باشد. توضیحات بیشتر را می‌توان در ریچارد (Richards) ۱۹۹۳ یافت.

دلیل آنکه چرا اصول اجزا و تحلیل فاکتور بازده بیشتری در شبکه‌های عصبی دارند واضح است. گالیناری و دیگران (Galinari et al) ۱۹۹۱ نشان دادند که یک برابری رسمی وجود دارد بین شبکه‌های عصبی خطی (مانند توابع فعال سازی خطی) با تحلیل فاکتورها، یک تشخیص قوی وجود دارد در مورد شبکه‌های عصبی غیرخطی (مانند شبکه‌های MLP و RBF) که خیلی مربوط هستند به تحلیل‌های حد و مرز به هر صورت استفاده از اجزاء اصلی یا تحلیل فاکتورها بری ساده‌سازی ساختار شبکه اجازه می‌دهد از زمانی که یک قسمت از تحلیل بی حد و استفاده از داده‌ها که هنوز استفاده می‌شود. بهر حال کمینه محلی کمتر ممکن است اتفاق بیافتد.

تکنیک‌های پیش پردازش جالب دیگر برای کاهش ورودی می‌توانند در آلמידا ۱۹۹۴ (Almeida) یافت شوند. تمام این روشها قالبی از ساختار شبکه‌های عصبی دارند. دقت کنید که به هر حال که همیشه اجزاء خالص و ورودی‌ها به محدوده‌های خوب شناخته شده مقیاس داده می‌شوند (به طول مثال ۱- تا ۱) برای قرار دادن مقادیر ورودی و یک محدوده متغیر با بیشترین حساسیت نسبت به توابع فعال ساز. این مطلب معمولاً نتیجه بهینه‌سازی استفاده شده در ورودی خواهد بود. در هر صورت این ممکن است توسط لایه‌های پنهان نرونها کم شود. برای مثال لوکان و

دیگران (Le Cum et al) ۱۹۹۱ نشان دادند که تصحیح هر مقدار ورودی خالص برای میانگین و انحراف از معیار استاندارد برای این جزء در مجموعه تمرین شده سرعت آزمایش را افزایش خواهد داد و به علاوه برای نشان دادن داده‌ها در یک محدوده قابل تغییر معمولاً از لگاریتم ورودی $\log(x)$ بجای خود ورودی x استفاده می‌شود.

راه ممکن دیگر برای محدود کردن تعداد متغیرهای داخلی حدس زدن از قبل آنها در یک ساختار شبکه عصبی می‌باشد و بطور خاص درباره ارتباطات بین ورودی و اولین لایه پنهان. برای مثال بجای آنکه لایه ورودی و لایه پنهان را کاملاً بهم مرتبط سازیم تنها ارتباطات محلی بین آنها برقرار می‌کنیم که این فرض منجر به ارتباط تنها سلولهای همسایه به هم خواهد شد. سپس ارتباط بین این گره‌های ورودی و چند نرون پنهان کافی خواهد بود. تنها نکته در این روش این است که این روش ممکن است تعداد لایه‌های پنهان را برای توصیف مناسب مسأله افزایش دهد.

به هر حال، اگر استفاده از اتصال محلی با عناصر وزنی مشترک آمیخته شود، یک کاهش قابل توجه در n_{total} بدست خواهد آمد. همچنین اتصالات دسته‌ای ورودی به گره‌های پنهان یک وزن منطقی را برای مسأله ما ایجاد خواهد کرد. حتی ارتباطات ورودی‌های دسته‌بندی شده به چندین گره می‌تواند عناصر وزنی را قوی‌تر کند. برای شبکه‌های بزرگتر، این روش ممکن است بطور قابل توجهی تعداد متغیرهای آزاد داخلی را کاهش دهد (به لوکان و دیگران ۱۹۸۹ LeCun et al مراجعه شود). متأسفانه نتایج بطرز قابل ملاحظه‌ای بستگی به ساختار شبکه عصبی ما دارند و هیچ شاخصی نیز برای معماری بهینه وجود ندارد.

تکنیک نرم وزنه‌های مشترک نولات و هینتون (Nowlan & Hinton) ۱۹۹۲ یک جایگزین جالب را پیشنهاد می‌کند. از این روش فرض می‌شود که وزنها در گروه‌های مختلفی دسته‌بندی می‌شوند که توزیع آنها براساس توزیع گوسی می‌باشد؛ در حین آزمایش بازده شبکه، میانه و توزیع وزن گوسی و روابط دیگر همزمان بهینه می‌شوند. از وقتی که یکی از توزیع‌های گوسی به حدود صفر می‌رسد این روش با هموارسازی تیخونوف در وزنه‌های مشترک ترکیب می‌شود. یکی از نتایج غلط این روش فرض بسیار قوی در توجه به توزیع وزنی می‌باشد. بعلاوه هیچ روشی برای تعیین تعداد بهینه گوسی وجود ندارد و همزمان با آن هنوز با بهینه‌سازی ساختار مشکل داریم.

۲-۳-۵- بهینه‌سازی آرایش: روشهای آسان سازی

مسأله ساختار پیوسته، می‌تواند به دو روش مختلف حل شود، استفاده از هم سازنده و هم مخرب، برای مثال، روشهای ساده‌سازی. در اولین روش با یک شبکه کوچک شروع کرده و سپس همزمان نرونهای جدیدی را اضافه کرده و آنها را آموزش می‌دهیم. روش دوم با یک شبکه بزرگ آغاز می‌کند و بتدریج شبکه را آموزش داده و گره‌ها و اتصالات اضافی را حذف می‌کند. در ابتدا چند روش ساده‌سازی توضیح داده می‌شود. این روشها به دو دسته تقسیم می‌شوند: آنهایی که تنها اتصال را حذف می‌کنند و آنها که کل گره را حذف می‌کنند. تمام روشهای ساده‌سازی به تکنیکهای حذف باز می‌گردند.

ساده‌ترین روش حذف وزنی آنهایی هستند که به حذف دامنه باز می‌گردند که آنها شامل این هستند که کوچکترین وزن حاضر را حذف می‌کنند و بعد از آن شبکه را دوباره آموزش می‌دهند. به هر حال این روش پاسخهای عالی ارائه نداده است (لوکان و دیگران ۱۹۹۵، حسیبی و استورک ۱۹۹۳ (Le Cun et al, Hassibi & Stork)) از آنجا که چنین وزنهایی، مانند کوچکترین، ممکن است نفوذ قابل توجه و یا بازده عالی در شبکه عصبی داشته باشد.

روش بهتر، بهتر کردن حساسیت تابع خطای تطبیق برای حذف وزنهای خاص می‌باشد. دو روش که به خوبی پاسخ داده‌اند، عبارتند از: آسیب بهینه مغزی (Optimal Brain Damage or OBS) (لوکان و دیگران ۱۹۹۰ (Le Cun et al)) و جراحی بهینه مغز (Optimal Brain Surgeon) or OBS دکتر حسیبی و استورک (Hassibi & Stork) ۱۹۹۳ می‌باشد.

هر دو روش متغیر δE از خطای تطبیق کمترین مقدار مربعات E قابل استناد برای حذف از یک وزن w_i توسط یک بسط درجه دوم سری تیلور که در زیر آمده است تخمین می‌زند:

$$\delta E = \sum_i \frac{\partial E}{\partial w_i} \Delta w_i + \frac{1}{2} \sum_i \frac{\partial^2 E}{\partial w_i^2} (\Delta w_i)^2 + \frac{1}{2} \sum_{i \neq j} \frac{\partial^2 E}{\partial w_i \partial w_j} \Delta w_i \Delta w_j \quad (۱۱-۲) \quad \text{فرمول}$$

عبارات با درجات بالاتر در نظر گرفته نشده‌اند. حذف وزن w_i عبارت $\Delta w_i = -w_i$ را می‌رساند. همچنین تمام روشهای حذفی تنها زمانی بکار برده می‌شوند که شبکه عصبی بکار رفته باشد و پس از استفاده، مقدار کمی آن بدست آمده باشد و آنگاه اولین عبارت در سمت راست می‌تواند صرف نظر شود. بعلاوه روش OBD فرض می‌کند که عبارت نظری ($i \neq j$) از عبارت $\partial E / \partial w_i \partial w_j$ برابر صفر است.

همچنین حساسیت (یا برجستگی) s_i برای تابع تطبیق برای حذف وزن w_i بوسیله عبارت زیر بیان می شود:

$$s_i = \frac{1}{2} \frac{\partial^2 E}{\partial w_i^2} w_i^2 \quad \text{فرمول (۱۲-۲)}$$

وزنهایی با حساسیتهای کمتر حذف می شوند و شبکه عصبی دوباره آموزش می شود. آموزش دوباره باید پس از حذف یک یا چند وزن انجام گردد. عبارت دقیق برای اجزاء قطری ماتریس بوسیله لوکان و دیگران معرفی شده است (لوکان و دیگران ۱۹۹۰ le Cun et al).

روش OBS یک شاخه از روش OBD می باشد، در زمانی که نیاز به آموزش دوباره نمی باشد. بجای صرف نظر کردن از عناصر غیر قطری، این تکنیک از ماتریس هسیان H کامل استفاده می کند. که ترکیبی از هر دو عبارت دوم و سوم در سمت راست معادله (۱۱-۲) است. دوباره، حذف عبارت w_i حاصل می دهد $\Delta w_i = -w_i$ ، که در اینجا بصورت زیر فرمول می شود $e_i' \Delta W + w_i = 0$ ، که بردار e_i ، آمین ستون را از ماتریس شناسایی را نشان می دهد. این مسأله ما را به متغیر δE راهنمایی خواهد کرد:

$$\delta E_i = \frac{1}{2} \Delta W' H \Delta W + \lambda (e_i' \Delta W + w_i) \quad \text{فرمول (۱۳-۲)}$$

(λ ضریب لاگرانژ می باشد). با کمینه کردن عبارت فرمول (۱۳-۲) خواهیم داشت:

$$\delta E_i = \frac{1}{2} \frac{w_i^2}{H_{ii}^{-1}} \quad \text{فرمول (۱۴-۲)}$$

و

$$\Delta W = -\frac{w_i}{H_{ii}^{-1}} H^{-1} e_i \quad \text{فرمول (۱۵-۲)}$$

وزن w_i نتیجه در کوچکترین تغییرات در خطای تطبیق δE در معادله (۱۴-۲) بیان شده است. بعد از آن، در معادله (۱۵-۲) می گوید که چگونه وزنهای دیگر بایستی با شبکه هایی که دوباره آموزش می شوند، تطبیق داده شوند. هنوز بعد از حذف چندین وزن شبکه های عصبی معمولاً برای رسیدن به بازده بالاتر تلاش می کنند.

همچنین روش بخوبی براساس اصول ریاضی بنا گذاشته شده است و یک نتیجه خطا نیز دارد. هم هسیان کامل (full Hessian) و هم معکوس آن بایستی محاسبه شود. ممکن است در این حالت اشکال حافظه پیدا کنیم. به هر حال استفاده از OBS بسیار جالب است اگر معکوس هسیان

یا $(\Delta A + \beta I)^{-1}$ تخمین زده شود از یک الگوریتم بهینه‌سازی درجه ۲ برای آموزش شبکه. توضیح دقیق برای ماتریس Hessian کامل می‌تواند در مقاله حسینی و استورک ۱۹۹۳ یافت شود. در آخر توجه کنید که معادله (۲-۱۱) تنها برای مقادیر کوچک w_i معتبر می‌باشد. همچنین OBD و OBS نباید برای حذف وزنه‌های خیلی بزرگ استفاده شوند. بعلاوه کنترل و دیگران (Cottrell et al) ۱۹۹۵ نشان دادند که هم OBD و هم OBS وزنه‌های تهی آماری را نیز حذف می‌کنند. و بعلاوه روش آماری آنها می‌تواند برای یک مرحله مشخص برای توقف حذف با روش OBD و OBS استفاده شود. چون آنها دریافت‌اند که وزنه‌های حذف نشده توسط دانشجویان ابتدایی که شاخص‌های آماری روشی دارند فراتر می‌رود. هرات و جوتن (Heranult & Jutten).

بجای حذف تنها ارتباطات، کل نرونها می‌توانند حذف شوند. دو روش که استفاده می‌شوند به نامهای روش موزر و اسمولنسکی (Mozer & Smolensky) ۱۹۸۹ و سیتسما و دوو (Sitiesma & Dow) می‌باشد. روش ساختاری موزر و اسمولنسکی شبکه را مانند روشهای OBD و OBS حذف می‌کند. به هر حال حذف کامل گره‌ها تطبیق را بهینه می‌کند و دوباره گره‌هایی که انحرافات کوچکی نشان می‌دهند حذف می‌گردند.

سیتسما و دوو (Sistesma & Dow) ۱۹۹۱ یک روند بسیار ساده را برای حذف نرونها بدست آوردند که نتایج بسیار عالی را داشت. آنها خروجی‌های گره‌ها را در لایه‌های مشابهی بررسی کرده و به واحدهای غیر گسترده یا (NCU) رسیدند. گره‌هایی که خروجی‌های تقریباً ثابت تولید می‌کنند برای همه نمونه‌های آموزشی یا آنها که یک خروجی تطبیق داده شده نشان می‌دهد که این گره‌ها همیشه موافق یا مخالف خروجی هستند. حذف گره‌ها می‌تواند بوسیله تطبیق ساده بایه‌ها و وزنه‌ها برای تمام گره‌هایی که به آن وصل شده‌اند تصحیح گردد. همچنین در اصل هیچ آزمایش دوباره‌ای نیاز نیست همچنین این اغلب برای افزایش بازده به کار می‌رود. همچنین سیتسما و دوو (Sitiesma & Dow) ۱۹۹۱ روش خود را فرموله نکردند و بوسیله عبارت آماری و یک شبکه آماری می‌تواند به سادگی ساخته و حذف شود بوسیله جستجو در میانگین‌ها و ماتریس همبستگی. بعلاوه این اجازه را می‌دهد برای تحلیل اجزای اصلی با هر لایه‌ای که دارای گره‌های نامرتبط حذف شده می‌باشد.

هر دو روش استخوان‌بندی و NCU همچنین اجازه می‌دهند که گره‌های ورودی را حذف کنیم. همچنین آنها بطور مؤثری تعداد متغیرهای داخلی را که شبکه عصبی را توصیف می‌کنند کاهش می‌دهند که آنها قابل توجه هستند در موردی از کمیت‌های محدود شده از نمونه‌های آموزشی.

تمام روشهای حذفی ظرفیت عمومی‌سازی را افزایش می‌دهند در شبکه چون آنها تعداد مقادیر کمینه محلی را کم می‌کنند. روشهای حذف دیگر را می‌توان در نوشته‌های کارنین (Karnin) (۱۹۹۵) و پلینو و فانلی (Pellilo & Famelli) (۱۹۹۳) و کوتزل و دیگران (Cottrel et al) (۱۹۹۵) یافت. یک جمع‌بندی کوتاه از تمام روشهای حذفی که شامل هموارسازی وزنی هم می‌باشد توسط رید (Reed) (۱۹۹۳) ارائه شده است.

در طول حذف، یک مسأله مشابه بعنوان زمان آموزش اتفاق می‌افتد. چه زمانی حذف بایستی متوقف شود؟ در عمل حذف کردن زمانی متوقف می‌شود که شبکه عصبی بعد از آن حذف نتواند بهترین تطبیق را بدهد ولی به هر حال این مسأله می‌تواند بهینه‌ترین پاسخ نباشد. روش بهتر دیگر آن است که از روشهای توصیفی در تقسیم‌بندی یا عمومی‌سازی استفاده شود. معیار تئوری غیر خطی می‌بایست بطور خاص مورد توجه قرار گیرد. چون آنها شامل تقابل پیچیدگی شبکه در مقابل عدم تطبیق می‌باشند.

۲-۳-۶- بهینه‌سازی آرایش: روشهای سازنده

به عنوان آخرین احتمال برای ساختن ساختارهای شبکه‌ای بهینه به روشهای سازنده توجه می‌کنیم. بیشتر روشها از تخریب کردن شروع می‌کنند تنها لایه‌های ورودی و خروجی تعریف می‌شوند و سپس آنها بطور خودکار اندازه شبکه را افزایش می‌دهند تا وقتی که به همگرایی دلخواه برسند. مسأله اصلی در این روش معیاری برای نقطه توقف می‌باشد. به عبارت دیگر مجموعه مورد آموزش بسیار ساده در حافظه قرار داده شده و بسیار ساده عمومی‌سازی می‌شود و در نهایت ضعیف خواهد بود. بیشتر اندازه‌گیریهای تئوری که انجام می‌شوند یک نسبت بین پیچیدگی شبکه و بازده دارند که از آن نسبت استفاده می‌شود.

۱. امروزه متدهای سازنده برای شبکه‌های MLP و هم RBF وجود دارند که حتی برای ترکیب این شبکه‌ها نیز بکار می‌روند. برای مثال برای شبکه‌های عصبی که توابع فعال مختلطی دارند. احتمالاً بهترین روش سازنده‌ای که شناخته شده است روش ساختن تطبیقی آبخاری از فالمن و لیبیر ۱۹۹۰ می‌باشد (Fahlman & Lebiere). این روش با لایه‌های ورودی و خروجی کاملاً آموزش یافته و متصل شده شروع می‌کند. سپس یک گره پنهان اضافه می‌شود که فقط به لایه ورودی متصل است. برای بدست آمدن بیشترین تطبیق خروجی این گره پنهان و خطای پیش‌بینی شده در شبکه مورد آموزش دارای بیشترین عدم تطبیق خواهند شد. سپس این گره به لایه خروجی وصل می‌شود. ماتریس وزنی از لایه ورودی خروجی بهینه می‌شوند. در تکرار بعدی یک گره پنهان دیگر اضافه می‌شود که به لایه ورودی وصل شده است و خروجی تمام گره‌های قبلی به آن اضافه می‌شود. دوباره همگرایی مطلق این خروجی و خطای پیش‌بینی شده شبکه عصبی بهینه می‌شود. بعد از آنکه اتصالات برقرار شد دوباره ثابت نگاه داشته می‌شود (Freeze می‌شود) و تمام اتصالات خروجی دوباره محاسبه می‌شود. این عملیات تا همگرایی نهایی ادامه پیدا خواهد کرد. دو گره جدید یک لایه پنهان را شکل خواهد داد. در هر صورت این الگو، شبکه بسیار گسترده و عمیقی را خواهد ساخت که در آن هر گره با تمام گره‌های دیگر متصل است. بعلاوه الگوی اولیه هیچ معیار توفقی را استفاده نمی‌کند چون ورودی بدون هیچ گونه نویزی فرض شده است.

دو روش فرض شده که هیچگونه نتایج بدی ندارند الگوهای افزایشی مودی ۱۹۹۴ (Moody) و جوتن و چنتف (J utten & Chentouf) ۱۹۹۵ می‌باشند. این الگوریتم‌ها با همبستگی آبخاری تفاوت دارد و تنها یک لایه پنهان منفرد برای تمام اتصالات استفاده می‌شود. دو روش در تعداد سلولهای عصبی که هر تکرار اضافه می‌شوند تفاوت دارد. یک سلول (جوتن چنتف ۱۹۹۵ یا چندین سلول مودی ۱۹۹۴) و همچنین در معیار توقفشان. همچنان که مودی ۱۹۹۴ از معیار پیش‌بینی عمومی خطا از مودی ۱۹۹۲ استفاده می‌کند. جوتن و چنتف ۱۹۹۵ تطبیق باقیمانده‌ها را تحلیل می‌کند. سازندگی بیشتر تمام خواهد شد اگر خصوصیات عدم تطبیق اندازه‌گیری شده با خصوصیات نطفه فرض شده مشابه باشد.

یک نوع از الگوی جوتن و چنتف (جوتن و چنتف ۱۱۹۶) اجازه ایجاد اتوماتیک یک شبکه عصبی را با چندین لایه می‌دهد. این راه یک راه عمومی و منطقی از راه اولیه می‌باشد. به هر حال این مسأله وقتی که یک سلول عصبی جدید بایستی در لایه میانی قرار داده شود و یا یک لایه

جدید باید ساخته شود، ایجاد می‌شود. نمونه دیگر (چنتف و جوتن ۱۹۹۶) به هر دو سلولهای گوسی و سیگموئیدی اجازه شرکت می‌دهد. این مسأله می‌رساند که چه نوع از تابع فعال‌سازی بزرگترین کاهش در عدم تطبیق را باعث می‌شود (همچنین به مقاله چنتف در ۱۹۹۷ رجوع شود). روش تعدیل کاهش پویاز برتلهلود دیاموند (۱۹۹۵) (Berthold & Diamond) یک روش افزایشی برای شبکه‌های RBF می‌باشد که بطور خودکار تعداد سلولهای عصبی را تعیین می‌کند و همچنین مرکز و انحراف تابع فعال‌سازی گوسی را به بهترین روش پیشنهاد می‌کند برای یک دسته‌بندی دقیق از نمونه‌های آموزشی. این نمونه‌های آموزشی مرکز تابع فعال‌سازی گوسی را در لایه پنهان مشخص می‌کند. وزن هر عنصر گوسی بطور نسبی انتخاب می‌شود و انحراف ناحیه تأثیر را مشخص می‌کند. برای تعیین این وزنها و انحرافها روش از تمام روشهای منفی و مثبت پایه استفاده می‌شود. حالت‌های منفی آغازین خروجی را برای کلاسهای خطا محدود می‌کند، هنگامی که مقدار مثبت یک مقدار کمینه را مشخص می‌کند، برای تصحیح کلاسهای مورد اعتماد. پس از آموزش، نمونه‌های آموزشی حداقل بایستی یک خروجی را با حداکثر مقدار مثبت بسازند برای یک کلاس صحیح و هیچ خروجی را برای کلاسهای غیر صحیحی که بزرگتر از مقدار اولیه منفی باشد، نسازند.

در طول آموزش، الگوریتم، بطور پی در پی نمونه‌های آموزشی را به شبکه ارائه می‌دهد. اگر یک نمونه آموزشی نتواند به درستی با شبکه جاری دسته‌بندی شود، سپس این نمونه مورد آموزش بعنوان یک نمونه اولیه استفاده می‌شود. در غیر این صورت، وزن نزدیکترین نمونه اولیه افزایش می‌یابد برای افزایش پیشامد نسبی. مقادیر تمام توابع گوسی را که دارای تضاد هستند تا آنجا که باعث تضاد و برخورد نشوند کاهش می‌یابند تا آنجا که مقدار کلاس دارای تضاد بزرگتر از مقدار منفی اولیه نمونه‌های مورد آزمایش باشد. خروجی دسته‌بندی نمی‌شود چون تابع فعال‌سازی خطی که وجود دارد گره‌های خروجی را مقدار می‌دهد. در هر صورت یک شبکه دقیق ساخته می‌شود که بسیار شبیه کلاسی برای یک نمونه آموزشی است هر چند این شباهت کامل نمی‌باشد. الگوریتم تعدیل پویای برتهولد و دیاموند ۱۹۹۵ شبکه شبیه عصبی احتمالاتی اسپچت ۱۹۹۰ (Specht) می‌باشد. این شبکه الگوی کاوسی را در هر نمونه آموزشی می‌سازد. در طول آموزش، فقط مقدار بهینه، مقادیر عمومی برای تمام گاوسی‌ها بایستی تخمین شوند. به هر حال عملکرد گره‌های پنهان برای هر نمونه آموزشی ساخته می‌شود که شبکه را می‌سازد بیشتر یا کمتر از مقدار

حافظه‌ای که برای آن در نظر گرفته شده و آن را بهینه خواهد کرد برای استفاده از مقادیر زیادی از نمونه که بایستی به کار گرفته شوند. تعدیل کاهش پویا به عبارت دیگر گره‌های جدید را تنها هنگامی می‌سازد که نیاز باشد.

الگوریتم‌های افزایشی دیگر شامل کمترین مقدار مربعات از چن و دیگران (Chen et al)، منابع جایگزین شبکه از پلات (Platt) ۱۹۹۱، پروژه آزمون جایگزین هوانگ و دیگران ۱۹۹۴ (Hwang et al) می‌باشد. برای مروری بر الگوریتم‌های سازنده می‌توان به کوک و یونگ ۱۹۹۷ (Kwok & Yenug) مراجعه کرد.

۲-۴- یک استراتژی عمومی

چگونه این روشها و تکنیکها می‌توانند در عملیات ژئوفیزیکی بکار روند؟ لیستی که در زیر ارائه می‌شود حاوی نکاتی است که باید برای هر عملیاتی در نظر گرفته شود دقت به تمام نکات آن بایستی در نظر گرفته شود.

انتخاب شبکه عصبی - برای مسائل ایستا، تحلیل از قبل داده‌ها یا بازبینی عمومی می‌تواند یک انتخاب بهینه را وقتی از شبکه‌های MLP یا RBF استفاده می‌کنیم به دست دهد. برای مثال، دسته‌بندی در مسائل خاص معمولاً در زمینه ورودی بکار می‌رود. به هر حال شبکه‌های RBF ممکن است جواب بهتری از شبکه‌های MLP داشته باشد. به هر صورت، هر دو نوع این شبکه‌ها یک تقریب وسیع برای معادلات می‌باشند که توانایی تولید پاسخ منطقی را دارند. علی‌رغم اینکه یک نمونه می‌تواند برای کاربرد خاصی مناسب باشد نمونه دیگر ممکن است نتایج بهتری را ارائه دهد و اگر هیچ انتخابی وجود نداشته باشد هر دوی آنها بایستی بکار روند.

انتخاب پارامترهای ورودی - در بیشتر مسائل این یک مسئله مورد سؤال می‌باشد. در بیشتر حالت‌های خاص، هر پارامتری که بایستی در مسئله گنجانده شود باید بیاید، بعد از تحلیل اصلی اجزاء (PCA) (Principal Component Analysis) یا تحلیل فاکتور بایستی استفاده شود. برای کاهش فضای ورودی و همچنین کاهش پارامترهای غیر مربوطه. علی‌رغم آن یک انتخاب مناسب از پارامترها مشخصاً بازده و کیفیت نتایج را بهتر خواهد کرد.

روشهای مناسب پیش برداشتی - هر تغییر مقیاسی، فیلتر کردن، یا هر عملیات دیگری که تاثیر روی ورودی بگذارد باید مورد بررسی قرار گیرد. بطور معمول PCA یا تحلیل فاکتور می تواند در این بخش هم مطرح شود.

مجموعه آموزشی و نمونه های آموزشی - تعداد نمونه های آموزشی مستقیماً روی تعداد کلی متغیرهای داخلی را که به شبکه عصبی اجازه نگهداری نمونه های پیش سازنده را می دهند اثر می گذارد. تعداد همه متغیرهای داخلی نباید هرگز از تعداد نمونه های آموزشی بیشتر باشد. بزرگترین شبکه عصبی با اتصالات بسیار زیاد بایستی توسط معادلات (۲-۳) و (۲-۴) محاسبه شود. طبیعتاً محدودیتی برای مجموعه های بزرگتر وجود نخواهد داشت.

الگوریتم آموزشی و اندازه گیرهای عمومی - معمولاً الگوریتم آموزشی انتخاب می شوند. روشهای سرهم بندی شده تدریجی بازده بهتری از الگوهایی که از پیش ثابت شده اند دارند، از زمانی که اولین همگرایی با تعداد محدودی از تکرار باید ایجاد شود. بعلاوه یک روش باید برای اطمینان از بازده خوب بطور عمومی اتخاذ شود. این مطلب می تواند بوسیله هر مطلب عمومی انجام شود، تقاطع معتبر و یا اندازه گیری تئوری، و یا تحلیل باقیمانده ها. به هر حال این اندازه گیری ها بایستی در طول آموزش محاسبه شوند و نه بعد از همگرایی.

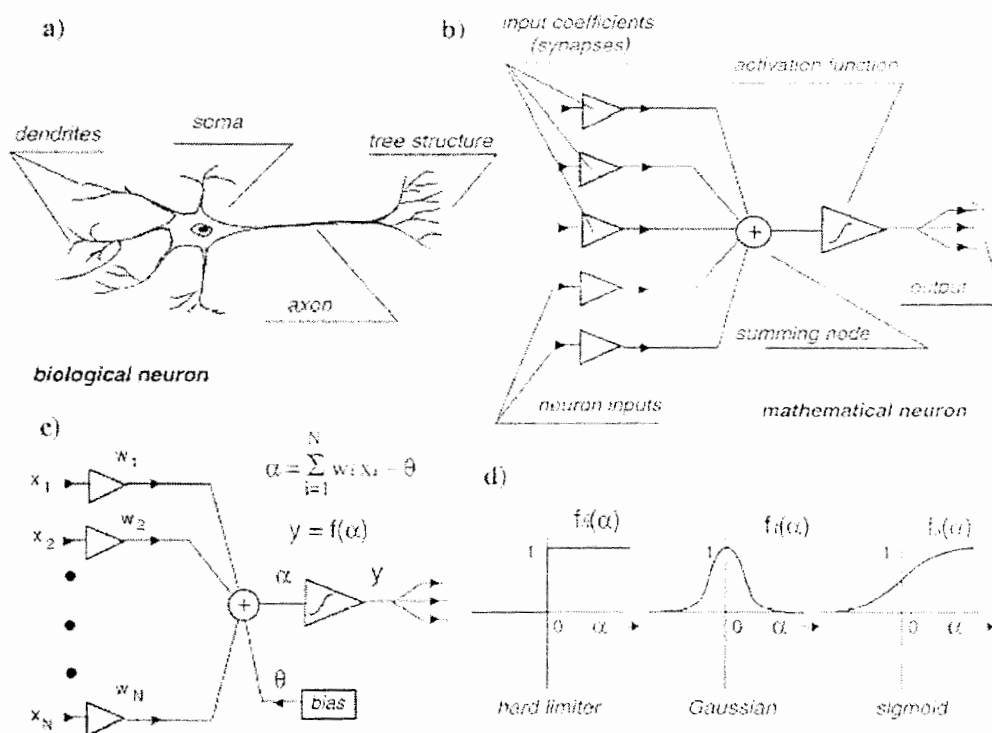
تعیین آرایش - انتخاب بین استفاده از یک سازنده یا روش ساده کننده بسیار مهم است. یک انتخاب عمومی بوسیله هر الگوریتم سازنده ای می تواند برای ایجاد یک آرایش مناسب و همچنین یک تکنیک حذف مناسب برای بهینه کردن آرایش استفاده شود. بعضی اوقات دوباره از نو شروع کردن و دوباره آموزش دادن یک شبکه عصبی ممکن است عدم تطبیق را برای ادامه حذف بهتر کند. در این حالتها دوباره از نو شروع کردن باید برای خروج از یک مقدار کمینه محلی انتخاب شود.

۲-۵- بحث و نتیجه گیری

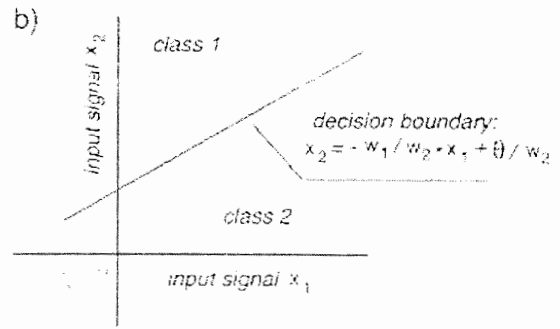
شبکه های عصبی یک تقریب زن کلی هستند. آنها شامل مقادیر اتفاقی بسیار نزدیکی به هر تابع پیوسته هستند که با یک روش مستقیم یا معکوس همراه است. بنابراین، آنها یک ابزار بسیار قوی برای ارتباط ژئوفیزیکی برای حل مسائل که دارای پیچیدگی می باشند هستند.

در اینجا ما روشها و تکنیکهای بسیار متفاوتی را تعریف کردیم برای بکار بردن آنها و افزایش بازده آنها. آخرین اصل این مطلب را می‌رساند که با مجموعه آموزشی مرتبط است. همچنانکه در روشهای دیگر نیز مشاهده می‌شود کیفیت نتایج حاصله با کیفیت داده‌های ورودی بالا و پایین می‌رود.

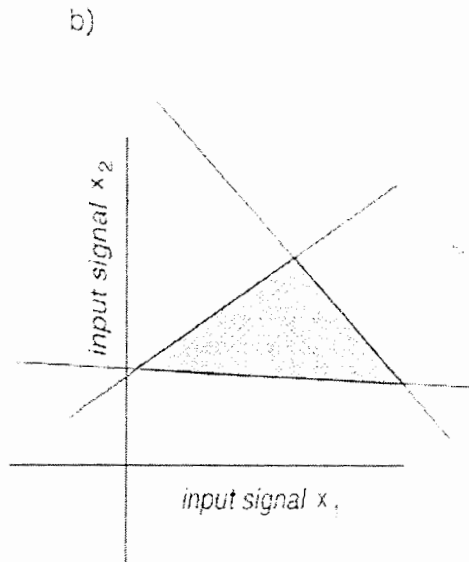
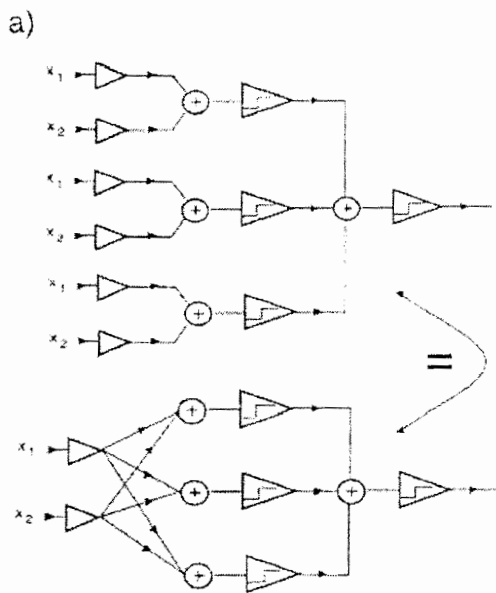
بعلاوه در ابتدا بایستی به این مطلب توجه شود که آیا استفاده از روش شبکه عصبی برای کاربردهای کوچک زمان تحقیق بسیار زیادی را می‌طلبد. عموماً این سؤال برای شرایط عملی که آیا نمونه‌های آموزشی خوب کافی می‌تواند برای تضمین یک فرآیند آموزشی تحمیل شده بدست آید. این مشکل ممکن است کاربرد موفق این روش حتی بعد از پیش پردازش مهم داده‌ها و کاهش تعداد پارامترهای ورودی بتاخیر اندازد. اگر یک پاسخ منفی بایستی داده شود به این سؤال مناسب، سپس یک جایگزین بهتر برای گسترش پایه‌های ریاضی جدید برای کاربردهای خاص ادامه می‌یابد. مطالب این فصل از منبع [۶] گرفته شده است.



شکل ۱-۲: نرون بیولوژیکی و ریاضی. نرون ریاضی (b) رفتار نرون بیولوژیکی (a) را تقلید می‌کند. جمع وزنی ورودیها توسط یک تابع محرک (c) ترسیم شده است. که چندین نمونه از آن در (d) نشان داده شده است. اقتباس شده از لیپمن (Lippmann 1987). هراولت و جوتن (Herault and Jutten) ۱۹۹۴ و رومئو (Romeo) ۱۹۹۴.



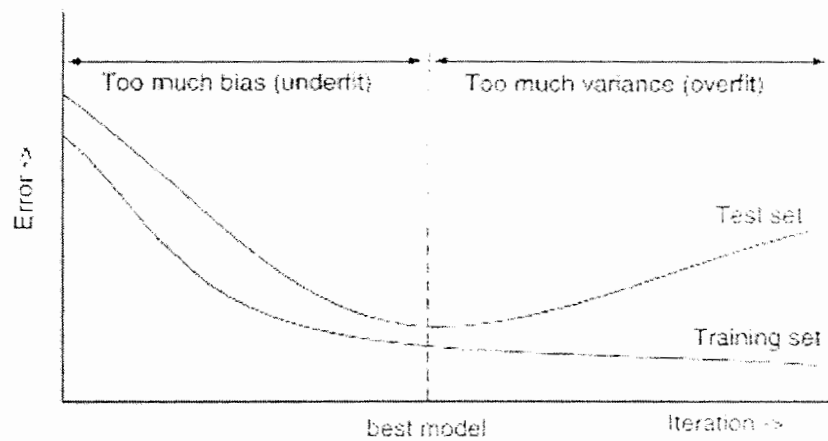
شکل ۲-۲: (a) لایه پرسپترون ساده و (b) مربوط به مرز تصمیم‌گیری می‌شود. اقتباس شده از رومنو ۱۹۹۴.



شکل ۲-۲: (a) لایه پرسپترون میانی ساده و (b) مرز تصمیم‌گیری را مشخص می‌کند. اقتباس شده از رومنو ۱۹۹۴.

MLP	Multilayer Perceptrons
NCU	Noncontributing Units
OBD	Optimal Brain Damage
OBS	Optimal Brain Surgeon
PCA	Principal Component Analysis
RBF	Radial Basic Functions
SCG	Scaled Conjugate Gradient

جدول ۱-۲. اختصارات



شکل ۴-۲: عمومیت دادن در مقابل خطای آموزش. اقتباس شده از موودی ۱۹۹۴ (Moody).

فصل سوم

کاربرد شبکه‌های عصبی در ژئوالکتریک

۳-۱- شبکه‌های عصبی مصنوعی برای تخمین پارامتر در ژئوفیزیک^۱

۳-۱-۱- خلاصه

اینجا ما یک مطالعه از قابلیت اجرایی شبکه‌های عصبی را برای حل برخی مسائل وارونه‌سازی ژئوفیزیکی شرح می‌دهیم. به ویژه، ما مسائل بدست آوردن مقاومت‌های ویژه سازند و ضخامت لایه‌ها را از داده‌های سونداژ الکتریکی عمودی (VES) (Vertical Electrical Sounding) مطالعه می‌کنیم. ما از یک شبکه عصبی پیشخور دولایه (Feedforward Neural Network) استفاده می‌کنیم که برای پیش‌بینی مدل‌های زمین از داده‌های اندازه‌گیری شده آموزش دیده است. بخش قابل توجه در استفاده FNNها برای وارونه‌سازی ژئوفیزیکی این است که آنها سیستم‌های انطباق‌پذیر هستند که یک نگاشت غیرخطی بین دو مجموعه از داده‌ها از یک حوزه داده شده انجام می‌دهند. در هر دو کاربرد زمان، ما FNNها را با استفاده از داده‌های مصنوعی به عنوان ورودی به شبکه‌ها و یک لایه از پارامترهای مدل‌ها به عنوان خروجی شبکه آموزش می‌دهیم.

۱- مطالب این بخش از منبع [۸] آورده شده است.

مدلهای زمین استفاده شده برای آموزش شبکه، از یک مجموع مدلهای تصادفی درون تعدادی محدوده‌های پارامتری از قبل مشخص شده استفاده شده‌اند. برای آموزش شبکه ما به ترتیب از الگوریتم پس انتشار و یک روش ذوب شبیه‌سازی شده پس انتشار هیبرید برای VES و مسائل معکوس لرزه‌ای استفاده می‌کنیم. موضوعات پایه‌ای دیگر برای به دست آوردن تخمین‌های دقیق پارامترهای مدل با استفاده از FNN های پارامتری کردن مدل هستند. شبیه‌سازی‌های ما نشان داد که FNN ها، اگر به قدر کافی آموزش یابند، منطقاً مدلهای درستی از زمین را هنگامی که داده‌های مشاهده شده، ورودی برای FNN ها باشند، تولید می‌کند.

۳-۱-۲- مقدمه

قابلیتهای یادگیری و انطباق‌پذیری این شبکه‌ها آن‌ها را برای به کارگیری در برخی مسائل در ژئوفیزیک اکتشافی جالب توجه می‌سازد.

در روشی که بر پایه ANN (شبکه عصبی مصنوعی) برای مساله زمان ورود موج است e.g. یک مفسر نمونه‌هایی از ورودی‌های تفسیر شده قبلی تهیه می‌کند که توسط شبکه برای ایجاد یک نگاشت تابع از یک حوزه ورودی (داده‌ها) به یک حوزه خروجی (ورودی‌ها) مورد استفاده قرار می‌گیرند. به عبارت دیگر، نگاشت بر پایه مقاومت تجارب قبلی تهیه شده توسط مفسر، انجام می‌شود. یکبار که شبکه امواج را برای نمونه‌های انتخاب شده توسط مفسر، دوباره تولید می‌کند. ANN به عنوان یک روش خودکار برای تفسیر داده‌های جدید مورد استفاده قرار می‌گیرد.

ANN ها همچنین برای مسائل دیگر به کار برده می‌شوند. مانند آنهایی که در تفسیر داده‌های چاه پیمایی، تحلیل موج، برشی شکافتی و تشخیص مخزن آب روبرو می‌شوند. کاربردهای مهم دیگر ANN ها حل مسائل وارونه ژئوفیزیکی بوده است. این شامل محل اهداف زیرزمینی از داده‌های صحرائی الکترو مغناطیسی و وارونه سازی داده‌های مشکل موج لرزه‌ای می‌شود.

حل یک مساله وارونه ژئوفیزیکی شامل استفاده از یک مجموعه از مشاهدات برای پی بردن به پارامترهای زیرزمین است. برای این منظور، ما ابتدا یک مدل ریاضی را فرض می‌کنیم که روابط بین داده‌های اندازه‌گیری شده و مدلهای شناخته نشده را توصیف می‌کند. معیارهای بازیافت خوب به یک تابع هدف یا تابع خطا که عموماً یک تابع غیرخطی از پارامترهای مدل است متمرکز

می‌شوند. معمولاً تابع هدف اختلاف بین مشاهدات و پیش‌بینی‌ها را اندازه‌گیری می‌کند، و کمینه‌سازی آن مساله را به یک حل بهینه هدایت می‌کند. اطلاعات اولیه به وارونه‌سازی برای اجتناب از تناقض و ناپایداری در راه حل داده می‌شود.

یک روش عمومی برای مسائل تخمین پارامتر کمینه‌سازی مکرر است که بر پایه مشتقات محلی تابع هدف است. برنامه‌های خطی شده فرض می‌کنند که نقطه شروع برای کمینه‌سازی مساله را به می‌نیمم مطلق هدایت می‌کند. اگر یک راه حل شروعی خوب وجود نداشته باشد، روشهای جستجوی سراسری مورد نیاز قرار می‌گیرند. تکنیک‌های مونت کارلو، مانند الگوریتم‌های ژنتیک و ذوب شبیه‌سازی شده، اخیراً برای چندین مسائل وارونه‌سازی در جایی که فقدان راه حل‌های شروعی خوب استفاده از روشهای گرادین را محدود می‌کند، به کار برده گرفته شده‌اند.

برخلاف روش ساده مونت کارلو در جایی که مدلها به طور یکسان در حالت تصادفی و به طور قطعی پذیرفته شده، انتخاب می‌شوند. SA و GA از ارزیابی‌های مدل قبلی برای جلو بردن جستجو در فضای پارامتر مدل برای بیشتر بخش‌های امیدبخش از سطح خطا، استفاده می‌کند. این روشها خیلی گرانتر از تکنیکهای خطی شده هستند و از اینرو تنها بیشتر مواقعی مفید هستند که یک راه حل آغازین خوب در دسترس نباشد.

شبکه‌های عصبی یک روش متفاوت برای مسائل وارونه ژئوفیزیکی پیشنهاد می‌کند. اینجا ما از یک شبکه عصبی به نام شبکه عصبی پیشخور (FNN) به عنوان یک تابع نگاشت بین داده‌های اندازه‌گیری شده و پارامترهای مدل زیرزمینی استفاده می‌کنیم. مثالهای این نگاشت در یک فرایند شناخته شده به عنوان آموزش شبکه استفاده شده‌اند. در اصل، یک شبکه به اندازه کافی آموزش دیده بایستی برای نشان دادن یک خروجی در حوزه مثالهای استفاده شده برای آموزش، قادر باشد. ما مطالعات ترکیبی خودمان را از کاربرد FNN ها برای مسائل وارونه‌سازی ژئوفیزیکی، مثلاً، تفسیر داده‌های سونداژ مقاومت ویژه ارائه می‌دهیم.

ANN ها یک روش متفاوت برای مسائل وارونه ژئوفیزیکی پیشنهاد می‌کنند. اینجا ما از یک ANN شناخته شده به عنوان شبکه عصبی پیشخور (FNN)، به عنوان یک تابع نگاشت بین داده‌های اندازه‌گیری شده پارامترهای مدل زیرزمین استفاده می‌کنیم.

مثالهای این نگاشت در یک فرایند به عنوان آموزش شبکه، مورد استفاده قرار گرفته‌اند. در اصل، یک شبکه به طول کافی آموزش یافته، بایستی قادر به پیش‌بینی یک خروجی در حوزه مثالهای

استفاده شده برای آموزش باشد. در این کار، ما از الگوریتم آموزشی پس انتشار برای مطالعه مساله مقاومت ویژه و یک برنامه آموزشی هیبرید که ترکیبی از روش پس انتشار و یک روش جستجوی سراسری شناخته شده به عنوان ذوب شبیه‌سازی شده خیلی سریع است برای مسائل لِرزه‌ای استفاده می‌کنیم. VFSA ثابت شده است که می‌تواند یک روش بهینه‌ساز سریع برای حل مسائل مختلف وارونه‌سازی ژئوفیزیکی باشد.

۳-۱-۳- شبکه عصبی پیشخور (FNN)

شبکه عصبی اغلب به طور وسیع در حل مسائل الگوشناسی در FNN استفاده شده است، بعلاوه سادگی اجرای آنها و قابلیت تعمیم آنها برای حل مسائل گوناگون. عناصر محاسبه‌ای منحصر به فرد یک FNN معمولاً به نرون‌ها یا عناصر فرایند ارجاع داده می‌شوند. هر نرون شامل یک بردار وزنهای قابل تعدیل یا توانایی‌های اتصال است. وظیفه یک نرون نگاشت یک بردار ورودی داده شده به یک خروجی منفرد است که به دیگر نرون‌ها منتقل شده است. در مدل نشان داده شده در شکل (۳-۱-۳)، هر عنصر بردار ورودی داده شده در یک وزن متناظر ضرب می‌شود و با هم جمع می‌شوند تا یک ورودی شبکه تولید کنند. نرون از یک تابع فعال‌ساز برای تبدیل ورودی شبکه به یک خروجی منفرد استفاده می‌کند. تابع فعال‌سازی که عموماً در یک FNN استفاده می‌گردد تابع سیگموئید است. این تابع یک بخش انتقال خطی دارد که به طور نمایی یکی از دو حالت را حاصل می‌کند.

نرون‌ها در یک FNN در لایه‌های متوالی با اتصال بین نرون‌های یک لایه و لایه‌هایی که در پی می‌آیند قرار گرفته‌اند اما نه با اتصال بین نرون‌های درون لایه مشابه (شکل ۳-۱-۳ b). در این آرایش لایه، داده‌ها در یک جهت شروع از لایه اول یا لایه ورودی جریان پیدا می‌کنند. وزن‌ها عموماً توسط کمینه کردن تفاضل بین خروجی‌های شبکه، یکبار بعد از اینکه مجموعه‌ای از بردارهای داده‌های ورودی یا نمونه‌ها، از میان شبکه عبور داده شدند، و مطابق نمونه‌های خروجی مطلوب، محاسبه می‌شوند. در زمینه شبکه عصبی، فرایند پیدا کردن وزن‌های بهینه به عنوان آموزش شبکه شرح داده می‌شود و مجموعه نمونه‌های ورودی - خروجی مجموعه آموزشی است. هدف یادگیری یک نگاشت بین نمونه‌های ورودی - خروجی شناخته شده برای یک مساله داده شده، بکار بردن

شبکه آموزش یافته برای نمونه‌های ورودی با خروجی‌های ناشناخته بعد از آن است. از این نقطه نظر، یک FNN می‌تواند به عنوان یک تابع غیر خطی شرح داده شود؛ فرایند آموزش، یک مسأله بهینه‌سازی است در جاییکه وزنه‌های شبکه پارامترهایی برای تخمین هستند.

اغلب بطور وسیع روش آموزش استفاده شده به عنوان روش پس‌انتشار شناخته شده است. روش پس‌انتشار یک تطبیق کمترین مربعات بین خروجی شبکه واقعی و یک خروجی مد نظر به وسیله محاسبه یک گرادیان محلی در جمله‌هایی از وزنه‌های شبکه، تولید می‌کند (Rummelhart et al. 1986). تنها روشی که همگرایی به یک بهینه محلی را تضمین می‌کند. روش دیگر برای محاسبه وزن روش مونت کارلو (Monte - Carlo) است که در آن وزنها بطور تصادفی آشفته شده و بطور قطعی پذیرفته می‌شوند. این روش ممکن است به یک راه حل بهتر از پس‌انتشار ساده هدایت کند اما آن به یک افزایش اساسی در زمان کامپیوتر به عنوان یک نتیجه از جستجوی نامشخص نیاز دارد.

در این کار، ما از الگوریتم آموزش پس‌انتشار برای مطالعه مسأله مقاومت ویژه استفاده می‌کنیم.

۳-۱-۴- آموزش FNN

فرض می‌کنیم که ما یک مجموعه از Q جفت بردارهای $(x^1, y^1), (x^2, y^2), \dots, (x^Q, y^Q), \dots, (x^N, y^N)$ داریم و فرض می‌کنیم که هر بردار x, N عنصر دارد و هر بردار y, M عنصر دارد. ما می‌خواهیم یک FNN را برای تقریب یک خروجی برای هر بردار ورودی آموزش دهیم. یک شبکه دو لایه‌ای مثل شبکه‌ای که در شکل (۳-۱) نشان داده شده در نظر می‌گیریم. بردار ورودی X^q از میان هر نرون در لایه میانی توسط معادله زیر عبور داده می‌شود:

$$a_j^q = f^h \left(\sum_{i=1}^N w_{ji}^h x_i^q + b_j \right), \quad \text{معادله (۳-۱)}$$

که در آن a_j^q پاسخ شبکه از لایه میانی برای q امین بردار ورودی است، $j=1, \dots, L$ ، که در آن L عدد نرونها در لایه میانی است، w_{ji}^h در آن i امین وزن نرون j است، بالانویس h به لایه میانی اشاره می‌کند، و b_j یک جمله انتقالی (همچنین بایاس نیز نامیده

می‌شود) تابع فعال‌ساز f^h است. تابع فعال‌ساز می‌تواند چندین شکل به خود بگیرد و عموماً یکی از آنها استفاده می‌گردد و آن تابع سیگموئید است

$$f(x) = \frac{1 - e^{-2\alpha x}}{1 + e^{2\alpha x}}, \quad \text{معادله (۲-۳)}$$

که α پارامتری است که تندی تابع را در نزدیکی $x=0$ کنترل می‌کند. در یک روش مشابه، معادلات برای ندهای خروجی به شکل زیر داده می‌شوند

$$o_k^q = f^o \left(\sum_{j=1}^L w_{kj}^o a_j^q + c_k \right), \quad \text{معادله (۳-۳)}$$

که o_k^q در آن k آمین خروجی شبکه است، $k=1, \dots, M$ ، در آن w_{kj}^o وزن نرون خروجی k است، و c_k ، k آمین جمله انتقال تابع فعال‌ساز f^o است که همچنین توسط معادله (۳-۳) تعریف شده است. خطا توسط معادله زیر محاسبه می‌شود

$$E = \frac{1}{2} \sum_{q=1}^Q \sum_{k=1}^M (y_k^q - o_k^q)^2. \quad \text{معادله (۴-۳)}$$

در روش پس‌انتشار، جهت در چیزی که وزنها جدید شده‌اند توسط منفی گرادیان E نسبت به هر عنصر ماتریس‌های وزن W^0 و W^h داده شده‌اند. توجه داشته باشید که جدید شدن‌های وزن در لایه خروجی شروع می‌شود و به سمت لایه اولی حرکت می‌کند، از این رو به آن “پس‌انتشار” می‌گویند. کمینه‌سازی نتایج خطا در معادلات زیر برای جدید کردن خروجی و وزنها لایه میانی داده شده است:

$$w_{kj}^o(l+1) = w_{kj}^o(l) + \eta \sum_{q=1}^Q \delta o_k^q a_j^q \quad \text{معادله (۵-۳)}$$

9

$$w_{ij}^h(l+1) = w_{ij}^h(l) + \eta \sum_{q=1}^Q \delta h_j^q x_i^q, \quad \text{معادله (۶-۳)}$$

که بترتیب

$$\delta o_k^q = (y_k^q - o_k^q) f^{o'}. \quad \text{معادله (۷-۳)}$$

9

$$\delta h_j^q = f^{h'}. \sum_{k=1}^M \delta o_k^q w_{kj}^o \quad \text{معادله (۸-۳)}$$

در این معادلات، f' مشتق f است، و عامل η به عنوان پارامتر میزان یادگیری شناخته شده است و معمولاً یک عدد کوچک می‌گیرد، گفته می‌شود بین $0/10$ و $0/5$ در اجرای روش‌مان، ما کسری از انحرافات وزن را از یک تکرار قبلی اضافه می‌کنیم، همچنین جمله مقدار نیز نامیده می‌شود، برای سمت راست معادله (۳-۵) و (۳-۶) برای بهبود همگرایی در طول مدت آموزش.

جملات وزن و بایاس اول به صورت تصادفی مقادیری بین $0/5$ و $-0/5$ ارزش دهی می‌شوند. عموماً هیچ قاعده محکمی برای تعیین شکل شبکه برای آموزش و پیش‌بینی بهینه شبکه وجود ندارد. روش پیش رفتن به این شکل است که ابتدا برای آموزش شبکه، با چندین نرون در لایه میانی و به صورت دوره‌ای، توانایی شبکه برای تولید جوابهای دقیق در داده‌های آزمایشی (که داده‌هایی هستند که در آموزش استفاده نشده بودند) اندازه‌گیری می‌گردد. اگر شبکه به طور ضعیفی در هر دوی آموزش و مثالهای آزمایشی انجام شود، محتمل است که داشتن نرون‌های زیاد در لایه میانی توانایی پیش‌بینی را در داده‌های آزمایشی بهبود بخشد. به عبارت دیگر، یک خطای کوچک در داده‌های آموزشی منسوب به خطای محاسبه شده برای داده‌های آزمایشی ممکن است یک نشان از عبور از تطبیق داده‌ها باشد، همچنین ارجاعی از آموزش زیاد.

۳-۱-۴-۱- آموزش FNN بوسیله روش پس انتشار و VFSA

روش ذوب شبیه‌سازی شده خیلی سریع یک نسخه بهبود یافته از SA است. برخلاف SA استاندارد که پارامترهای آزمایشی جدید از یک توزیع یکسان انتخاب می‌شوند، VFSA آشفته‌گی‌های پارامترهای را از یک دمای وابسته توزیع متمرکز شده در اطراف پارامترهای رایج انتخاب می‌کند. آشفته‌گی‌های وزن بر پایه‌های معادلات زیر هستند:

$$w_{ij}(l+1) = w_{ij}(l) + r_{ij} \Delta w \quad (۳-۹)$$

که در آن $w_{ij} \leq w_{\max}$ ، $w_{\min} \leq w_{ij}$ ، $\Delta w = w_{\max} - w_{\min}$ ، w_{\min} و w_{\max} بیشتر به عنوان حدهای وزن تعیین می‌شوند و $r_{ij} \in [-1, 1]$ یک عدد تصادفی است. VFSA از یک محصول توابع توزیع شبه برای تولید r_{ij} استفاده می‌کند. یک عدد تصادفی u_{ij} انتخاب شده از یک توزیع یکسان $U[0, 1]$ ، می‌تواند به این توزیع با استفاده از فرمول زیر نگاشته شود:

$$r_{ij} = \operatorname{sgn}\left(u_{ij} - \frac{1}{2}\right) T' \left[\left(1 + \frac{1}{T'}\right)^{2u_{ij}-1} \right] \quad (۳-۱۰)$$

که Sgn تابع علامت است و T^1 پارامتر دما در a تکرار است. یک مقدار بالا از T^1 در شروع فرایند آن را متحمل می‌سازد که تمام فضای جستجو نمونه‌گیری خواهد شد با احتمال برابر برای هر آشفتگی هر وزن، در حالی که برای مقادیر کم T^1 ، نمونه‌گیری به نزدیکی وزنه‌های رایج تحمیل می‌شود. این نتایج دما با یک برنامه زمانی خنک‌سازی سریع بوسیله زیر کاهش می‌یابد:

$$T^i = T^0 e^{-\frac{1}{2}(i-1)} \quad \text{معادله (۱۱-۳)}$$

که T^0 دمای آغازین است. وزنه‌های جدید از فرمول (۹) پذیرفته می‌شوند، بر مبنای معیار از SA که از یک پارامتر دما استفاده می‌کند که از منحنی خنک‌سازی (۱۱-۳) استفاده می‌کند. در روش هیبرید برای آموزش شبکه، وزنه‌های آغازین با استفاده از VFSA بعد از به طور نسبی چند تکرار یا کاهش‌های دما پیدا می‌شوند. برخلاف روش گرادیان، VFSA احتیاجی به محاسبه مشتقات خطا نسبت به وزنه‌ها ندارد زیرا، جدیدسازی وزنه‌ها توسط VFSA برای یک تکرار ساده سریعتر از روش گرادیان است. هدف استفاده از VFSA بدست آوردن یک مجموعه قابل قبول از وزنه‌ها است که می‌توان سپس به عنوان نقطه آغازین برای روش پس انتشار مورد استفاده قرار گیرد. یکبار که روش پس انتشار به یک راه حل همگرا می‌شود، روش برمی‌گردد به VFSA (اگر مطلوب باشد) با هدف سوق دادن موقعیت رایج در فضای پارامتر وزن به یک خطای کوچکتر، این روش می‌تواند برای تهیه وزنه‌های بهینه در زمان محاسبات کمتری از روش گرادیان ساده، خیلی مفید باشد. عمدتاً، VFSA آشفتگی‌های وزنی را تولید خواهد کرد که اجازه می‌دهد به روش گرادیان برای گریز از مینیمم محلی.

۳-۱-۵- تعبیر و تفسیر داده‌های سونداژ مقاومت ویژه

تفسیر منحنی‌های سونداژ الکتریکی عمودی (VES)، Vertical Electrical Sounding، تخمین مقاومت ویژه را به عنوان یک تابع عمق از مقادیر صحرایی مقاومت ویژه ظاهری که برای چندین فاصله الکترودی اندازه‌گیری شده است، وارد می‌کند. شکل (۲-۳) یک منحنی صحرایی بدست آمده با یک آرایش الکترودی شلومبرژه را در یک منطقه سنگهای رسوبی با بسترهای متناوبی از ماسه و رس قرار گرفته در زیر یک لایه نازک از آبرفت، نشان می‌دهد.

وارونه‌سازی داده‌های VES عموماً توسط تطبیق کمترین مربعات و یک نسخه طولی شده معادلات حاکم انجام گردید (Glen et al. 1973; Hoversten . Dey and Morrison 1982). این روش احتیاج به یک مدل شروع برای بهینه شدن دارد. اگرچه در اغلب موقعیتهای مدل‌های شروع مختلف می‌تواند جوابهای مشابهی را تولید کند (Glenn et al. 1973)، این همیشه موردی نیست. Sen . Bhattacharya و Stoffa (۱۹۹۳) از روش بهینه‌سازی ذوب شبیه‌سازی شده برای تفسیر ترکیبی و داده‌های مقاومت ویژه ظاهری صحرایی استفاده کردند. مؤلفین نتایج خوبی را بدون نیاز به یک مدل شروع خوب گزارش دادند، اما آنها همچنین زمان محاسبه‌ای صرف شده بالایی را توسط این روش ثبت کردند. در اینجا ما کاربرد FNNها را برای مسائل وارونه‌سازی VES گزارش خواهیم داد.

۳-۱-۵-۱- آموزش FNN با داده‌های مصنوعی

مقصود اصلی از روش وارونه‌سازی مقاومت ویژه ارائه شده در اینجا، ابتدا آموزش یک F NN با منحنی‌های مقاومت ویژه مصنوعی است و سپس آزمایش شبکه آموزش یافته با داده‌های مقاومت ویژه جدید است. مسأله تولید داده‌های مصنوعی بموقع خوب فهمیده شد (Ghosh 1971; Koefoed 1979) و آن در اینجا بحث نخواهد شد. برطبق نشانه‌گذاری ارائه شده در بخش اخیر، بردار ورودی توسط N مقدار مقاومت ویژه ظاهری به عنوان $x = [\rho_{a1}, \rho_{a2}, \dots, \rho_{aN}]^T$ داده شده است، در حالیکه بردار خروجی توسط پارامترهای مدل m لایه به عنوان $y = [\rho_1, \rho_2, \dots, \rho_m, n_1, n_2, \dots, n_{m-1}]^T$ شامل $M=(2m-1)$ پارامترهای مدل داده شده است. ρ_i و h_i مقاومت ویژه صحیح و ضخامت لایه i هستند.

شکل (۳-۳) a یک گروه از منحنی‌های VES مصنوعی نمونه‌گیری شده در ۲۰ فاصله الکترودی بدست آمده از مدل‌های سه لایه‌ای با پارامترهای بطور تصادفی انتخاب شده با مقادیر حدی داده شده در جدول (۱-۳) را نشان می‌دهد. این منحنی‌ها مطابق با منحنی‌های نوع H هستند (Bhattacharya and Patra 1968). مقاومت ویژه لایه اول برای همه مدلها در Ωm قرار می‌گیرد، به عبارت دیگر، مقاومت ویژه‌ها برای لایه‌های دوم و سوم توسط مقاومت ویژه لایه اول مقدار دهی می‌شود. سپس، ما فرض خواهیم کرد که مقاومت ویژه لایه اول می‌تواند از

اندازه‌گیریهای فواصل الکتروودی که نسبت به ضخامت لایه اول کوچک هستند بدست آید. بنابراین، در این مجموعه از مدلها، پارامترهایی که تغییر می‌کنند مقاومت ویژه‌های لایه‌های دوم و سوم، و ضخامت لایه‌های اول و دوم هستند.

یک FNN دو لایه‌ای با ۸ نرون در لایه میانی با استفاده از روش پس‌انتشار آموزش داده شده بود. یک مجموعه کامل از ۱۰۰ منحنی VES مصنوعی، مانند آنهایی که در شکل (۳-۳ a) نشان داده شد، با تطبیق پارامترهای لایه خودشان برای آموزش شبکه استفاده شده بودند. شکل (۳-۴) پیشینه خطای بدست آمده در طول آموزش شبکه را نشان می‌دهد. خطای رسم شده با خطای انباشته شده تقسیم شده توسط عدد مثالهای آموزش بدست آمده از مقایسه مدل‌های صحیح با خروجی‌های شبکه، مطابقت می‌کند. شکل‌های (۳-۳ b) و (۳-۳ c) قطعات رنگی توزیع دامنه لایه میانی آخری و خروجی بترتیب ماتریسهای W^0 و W^h را نشان می‌دهد. این پلانها نشان می‌دهند که چگونه ساختارهای بردارهای داده‌های ورودی به فضای وزن نگاشت می‌شود. نگاشت دامنه شکل (۳-۳ b) اشاره بر این می‌کند که بعد از آموزش، تعدادی از وزنها در لایه میانی یک پاسخ بزرگ برای داده‌های ورودی دارد و چیزی که آنجا همبستگی زیادی بین قطعه رنگی برای W^h و منحنی‌های نشان داده شده در شکل (۳-۳ a) وجود دارد. مقدارهای بزرگ برای وزنها در شروع منحنی و در مرکز مشاهده می‌شوند، که می‌تواند این مفهوم را داشته باشد که جدایی الکتروودها از ۲ تا ۶ متر و از ۳۰ تا ۷۰ متر خیلی برای تغییرات در پارامترهای مدل حساس هستند.

۳-۱-۵-۲- اثر اندازه FNN در پیش‌بینی

همگرایی خطا در طول آموزش شبکه بستگی به تعداد نرون‌ها در لایه میانی شبکه دارد. مهمتر از کاهش خطا، توانایی شبکه برای پیش‌بینی موفق جوابها از داده‌های جدید است که بستگی دارد به هر دوی تعداد نرون‌های معرف شبکه و تعداد مثالهای استفاده شده برای آموزش. برای پیدا کردن عدد بهینه نرونها و اندازه مجموعه آموزشی برای این کاربرد ویژه، NNها با ۶، ۸ و ۱۲ نرون در لایه میانی بترتیب با ۵۰، ۱۰۰ و ۲۵۰ مثال آموزشی آموزش یافته بودند. هر شبکه برای ۲۰۰۰۰ تکرار آموزش یافته بود. برای اجتناب از مینیمم محلی کم عمق که ممکن است یک اجرای ضعیف از شبکه بدهد، هر شبکه ۱۰ بار آموزش یافته بود و تنها شبکه‌ای که کوچکترین خطا را داد

ذخیره شده بود. شبکه‌های آموزش یافته سپس در ۱۰۰ مثال جدید بطور تصادفی انتخاب شده از یکدسته مشابه از مدل‌هایی که برای آموزش استفاده شده بودند، آزمایش شدند. برای هر دوی آموزش و پیش‌بینی مثالها، یک خطا برای هر پارامتر لایه محاسبه شده بود. این خطاها مطابق بود با انحراف معیارهای بدست آمده از مقایسه درست و پارامترهای تقریب زده میانگین شده بر روی تعداد آموزش یا پیش‌بینی مثالها.

جداول (۲-۳) و (۳-۳) آموزش و پیش‌بینی نتایج بدست آمده با استفاده از FNN های آموزش یافته با ۶ و ۸ نرون بترتیب در لایه میانی را خلاصه می‌کند. از جدول ۳، مشاهده می‌شود که خطای آموزش اندکی افزایش یافته وقتی که اغلب مثالها به آموزش پیوسته شده بودند در حالیکه خطای پیش‌بینی کاهش یافته. یک اندازه آموزش از ۲۵۰ مثال نتایج بهتری در جملات خطای پیش‌بینی تولید کرد. هنگام آموزش یک شبکه با ۸ نرون (جدول ۳-۳)، خطاهای آموزش نهایی برای اندازه‌های مختلف مثالهای آموزش شبیه بودند در حالیکه خطای پیش‌بینی کاهش پیدا کرد بطوریکه بیشتر مثالها به آموزش پیوسته بودند. همچنین توجه کنید که، در حالیکه مقاومت ویژه‌های لایه‌های دوم و سوم با دقت یکسان بدست آمده بودند (با تردید حدود $0.15 \text{ m } \Omega$)، تخمین‌های ضخامت لایه دوم دقت کمتری از تخمین‌های لایه اول داشتند. این مربوط به حقیقتی است که حساسیت داده‌ها برای تغییرات ضخامت برای لایه دوم خیلی کوچکتر از آن برای لایه اول است. این در هر دوی نتایج آموزش و پیش‌بینی شرح داده شده در جدول (۳-۳)، منعکس شده است. نتایج از آموزش شبکه با بیش از ۲۵۰ مثال (نتایج در جدول (۳-۳) نشان داده نشده است) پیش‌بینی را بیشتر از این بهبود نمی‌بخشد. سرانجام، آموزش یک شبکه با ۱۲ نرون در لایه میانی (نتایج نشان داده نشده) یک بهبود اساسی را در تخمین‌های بدست‌آمده توسط شبکه ۸ - نرونی تولید نمی‌کند. از این آزمایشات، ما نتیجه می‌گیریم که شبکه بهینه برای این مسأله ویژه، ۸ نرون دارد و بایستی با ۲۵۰ یا مثالهای داده‌های بیشتری آموزش یابد.

داشتن یک عدد بزرگ از وزنها و چند مثال برای آموزش معادل است با داشتن یک مسأله وارونه که تحت تعیین است. شبکه به دقت داده‌های را تطبیق می‌دهد اما قادر به پیش‌بینی جوابهای صحیح از داده‌های جدید نیست. اگر تعداد وزنها خیلی کوچک باشد، شبکه یک کارایی ضعیف در هر دوی آموزش و پیش‌بینی خواهد داشت. همچنین، زمانی را که آن برای آموزش یک شبکه می‌گیرد با بزرگتر شدن شبکه افزایش می‌یابد زیرا عدد محاسبات متناسب عدد وزنها است و

چون، عموماً، یک شبکه داشتن نرونهای زیاد مثالهای آموزشی زیاد و تکرارهای آموزشی زیادی را لازم دارد.

نتایج پیش‌بینی بدست آمده با شبکه ۸ نرونی آخری در شکل (۵-۳) نشان داده شده‌اند. سه مدل مقاومت ویژه بدخواه انتخاب شده از نتایج پیش‌بینی برای محاسبه منحنی‌های VES ترکیبی استفاده شده بودند. منحنی‌های صحیح و محاسبه شده در شکل (۵-۳) a) نشان داده شده‌اند و مدل‌های صحیح و تخمین زده شده در شکل (۵-۳) b) نشان داده شده‌اند. اگرچه تطبیق بین داده‌های بدست آمده از مدل‌های تقریبی و داده‌های صحیح کامل نیست، مدل‌های نهایی با تغییرات در مقاومت ویژه و ضخامت لایه نامتناقض هستند.

۳-۵-۱-۳- آموزش FNN و پیش‌بینی با منحنی‌های VES چهار لایه‌ای

آنالیز بالا می‌تواند برای تفسیر منحنی‌های VES با بیشتر از چهار لایه دنبال شود. اندازه مجموعه آموزشی برای بدست آوردن نتایج بهینه ممکن است برای این مورد رشد کند، از آنجاییکه یک فضای مدل بزرگتر برای نمونه‌گیری اجبار شده است. برای آزمایش این، آموزش شبکه با منحنی‌های چهار لایه‌ای فرض می‌شود. مدل‌های جدید به عنوان منحنی‌های نوع HK فرض می‌شوند. $(\rho_1 > \rho_2 < \rho_3 > \rho_4)$. حدها برای هر پارامتر مدل شبیه آنهایی هستند که در مثال قبلی و در جدول (۴-۳) داده شده‌اند.

نتایج آموزش و پیش‌بینی هنگام آموزش یک شبکه با ۱۲ نرون در لایه میانی برای ۱۰۰، ۲۵۰ و ۴۰۰ مثال در جدول (۵-۳) داده شده‌اند. اگرچه خطاهای آموزش برای اندازه‌های مختلف مجموعه آموزشی شبیه هستند، خطای پیش‌بینی کاهش می‌یابد بطوریکه عدد مثالهای آموزش رشد می‌کند، با کوچکترین خطاهای تولید شده توسط شبکه آموزش یافته با ۴۰۰ مثال. از جدول (۳-۳) و (۵-۳)، انحراف معیارها برای مقاومت ویژه‌های لایه‌های اول و دوم برای مثالهای منحنی سه لایه و چهار لایه مشابه هستند. تقریبهای مقاومت ویژه برای لایه سوم با کوچکترین خطاها در مورد سه لایه‌ای نسبت به مورد چهار لایه‌ای بدست آمده بودند. همچنین، خطاها برای ضخامت‌های لایه‌های دوم و سوم بزرگتر از آن برای لایه اول هستند. این مربوط به حساسیت ضعیفتر داده‌ها به این پارامتر می‌شود. شکل‌های (۳-۲۶) و (۳-۲۶b) نتایج پیش‌بینی بدست‌آمده برای ۳ منحنی چهار

لایه‌ای بدلیخواه انتخاب شده را از مجموعه پیش بینی نشان می‌دهد. نتایج برای منحنی‌های چهار لایه‌ای قابل مقایسه با نتایج بدست آمده از منحنی‌های سه لایه‌ای شرح داده شده در شکل‌های (۳-۳) و (۳-۴) هستند.

۳-۱-۵-۴- تفسیر داده‌های صحرایی VES با FNNهای آموزش یافته

تفسیر منحنی‌های VES صحرایی با روش پیشنهاد شده، یک مجموعه از FNNهای آموزش یافته بوسیله مدل‌های مقاومت ویژه با حدهای پارامتری مدل متفاوت در هر یک از شبکه‌ها، لازم دارد. ویژگی‌های مجموعه منحنی‌های VES استفاده شده برای آموزش شبکه بایستی بستگی به زمین‌شناسی منطقه ویژه تحت مطالعه داشته باشد. برای سهولت آموزش شبکه، مدل‌های با چند لایه (۳-۵) بایستی در نظر گرفته شوند. داده‌های صحرایی شکل (۳-۲)، ورودی برای FNN با ۸ نرون در لایه میانی بودند و با ۲۵۰ مثال آموزش یافته بودند (جدول ۳-۳ را ببینید). پارامترهای مدل تخمین زده شده به صورت زیر می‌باشند:

- $\rho_1 = 1.00 \Omega m, h_1 = 3.81 m,$
- $\rho_2 = 0.090 \Omega m, h_2 = 18.96 m,$
- $\rho_3 = 0.454 \Omega m.$

شکل (۳-۷) یک مقایسه از داده‌های مشاهده شده با داده‌های محاسبه شده را نشان می‌دهد و شکل (۳-۷) مدل مقاومت ویژه تخمین شده را نشان می‌دهد. توجه کنید که مقاومت ویژه ظاهری در فواصل الکترونی متوسط غلو آمیز برآورد شده در حالیکه در لبه‌های منحنی تقریباً یک تطبیق کامل بین داده‌های مشاهده شده و محاسبه شده وجود دارد. برای بهبود این نتیجه، مدل برآورد شده به عنوان مدل شروع برای وارونه‌سازی خطی شده حداقل مربعات با روش نیوتن کمینه‌سازی، استفاده شده بود. روشی که ما استفاده کردیم یک نسخه ساده شده روش نیوتن است (Sen and Stoffa 1995, P.64) که بطور عددی مشتق دوم تابع هدف را نسبت به پارامترهای مدل محاسبه می‌کند. خطای زیرین کمینه شده بود:

$$E = \frac{1}{N} \sum_{i=1}^N (\rho_{a_i}^o - \rho_{a_i}^n)^2. \quad (۳-۱۲)$$

که در آن ρ_a^o و ρ_a^n مقادیر مقاومت ویژه ظاهری ترکیبی و مشاهده شده هستند. مدل نهایی بدست آمده بعد از سه تکرار به شکل زیر است:

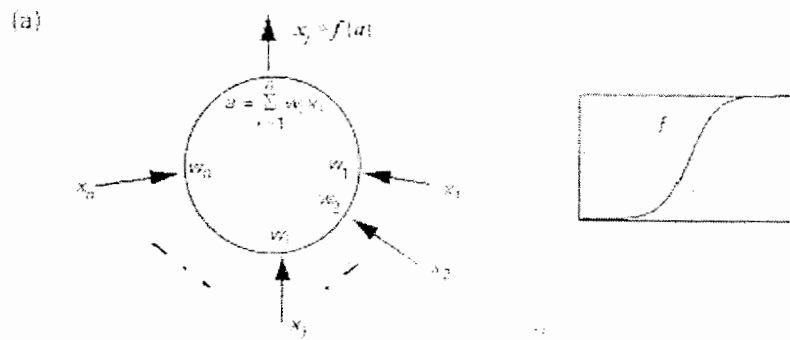
- $\rho_1 = 0.982\Omega m, h_1 = 3.92m.$
- $\rho_2 = 0.080\Omega m, h_2 = 21.79m.$
- $\rho_3 = 0.438\Omega m.$

توجه داشته باشید که مقاومت ویژه لایه دوم در مقایسه با برآوردهای شبکه کاهش می‌یابد و ضخامت لایه دوم توسط تقریباً ۳ متر افزایش می‌یابد. فرض نتیجه وارونه‌سازی خطی شده به عنوان مدل " صحیح "، تفاضلات بین برآوردهای مقاومت ویژه برای همه لایه‌ها بخوبی ضخامت لایه اول درون برآوردهای خطای جدول (۳-۳) هستند. همچنین، توجه داشته باشید که اختلاف بین برآوردهای شبکه و مدل " صحیح " برای ضخامت لایه دوم دو برابر شد وقتی که با برآوردهای خطای داده شده در جدول (۳-۳) مقایسه شد. ما این را توسط این حقیقت که ضخامت " صحیح " برای لایه دوم، بیرون از حدود مدل‌های استفاده شده برای آموزش شبکه است، توضیح می‌دهیم (حد ماکزیمم ۲۰ متر بود). با وجود این، شبکه قادر به بدست آوردن برآوردهای دقیق برای پارامترهای دیگر بود. منحنی محاسبه شده جدید و مدل برآورد شده بدست آمده توسط روش نیوتن بعد از تفسیر منحنی با FNN در شکل‌های (۳-۸ a و b) نشان داده شده‌اند.

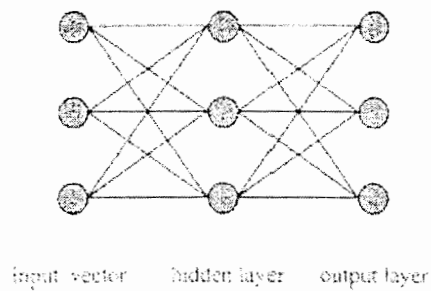
برای ارزیابی نتیجه بدست آمده توسط FNN ترکیب شده و تکرار برنامه حدافل مربعات، تفسیر منحنی با روش نیوتن انجام پذیرفته بود، با شروع در یک نقطه دلخواه درون حدود پارامتر مدل استفاده شده برای آموزش شبکه. روش نمی‌تواند به یک مدل مناسب همگرا شود و خیلی نزدیک به مدل شروع خاتمه یابد. یک تحلیل از سطوح خطا برای ترکیبات مختلف پارامترهای مدل برای یافتن اجسام هندسی از داده‌های پروفیل مقاومت ویژه توسط (Chunduru et al. (1995 انجام شده بود. در کار آنها، آنها نشان دادند که اگرچه برخی از پارامترهای مدل یک مینیمم منفرد خوب تعریف شده داشت، تغارهای چندگانه در برخی از سطوح خطا مشاهده شده بودند. آنها اشاره داشتند که یک اشکال از برنامه گرادیان برای حل چنین مسائلی طبیعت حریص برنامه خطی شده بود. در این مورد، یک FNN آموزش یافته می‌تواند برای تهیه یک جواب که نزدیکتر به راه حل صحیح است، استفاده گردد.

۳-۱-۶- بحث و نتیجه گیری

ما کاربرد FNNها را برای حل یک مسأله وارونه ژئوفیزیکی ارزیابی کردیم. مسأله مربوط به تخمین یک بُعدی مدل‌های عمق مقاومت ویژه از منحنی‌های مقاومت ویژه ظاهری می‌شود. آموزش یک FNN با منحنی‌های VES، یک فرآیند سریع اصلاً مربوط به نسبتاً چندین نمونه است که معمولاً این منحنی‌ها را تعریف می‌کند، و همچنین مربوط به سادگی عددی محاسبه داده‌های ترکیبی می‌شود. در این کاربرد، روش گرادیان برای جدید سازی وزن استفاده شده بود برای بدست آوردن وزنهای بهینه. برای گریز مینیمم محلی کم عمق، فرآیند آموزش یک عدد از دفعات با وزنهای شروع متفاوت را تکرار کرده بود. نقشه‌های رنگی توزیع دامنه ماتریس‌های وزن نشان می‌دهند که تنها یک عدد کوچک از این وزن‌ها، نسبتاً مقادیر بزرگی دارد. این به این معنی است که تنها یک بخش محدود از منحنی‌های VES در یک روش مهم برای نگاشت سهم دارد. یک توصیف از داده‌ها در جمله‌های یک چند نقطه بحرانی تعریف ویژگی‌های اصلی داده‌ها ممکن است اساساً سرعت فرایند آموزش را بالا ببرد. فایده اصلی استفاده از FNNها برای تفسیر منحنی، یکی این است که FNN آموزش داده می‌شود، آن توانایی تفسیر اندازه‌گیری‌های خیلی سریع صحرائی را دارد. در روش تفسیر منحنی با یک FNN، توسط وارونه‌سازی خطی شده دنبال می‌شود، انتظار می‌رود که FNN داده‌ها را وارونه کند برای یک نقطه در چیزی که برنامه خطی شده می‌تواند تولید کند یک چندین مدل دقیق درون یک چند تکرار. [۸]



(b)

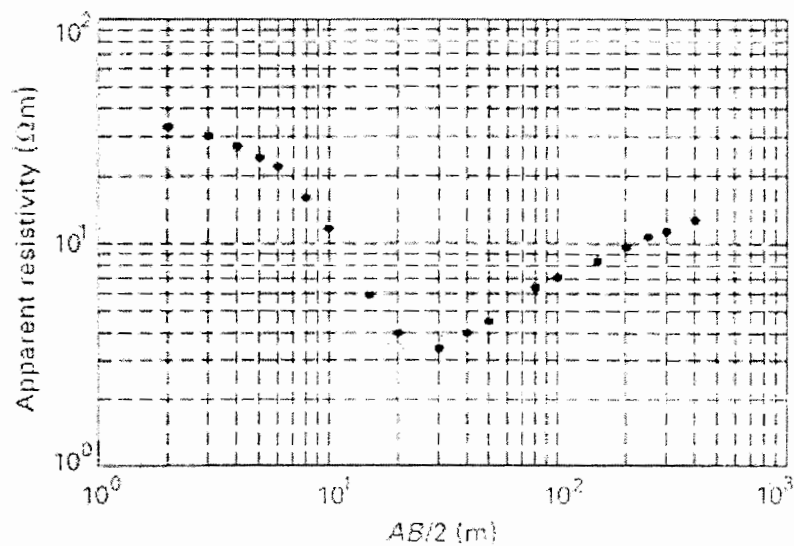


شکل ۱-۳: (a) مدل ریاضی یک نرون ساده و تابع محرک سیگموئید. (b) FNN دولایه. تعداد نرونها در لایه میانی مورد لزوم برای بدست آوردن قابلیت یادگیری و عمومیت دادن، مربوط به مساله است. تعداد نرونها در لایه خروجی برابر با تعداد پارامترهای مطلوب از بردار خروجی است.

جدول ۱-۳: فواصل پارامترهای مدل استفاده شده برای تولید آموزش و پیش بینی برای مدل‌های مقاومت ویژه سه لایه‌ای و منحنی‌های متناظر

Table 1. Model parameter intervals used to generate training and prediction for three-layer resistivity models and the corresponding curves.

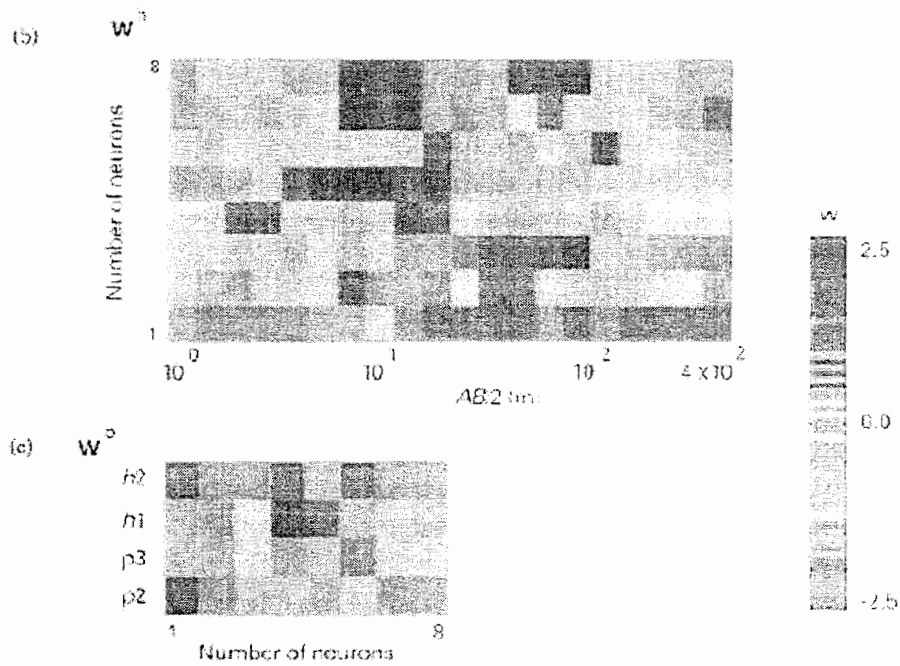
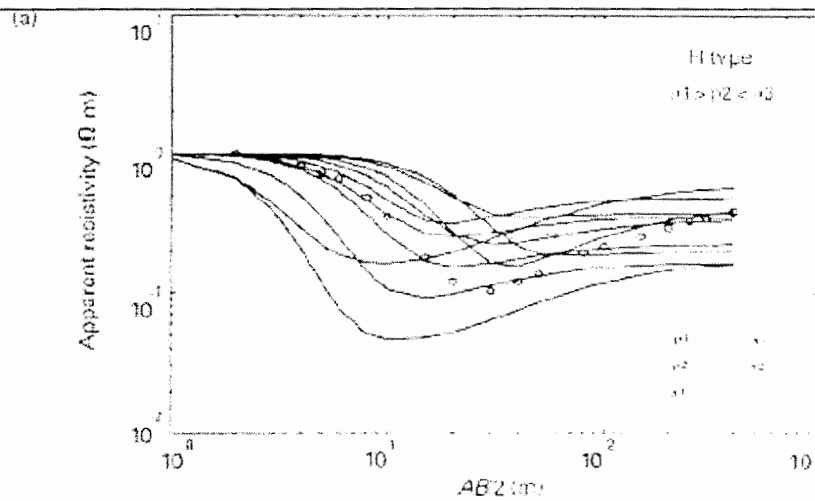
Layer no.	ρ_{min} (Ωm)	ρ_{max} (Ωm)	h_{min} (m)	h_{max} (m)
1	1.00	1.00	1	10
2	0.03	0.20	3	20
3	0.15	0.61	-	-



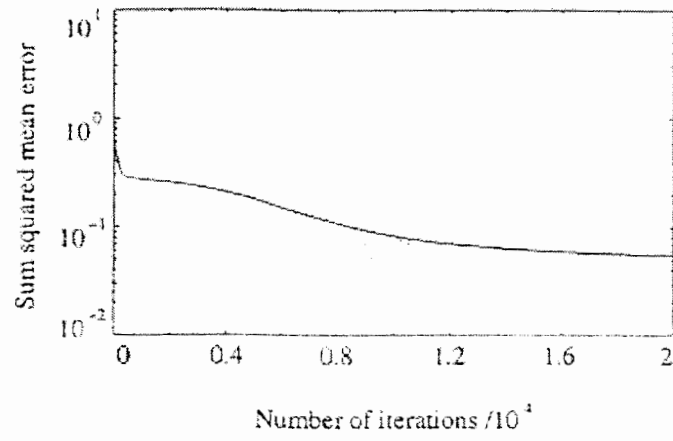
شکل ۳-۲: منحنی مقاومت ویژه ظاهری بدست آمده با آرایش الکترودی شلومیرژه، AB مربوط به فاصله افقی الکترودی می‌شود.

جدول ۳-۲: نتایج پیش بینی و آموزش بدست آمده با استفاده از یک FNN با ۶ نرون در لایه میانی، شبکه با ۵۰، ۱۰۰ و ۲۵۰ منحنی سه لایه‌ای آموزش دیده است. جدول تخمین‌های خطا را برای هر پارامتر مدل برای آموزش به خوبی نمونه‌های پیش بینی را ارائه می‌دهد. ردیف آخر مربوط به مربع جمع آخری اندازه گیری شده توسط تعداد نمونه‌های استفاده شده برای آموزش شبکه است. T، آموزش؛ P، پیش بینی.

	Size of training set					
	50		100		250	
	T	P	T	P	T	P
ρ_2 (Ωm)	0.014	0.039	0.014	0.019	0.014	0.010
ρ_3 (Ωm)	0.025	0.029	0.014	0.018	0.019	0.020
h_1 (m)	0.652	0.764	0.590	0.649	0.230	0.242
h_2 (m)	1.261	3.895	1.613	2.682	1.977	2.053
Sum sq. error	0.068	0.407	0.077	0.166	0.104	0.119



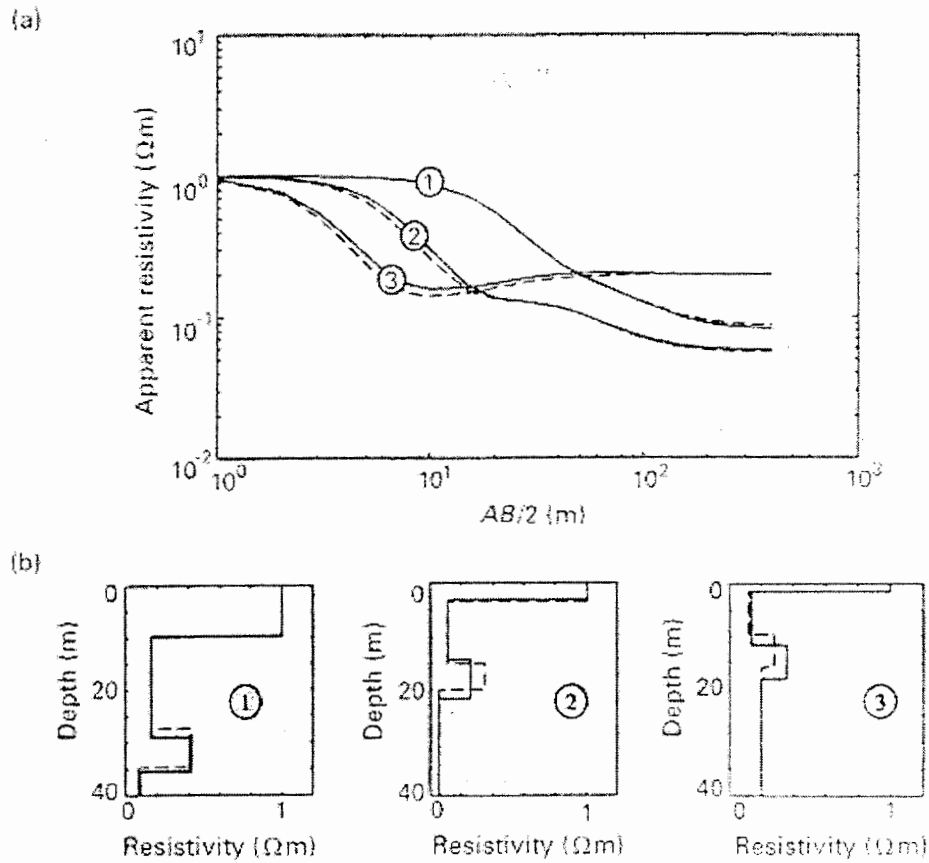
شکل ۳-۳: (a) منحنی‌های مقاومت ویژه بدست آمده از یک مجموع مدل‌های تصادفی (جدول ۳-۱ را ببینید). منحنی‌ها برای ۲۰ فاصله الکترودی محاسبه شده‌اند. منحنی صحرائی توسط دایره‌های خالی نشان داده شده است. (b) قطعه رنگی توزیع دامنه ماتریس لایه میانی را نشان می‌دهد. (c) قطعه رنگی توزیع دامنه ماتریس لایه خروجی را نشان می‌دهد.



شکل ۳-۴: پیشینه خطای آموزش. خطای اولیه ۲/۳ است.

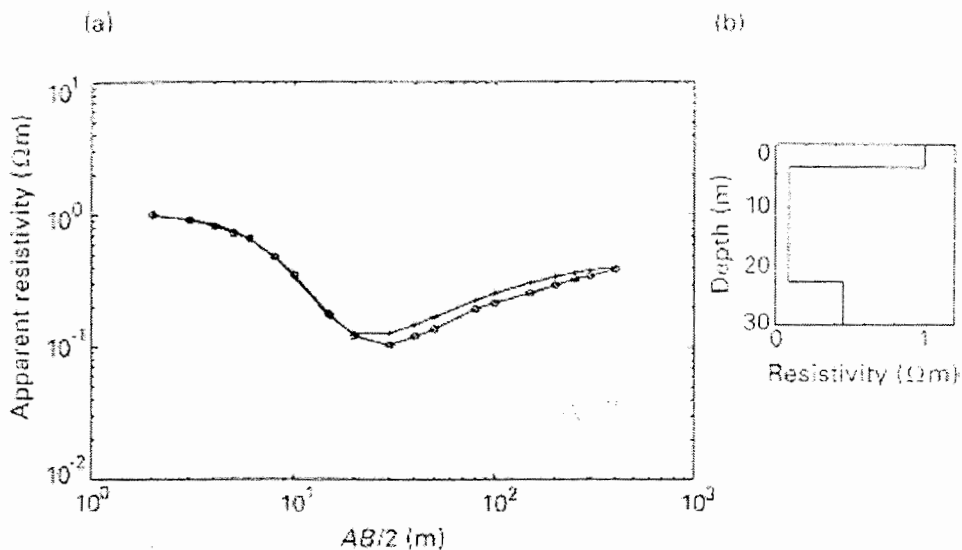
جدول ۳-۲: نتایج پیش بینی و آموزش بدست آمده با یک FNN با ۸ نرون در لایه میانی برای مدل‌های مقاومت ویژه سه لایه . T، آموزش؛ P، پیش بینی.

شکل ۳-۵: (a) مقایسه بین منحنی‌های VES محاسبه شده از مدل مقاومت ویژه سه لایه‌ای از مجموعه پیش بینی. خط تیره نمایانگر داده‌های ورودی شبکه و خط پر نشانگر داده‌های مصنوعی بدست آمده از خروجی شبکه است. (b) مدل‌های مقاومت ویژه تخمین زده شده و حقیقی.



جدول ۳-۴: فواصل پارامتری مدل استفاده شده برای تولید آموزش و پیش بینی مدل‌های مقاومت ویژه چهار لایه‌ای و منحنی‌های مربوطه

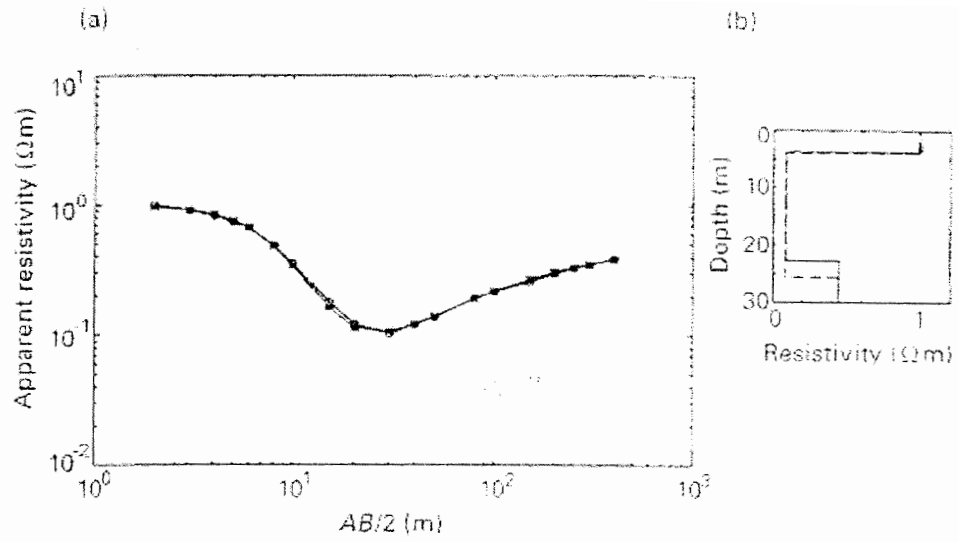
شکل ۳-۶: (a) مقایسه بین منحنی‌های VES محاسبه شده از مدل مقاومت ویژه چهار لایه‌ای از مجموعه پیش بینی. خط تیره‌ها نشانگر داده‌های ورودی شبکه و خط پر نشانگر داده‌های مصنوعی بدست آمده از خروجی شبکه هستند. (b) مدل‌های مقاومت ویژه تخمین زده شده و حقیقی.



شکل ۷-۲: مقایسه بین منحنی VES صحرایی (O) و منحنی مصنوعی (+) محاسبه شده از ۸ نرون خروجی‌های FNN وقتی که داده‌های صحرایی به عنوان ورودی شبکه استفاده شده‌اند. (b) تخمین مدل FNN.

جدول ۵-۲: نتایج آموزش و پیش‌بینی بدست آمده با یک FNN با ۱۲ نرون در لایه میانی برای مدل‌های مقاومت ویژه چهار لایه‌ای. T، آموزش؛ P، پیش‌بینی.

	Size of training set					
	100		250		400	
	T	P	T	P	T	P
ρ_2 (Ωm)	0.009	0.012	0.009	0.014	0.007	0.009
ρ_3 (Ωm)	0.036	0.066	0.033	0.037	0.034	0.040
ρ_4 (Ωm)	0.008	0.011	0.007	0.008	0.004	0.004
h_1 (m)	0.170	0.196	0.216	0.212	0.256	0.255
h_2 (m)	1.387	2.377	1.428	1.702	1.392	1.512
h_3 (m)	0.961	1.252	0.977	0.970	0.969	0.894
Sum sqr. error	0.083	0.210	0.100	0.119	0.091	0.108



شکل ۸-۳: مقایسه بین منحنی VES صحرائی (O) و منحنی مصنوعی (*) محاسبه شده از مدل مقاومت ویژه- عمق بدست آمده با استفاده از وارونه‌سازی مکرر خطی. خروجی شبکه (شکل ۳-۵b) به عنوان مدل شروع برای وارونه‌سازی خطی استفاده شده است. (b) مدل مقاومت ویژه- عمق نهایی

۳-۲- آشکارسازی صفحه ضخیم هادی توسط روش عنصر محدود و

شبکه‌های عصبی^۲

۳-۲-۱- خلاصه

آشکارسازی دایک در زمین فرو رفته با تباین قابلیت هدایت، یک وظیفه مهم برای ژئوفیزیک کاربردی است. تخمین طول در عمق از یک جسم معدنی بزرگ امتداد یافته، به ویژه برای روش‌های مقاومت ویژه سطحی مشکل است زیرا آنومالی آشکار شده ضعیف است. در این تحقیق، مدل‌سازی عنصر محدود و روش شبکه عصبی برای مطالعه مساله صفحه فرو رفته ترکیب شدند. نیم فضا با یک صفحه ضخیم دفن شده در نظر گرفته شده است. یک جفت خط الکترودهای جریان روی زمین به صورت ثابت برای تولید یک میدان الکتریکی 2D به کار گرفته شده‌اند. منحنی‌های مقاومت ویژه ظاهری برای آرایه دو قطبی - دو قطبی توسط روش عنصر محدود محاسبه شده‌اند و برای موارد (i) صفحه با یک بعد ثابت و اما مقاومت ویژه‌ای متغیر (ii) یک صفحه هادی ثابت با طول‌های در عمق متغیر. اثرات دوباره همچنین در نظر گرفته شده‌اند. یک شبکه عصبی پس انتشار سه لایه برای وارونه‌سازی پارامتری ژئومتری و فیزیکی صفحه ایجاد شده است. دقت میانگین تخمین طول ۲۰۱٪ است و در حالیکه به ترتیب در ارزیابی مقاومت ویژه، ۵٪ برای صفحه هادی و ۲/۸٪ برای صفحه مقاوم.

۳-۲-۲- مقدمه

روش الکتریکی در حال حاضر به طور فزاینده‌ای اکتشافات معدنی، بررسی‌های ژئوترمال، مهندسی و نقشه‌های زمین‌شناسی محیط‌زیستی مورد استفاده قرار می‌گیرد. مطالعات بیمشار مدل‌سازی وارنه و مستقیم VES بر پایه راه حل لابلای برای مساله یک منبع جریان نقطه‌ای روی نیم فضای چند لایه‌ای افقی بنا نهاده شده‌اند. قبلاً تفسیر مقاومت ویژه بر پایه راه حل‌های تحلیلی برای موارد یک جسم کروی هادی مدفون بنا نهاده شده بود. به عبارت دیگر، مدل‌سازی فیزیکی، کمک زیادی برای تفسیر ساختار زمین‌شناسی پیچیده، فراهم می‌کند اما نمی‌تواند یک راه حل

^۲ - مطالب این بخش از منبع [۳۴] گردآوری شده است.

کمی، به علت محدودیتهای آن روی بعد مدل و پارامترهای فیزیکی مواد ارائه دهد. از اینرو مدلسازی عددی یکی از قویترین ابزار در ژئوفیزیک اکتشافی شد. فقط تعدادی را نام می‌بریم، برای مساله الکتریکی 2D، راجح حل‌های عنصر محدود و تفاضل محدود وجود دارند؛ و برای 3D، راه حل معادلات انتگرال، تفاضل محدود و عنصر محدود وجود دارد. روش معادلات انتگرال، برای مسائل با ناهمگنی‌های کوچک موثرتر است و کاربردهایی برای مسائل عملی قطعی دارد، مثلاً، روش و روش‌های الکتریکی. به طور نظری، راه حل یک منبع جریان در نیم فضای چند لایه‌ای؛ با ناهمگنی‌های کوچک توسط روش معادلات انتگرال می‌تواند به دست آمده باشد. به هر حال، روش معادلات انتگرال می‌تواند تنها مسائلی با مدل ژئومتریک نسبتاً ساده حل کند. برای مطالعه اثرات برقی برای پروفیل الکتریکی پیچیده، ما برنامه‌های عنصر محدود 2D و 3D را برای محاسبه مقاومت ویژه ظاهری برای مسائل زمین شناسی پیچیده ایجاد کردیم. با استفاده از روش عنصر محدود و تفاضل محدود 3D برای حل مسائل ارائه شده توسط ساختارهای زمین شناسی واقعی همیشه حافظه و زمان اجرای قابل توجه در محاسبات لازم است. در این تحقیق، ما مدلسازی عنصر محدود 2D را با شبکه عصبی برای مطالعه مسائل صفحه دفن شده ترکیب می‌کنیم.

ما از یک جفت الکترودهای خطی برای تولید تقریبی یک میدان جریان الکتریکی 2D استفاده می‌کنیم. (شکل ۳-۹ و شکل ۳-۱۰). در این مقاله ما تنها نتایج سونداژ دو قطبی - دو قطبی را برای نشان دادن یک صفحه دفن شده در یک فضای همگن ارائه می‌دهیم (با یا بدون یک روبراه). نتایج مدلسازی مستقیم می‌تواند برای تهیه اطلاعات برای تفسیر صحرائی استفاده شود. عمق سطح بالای صفحه ممکن است توسط روش‌های مقاومت ویژه مرسوم با اندازه‌گیری‌های سطحی تعیین شود. بنابراین ما برای تخمین متمرکز می‌شویم روی (i) مقاومت ویژه، (ii) طول عمق صفحه. مساله برای عمق اکتشاف بیشتر مورد بررسی قرار می‌گیرد زیرا آرایه گرادیان، آنومالی‌های ضعیف برای ساختارهای شیب تند فرو رفته حاصل می‌دهد. یک الگوریتم شبکه عصبی پس انتشار سه لایه برای بررسی مسائل وارونه بسط داده شده است. به طویلی اطلاعات اولیه ناحیه بررسی که ما داشتیم ما نمی‌توانیم یک سری از داده‌ها را با مقاومت ویژه گوناگون یا درازاهای گوناگون از صفحه توسط مدلسازی عنصر محدود برای آموزش شبکه عصبی تولید کنیم. بعد از یادگیری، درازا و مقاومت ویژه‌های صفحه از نتایج اندازه‌گیری صحرائی ممکن است توسط شبکه عصبی ارزیابی شود.

۳-۲-۳- طراحی آرایش

شکل (۳-۹)، یک تولید جریان الکتریکی دو قطبی D 2 در صحرا را توسط کار گذاری یک سری از الکترودها در طول جهت امتداد ساختار، شرح می‌دهد. پروفیل آرایش سونداژ دو قطبی - دو قطبی، عمود بر جهت امتداد ساختار است.

شکل (۳-۱۰) یک شکل روی سطح آب را نشان می‌دهد. منبع دو قطبی خطی روی سطح آب جایگذاری شده‌اند یا در ته لایه آب ثابت گذاری شده‌اند.

۳-۲-۴- نتایج مدلسازی عنصر محدود (Finite Element)

شکل (۳-۱۱) تا شکل (۳-۱۴) نتایج محدودی را برای آرایش سونداژ دو قطبی - دو قطبی شرح می‌دهد.

شکل (۳-۱۱) مقاومت ویژه ظاهری روی یک صفحه دو بعد با مقاومت ویژه‌های گوناگون (ohm-m) ۵۰۰-۵۰ در یک عمق ۱۰ واحدی را نشان می‌دهد. بعدها صفحه ۱۰ واحد در عرض و ۳۰ واحد در عمق هستند. مقاومت ویژه سنگ میزبان hm-m ۱۰۰ o است.

شکل (۳-۴) شرح می‌دهد که چطور تغییرات طول صفحه می‌تواند بر پروفیل مقاومت ویژه اثر بگذارد. به طوری که با افزایش طول، مقاومت ویژه‌های ظاهری کاهش می‌یابند.

شکل (۳-۱۳) و شکل (۳-۱۴) منحنی‌های مقاومت ویژه ظاهری را برای یک صفحه هادی دفن شده با یک روباره مقاوم از ۵ تا ۱۰۰ hm-m o؛ و یک روباره هادی با مقاومت ویژه به ترتیب ۲۰۰ تا ۵۰۰ ohm-m را نشان می‌دهد.

۳-۲-۵- وارونه‌سازی شبکه عصبی

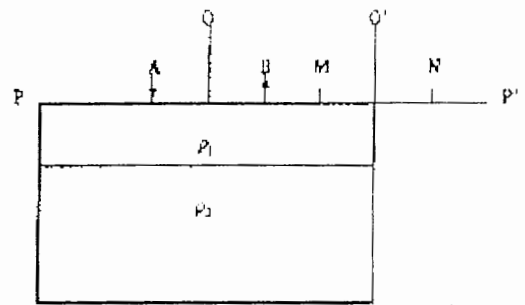
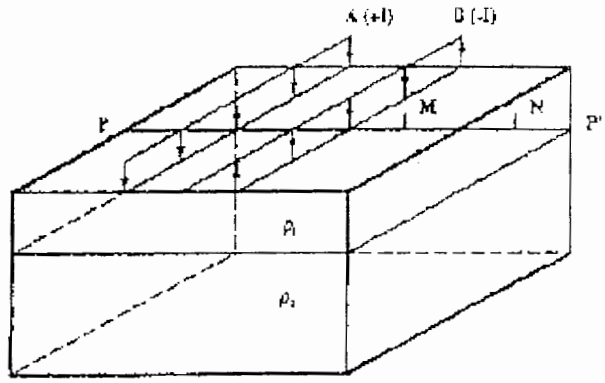
این مقاله، نگاشت توابع پیوسته بین دو فاصله بسته را مطرح می‌کند. این روش از اغلب کاربردهای شبکه عصبی برای تشخیص الگو در مشخصه‌های ژئوفیزیکی متفاوت است (Murat and Rudman, 1992; McCormack, et al., 1993) که با نگاشت کدهای دیجیتالی گسسته سر و کار دارد. یک شبکه عصبی سه لایه برای وارونه‌سازی پارامترهای صفحه استفاده شده است. طراحی داده‌های آموزشی بر پایه اطلاعات قبلی منطقه مورد تحقیق بنا نهاده شده است. مدلسازی عنصر

محدود می‌تواند یک سری از داده‌ها برای آموزش شبکه تولید کند. داده‌ها با پارامترهای نامشخص می‌تواند به عنوان داده‌های ورودی در شبکه آموزش یافته برای وارونه‌سازی پارامترها استفاده شود. شکل (۳-۱۵) نشان می‌دهد که چطور تعداد تکرارها شدیداً بر روی خطای میانگین شبکه به طوری که در جدول ۱ نشان داده شده است، تاثیر می‌گذارد. وقتی که تعداد تکرارها از ۵۰۰۰-۰ تغییر می‌کند، خطای میانگین از ۰/۱ تا ۰/۰۰۰۰۱ کاهش می‌یابد. شکل (۳-۱۶) رابطه بین خطای میانگین و تکرارها را برای جدول (۳-۷) نشان می‌دهد.

جدول (۳-۶) نتایج بررسی طول صفحه را نشان می‌دهد وقتی که مقدار درست از ۵-۳۰ واحد تغییر می‌کند، دقت میانگین ۱/۲٪ است. جدول (۳-۷) و جدول (۳-۸) نتایج بررسی مقاومت ویژه را نشان می‌دهد. دقت میانگین ۵٪ برای صفحه هادی و ۲/۸٪ برای صفحه مقاوم است.

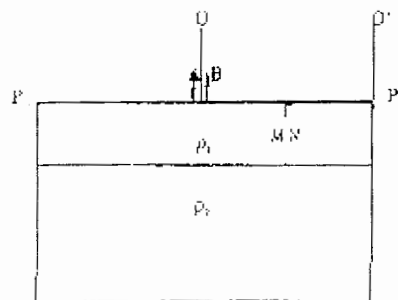
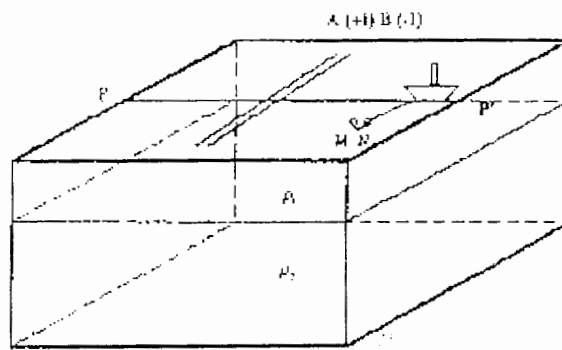
۳-۲-۶- نتایج

با کمک نتایج به دست آمده به طور سیستماتیک توسط مدلسازی عناصر محدود، تغییر نتایج صحرائی بدست آمده از آرایش الکترونی، مانند ترکیب دو قطبی - دو قطبی، امکان‌پذیر است. شبکه عصبی می‌تواند برای وارونه‌سازی پارامترهای صفحه ضخیم عمودی استفاده شود. با ترکیب روش عناصر محدود با شبکه عصبی، ما ممکن است از عهده مساله برای کشف عمق ساختارهای با شیب تند فرو رفته برآییم. [۳۴]



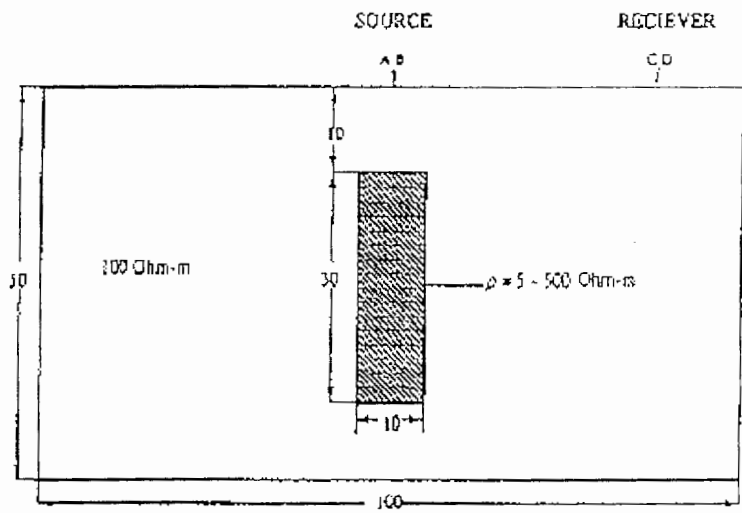
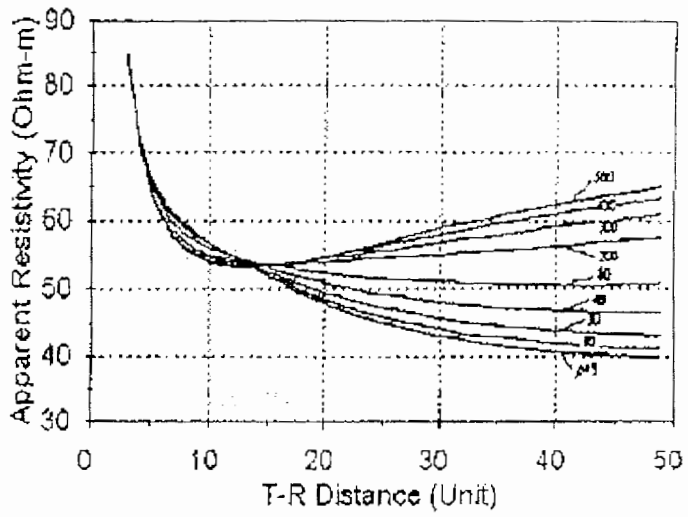
DIPOLE-DIPOLE ARRAY ON LAND SURFACE
 P-P - INVESTIGATION PROFILE
 O-O' T-R DISTANCE
 A- parallel source electrodes
 B- Parallel sink electrodes
 I- electrical current, MN- measurement electrodes
 AB generate a 2-D dipole electrical field

شکل ۹-۳: سونداژ دوقطبی - دوقطبی در روی زمین



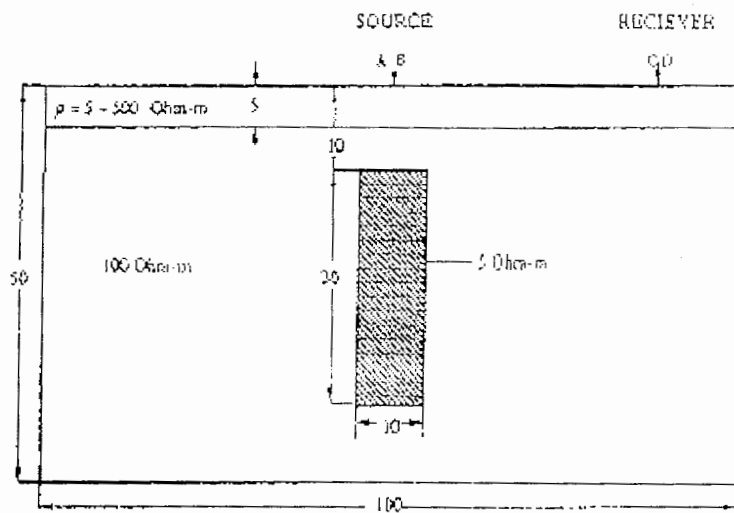
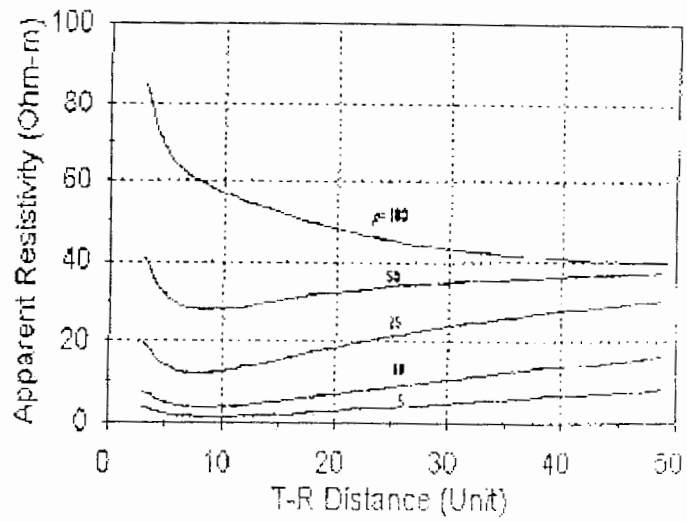
DIPOLE-DIPOLE ARRAY ON WATER SURFACE
 P-Q INVESTIGATION PROFILE
 Q-D' T-R DISTANCE
 A- line source electrode
 B- line sink electrode
 I- electrical current, MN- measurement electrodes
 All generate a 2-D dipole electrical field

شکل ۱-۳: سونداژ دوقطبی - دوقطبی بر روی سطح آب



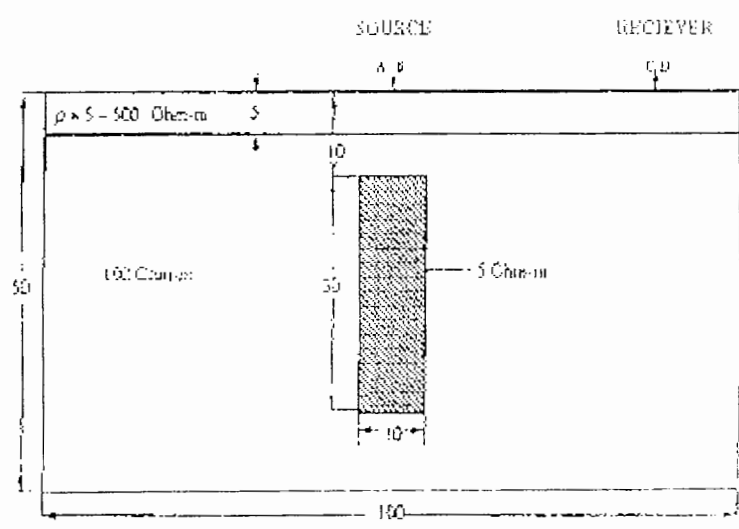
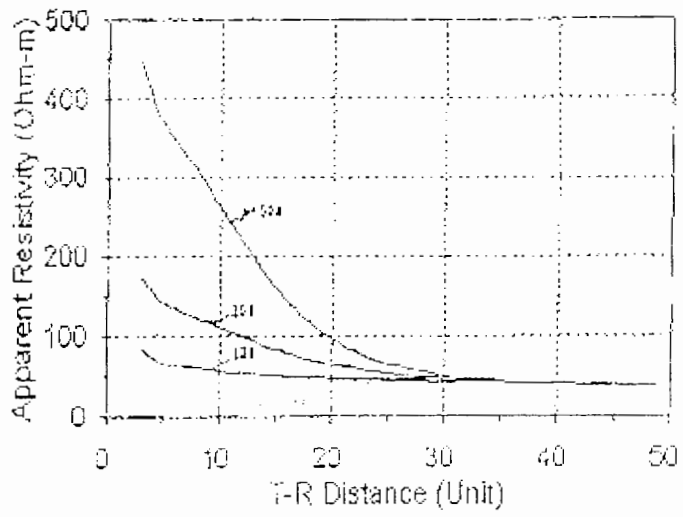
DIPOLE-DIPOLE ARRAY
 The Length's Unit = Unit
 The Distance of AB and CD= 2 Units

شکل ۱۱-۳: مقاومت ویژه بر روی یک صفحه دفن شده با مقاومت ویژه‌های گوناگون.



DIPOLE-DIPOLE ARRAY
 The Length's Unit = Unit
 The Distance of AB and CD = 2 Units

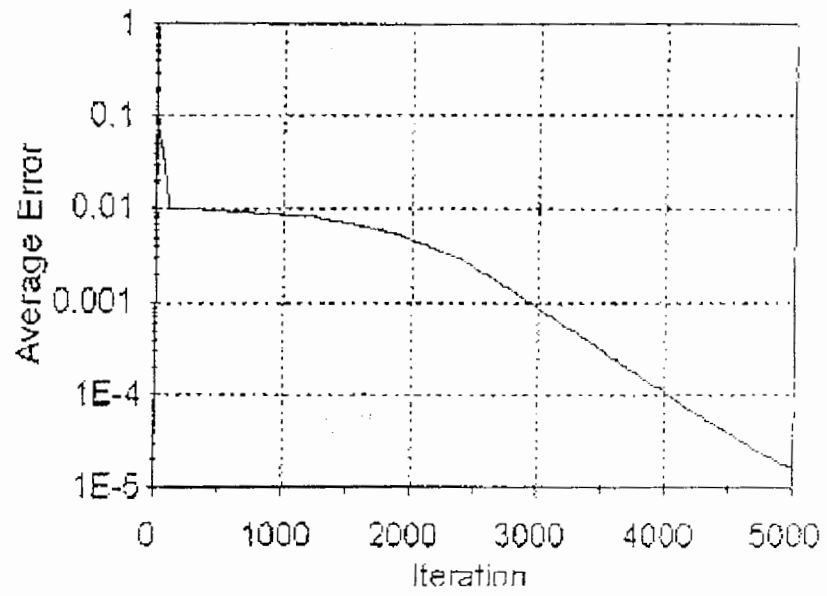
شکل ۳-۲۳: اثرات روباره هادی



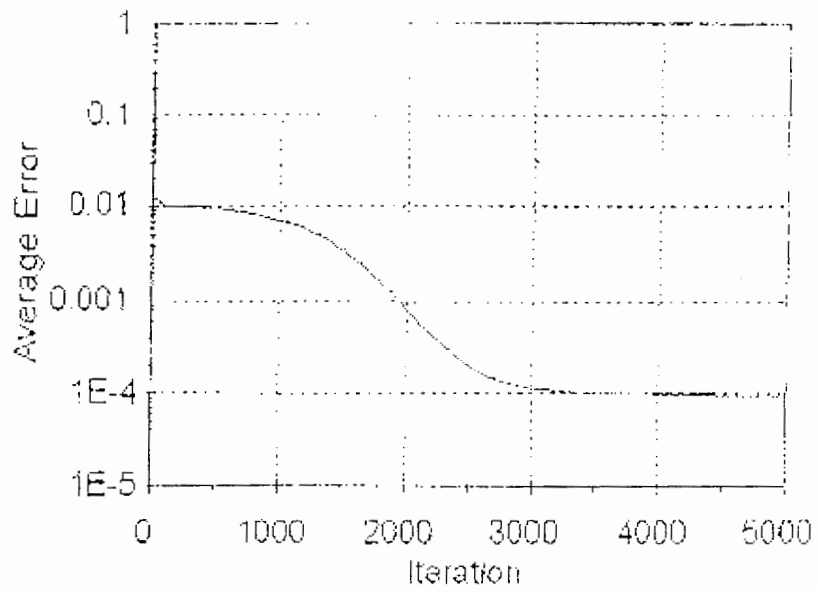
DIPOLE-DIPOLE ARRAY

The Length's Unit = Unit
 The Distance of AB and CD = 5 Units

شكل ٢-١٤: اثرات روباره مقاوم



شکل ۳-۱۵: رابطه بین خطای میانگین و تکرار برای جدول (۴-۳).



شکل ۳-۱۶: رابطه بین خطای میانگین و تکرار برای جدول (۷-۳).

جدول ۶-۳: آشکارسازی طول صفحه

Iteration=25,000, Length unit=Unit		
NNT output	True value	Accuracy
5.2457	5.0	4.9%
9.7636	10.0	2.4%
14.7922	15.0	1.4%
20.0375	20.0	1.8%
25.2527	25.0	1.0%
30.2085	30.0	1.0%
		Average accuracy= 2.1%

جدول ۷-۳: آشکارسازی مقاومت ویژه برای صفحه هادی

Iteration=15,000, Resistivity unit=ohm-m		
NNT output	True value	Accuracy
9.530	10.0	4.7%
13.651	13.0	5.0%
18.562	20.0	7.1%
42.529	40.0	6.3%
78.360	80.0	2.1%
		Average accuracy= 5.0%

جدول ۸-۳: آشکارسازی مقاومت ویژه برای صفحه مقاوم

Iteration=5,000, Resistivity unit=ohm-m		
NNT output	True value	Accuracy
188.97	200.0	5.5%
306.16	300.0	2.0%
359.03	350.0	2.5%
406.79	400.0	1.6%
486.49	500.0	2.7%
		Average accuracy= 2.8%

۳-۳- تفسیر ژئوفیزیکی توسط سیستم‌های عصبی مصنوعی^۳

۳-۳-۱- خلاصه

ما قابلیت اجرا شبکه‌های عصبی در وارونه‌سازی ژئوفیزیکی را بررسی می‌کنیم. شبکه‌های عصبی در مسائل تشخیص الگو مفید بودند و آن‌ها کاملاً تجربه را یاد می‌گیرند. ما یک سیستم عصبی پیشخور مصنوعی را برای وارونه‌سازی برای پارامترهای دل زمین از دو مجموعه داده‌های ژئوفیزیکی مختلف پیاده کردیم: داده‌های سونداژ مقاومت ویژه و تنها امواج p اولیه لرزه نگارهای انعکاسی برای مدل‌های زمین ID. در هر مورد ما شبکه را با استفاده از یک مجموعه داده‌های مصنوعی برای مدل‌هایی که شناخته شده بودند، آموزش دادیم. شبکه‌های آموزش یافته سپس برای این موارد مورد استفاده قرار گرفتند: (۱) پیش بینی یک مدل برای داده‌های سونداژ مقاومت ویژه صحرائی و (۲) پیش بینی مدلها از گردآوری‌های لرزه‌ای که متعلق به مجموعه آموزشی نبودند. مدل برای داده‌های مقاومت ویژه به خوبی با مدل بدست آمده توسط یک نتیجه وارونه‌سازی مستقل مطابقت دارد. در مورد داده‌های لرزه‌ای، درجه پیش بینی از خوب تا متوسط است. ما این نتایج را بر حسب فیزیک‌های مساله مستقیم و پارامترهای (مانند تعداد مجموعه داده‌های استفاده شده برای آموزش، تعداد ندها در لایه میانی و غیره) استفاده شده در شبکه مطرح می‌کنیم. ما از یک قاعده پس انتشار برای آموزش استفاده کردیم که بر پایه روش کاهش شیب بنا نهاده شده است. این روش با مجموعه داده‌های آموزشی محدود شده ما خوب کار می‌کند. برای مجموعه داده‌های بزرگتر و برای داده‌های لرزه‌ای که شامل هر دو اطلاعات فاز و دامنه است ما ممکن است به استفاده از روش‌ای بهینه‌سازی مطلق برای آموزش نیاز داشته باشیم.

۳-۳-۲- مقدمه

وارونه‌سازی و یا تفسیر ژئوفیزیکی، نگاشت داده‌های ژئوفیزیکی را به مدل‌های زمین‌شناسی واقعی را دربر می‌گیرد. مدل‌های بهینه‌سازی اخیر (هم موضعی و مطلق) در وارونه‌سازی ژئوفیزیکی، در جهت تلاش برای تطبیق داده‌های مصنوعی برای یک مدل زمینی فرض شده، با اندازه‌گیری‌های مورد

۳- مطالب این بخش از منبع [۳۲] آورده شده است.

استفاده بودند. روشهای متفاوتی برای استخراج پارامترهای مدل از داده‌های اندازه‌گیری شده وجود دارد. روش خطی خیلی عمومیت دارند هنگامی که اطلاعات زمین‌شناسی معتبر وجود دارد (Tarantola, 1987) و نتایج خوب با روش‌های بهینه‌سازی مطلق بدست آمدند (Sen and Stoffa, 1991)، در جایی که داشتن یک مدل شروعی خوب مهم نیست. سیستم‌های عصبی مصنوعی یک روش متفاوت برای مسائل وارونه‌سازی ژئوفیزیکی ارائه می‌دهد. مساله وارونه‌سازی می‌تواند به عنوان یک نگاهت از یک فضای داده ورودی به پارامترهای مدل به نظر برسد. آموزش شبکه توسط خودش، یک فرایند بهینه‌سازی است و ممکن است استفاده محاسبه‌ای روش بهینه‌سازی مطلق متمرکز را تضمین کند. در این مقاله، ما قابلیت اجرا، قدرت و حساسیت شبکه‌های عصبی را برای وارونه‌سازی ژئوفیزیکی بررسی می‌کنیم.

۳-۳-۳- سیستم‌های عصبی مصنوعی

اغلب شبکه‌های عصبی در مورد استفاده بودند، تنها اخیراً تلاشهایی در جهت کاربرد آنها برای وارونه‌سازی شکل لرزه‌ای صورت گرفته است. تعداد نرونها در ورودی و لایه خروجی به ترتیب توسط تعداد نقطه داده‌های ورودی و تعداد نقاط در بردارهای مدل خروجی مطلوب داده می‌شود. تعداد لایه‌های میانی و تعداد نرونها در هر یک از آنها تا حدی اختیاری است. ANS (Neural Systems Artificial) با یک روند جهتی، شبکه‌های پیشخور نامیده می‌شوند. در این کار، ما یک شبکه پیشخور سه لایه برای وارونه‌سازی پارامترهای مدل ID از مجموعه داده‌های ژئوفیزیکی به کار بردیم.

برای آموزش شبکه، نمونه‌های ورودی به لایه اول (لایه میانی) از واحدهای شبکه (شکل ۳-۱۷) ارائه می‌شوند که در چرخش ورودی‌های جدید برای لایه‌های متوالی تولد می‌کند تا یک خروجی تولید شود. خرجی شبکه و خروجی مطلوب یک علامت خطا می‌دهد که پس انتشار می‌یابد و در طول همه واحدهای شبکه بر پایه این خطا، داده‌های جدید برای وزنه‌های اتصالی به طوری محاسبه می‌شوند که شبکه به حالت پایه یا مینیمم موضعی همگرا شود.

در این مرحله، بوسیله آموزش مناسب با یک مجموعه داده‌های نماینده، شبکه یک توانایی برای نگاهت نمونه‌های مشابه به فضای خروجی را بسط می‌دهد، که این فرایند به عنوان عمومیت

دادن شناخته شده است. خصیصه‌های اصلی از ورودی، در حال حاضر به طریقی در یک شکل از ارتباطات بین واحدها ذخیره شده‌اند.

فرض کنید یک شبکه با یک لایه از واحدهای میانی مستقیماً متصل شده به واحدهای خروجی (شکل ۳-۱۷) که حالت‌هایش یک تابع هموار پیوسته از ورودی‌های کلی آن‌ها است. بردار ورودی X_i به تمام واحدها در لایه میانی توسط معادلات زیر انتشار می‌یابد:

$$A_j^h = \sum_{i=1}^N W_{ji}^h X_i + \Theta_j^h \quad \text{فرمول (۳-۱۳)}$$

$$i_j = f_j^h(A_j^h) \quad \text{فرمول (۳-۱۴)}$$

که بالانویس h به واحد میانی اشاره دارد، A ورودی شبکه به ز واحد میانی است، W_{ji} وزن یک اتصال از i امین واحد ورودی است، θ_j هم بایاس است، و f_j^h از اند خروجی است، f_j^h توسط یک تابع از نوع سیگموئید به شکل $f_j^h = (1 + e^{-x})^{-1}$ داده شده است. معادلات برای ندهای خروجی می‌تواند به شکل زیر نوشته شود:

$$A_k^o = \sum_{j=1}^l W_{kj}^o i_j + \Theta_k^o \quad \text{فرمول (۳-۱۵)}$$

$$y_k = f_k^o(A_k^o) \quad \text{فرمول (۳-۱۶)}$$

که o بر لایه خروجی دلالت دارد، و l تعداد واحدهای در لایه میانی است. در این مورد، $f_k^o = A$ فرض کنید که ما اکنون می‌خواهیم شبکه را با برخی خروجی‌های d_k مطلوب آموزش دهیم، یک اندازه‌گیری از خطا می‌تواند به شکل زیر تعریف شود:

$$E = \frac{1}{2} \sum_k |d_k - y_k| \quad \text{فرمول (۳-۱۷)}$$

خطا می‌تواند کمینه شود توسط شروع با یک مجموعه از وزنها تصادفی و مکرراً تغییر هر وزن توسط یک مقدار متناسب با:

$$W_{kj}^o = \varepsilon \frac{\partial E}{\partial W_{kj}^o} \quad \text{فرمول (۳-۱۸)}$$

که ε میزان یادگیری است، به طوری که ε متمایل به صفر باشد، و تعداد داده‌های جدید متمایل به بی‌نهایت باشد، این روش یادگیری برای پیدا کردن مجموعه‌های وزنها تضمین شده است که خطای حداقل مربعات را می‌دهد. معادلات دیفرانسیل (۳-۱۵) و (۳-۱۶)، $\partial E / \partial W$ می‌تواند به صورت زیر نوشته شوند:

$$\frac{\partial E}{\partial W_{kj}^o} = -(d_k - y_k) f_k^{o'}(A_k^o) i_j = \delta_k^o i_j \quad \text{فرمول (۱۹-۳)}$$

روش پس انتشار یک تعمیم از روش حداقل مربعات برای شبکه با لایه‌های میانی بین لایه‌های ورودی و خروجی است. ایده مرکزی پس انتشار این است که مشتقات، با شروع در لایه خروجی و به عقب برگشتن برای محاسبه مقادیر جدید برای هر لایه، محاسبه می‌شوند. معادلات مشابه شرح داده شده در معادله (۱۹-۳) می‌تواند برای لایه میانی به کار برده شود به طوری که:

$$\frac{\partial E}{\partial W_{kj}^h} = -\sum_k (d_k - y_k) f_k^{o'}(A_k^o) W_{kj}^o f_j^{h'}(A_k^h) X_i = \delta_i^h X_i \quad \text{معادله (۲۰-۳)}$$

از معادلات (۱۹-۳) و (۲۰-۳)، وزنهای جدید می‌تواند به صورت زیر خلاصه شوند:

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_i^h X_i \quad \text{معادله (۲۱-۳)}$$

که X ورودی را به لایه در مطالعه ارائه می‌دهد، و η میزان یادگیری است. در ای معادلات بالانویس‌ها از قلم افتاده‌اند، زیرا روابط مشابه برای هر دو لایه میانی و خروجی نگه داشته می‌شوند. معادله (۲۱-۳) عموماً به عنوان *قاعده دلتای تعمیم یافته* منسوب شده است.

یک راه برای افزایش سرعت همگرا شدن برای مقادیر جدید استفاده از یک تکنیک شناخته شده به عنوان *ممنتوم* است، در اینجا یک قسمت از وزن قبلی اضافه شده است:

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_i^h X_i + \alpha \Delta W_{ji}(t-1) \quad \text{معادله (۲۲-۳)}$$

که α ضریب ممنتوم است و مجموعه‌ای عادی با یک مقدار کمتر از ۱ است. این جمله اضافی، تغییرات وزن را مجبور می‌کند که در یک جهت یکسان نگه داشته شوند.

۳-۳-۴- کاربرد برای داده‌های سونداژ مقاومت ویژه

داده‌های سونداژ مقاومت ویژه شامل یک مجموعه از اندازه‌گیری‌های مقاومت ویژه ظاهری به عنوان یک تابع از فاصله الکترونی است. این داده‌ها سپس به مدل‌های زمین، شامل لایه‌های همگن از مقادیر مختلف مقاومت ویژه و ضخامت (h) نگاشته می‌شوند. وارونه‌سازی داده‌های مقاومت ویژه (p) عموماً بوسیله انطباق منحنی با استفاده از منحنی‌های سر یا توسط روش خطی قراردادی انجام می‌شود. اخیراً، یک روش بهینه‌سازی مطلق غیرخطی برای وارونه‌سازی داده‌های مقاومت ویژه با نتایج خوب به کار برد. در این کار، ما ابتدا یک مجموعه از ۵۰ مدل هر کدام با سه لایه به

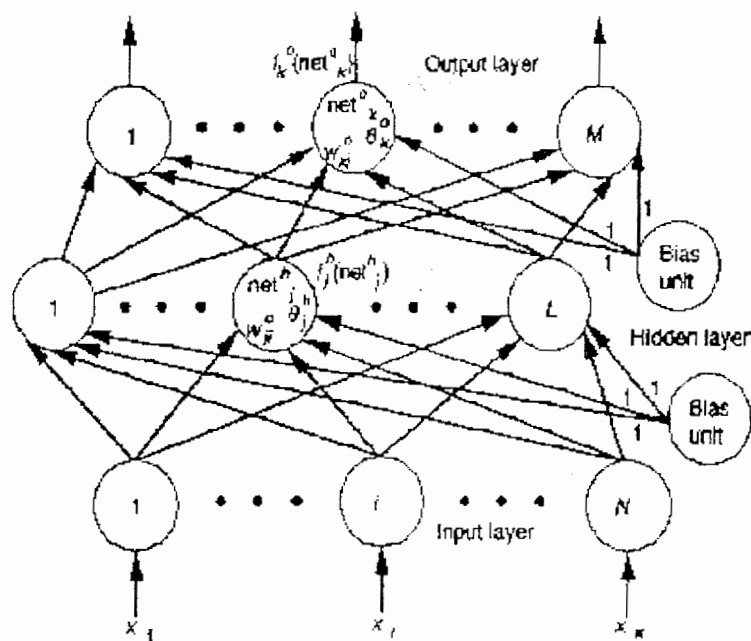
تولید کردیم. مقادیر مقاومت ویژه ظاهری نظری، برای یک آرایش سلومبرژه در ۱۹ فاصله الکترونی محاسبه شده بودند. از این مجموعه، ۳۰ مدل به دلخواه انتخاب شده از مجموعه ورودی برای آموزش شبکه استفاده شده بودند. بیست و سه واحد در لایه میانی استفاده شده بودند، و ۱۰۰۰۰ تکرار برای آموزش شبکه مورد استفاده قرار گرفته بود. پارامترهای η و α که به ترتیب نرخ یادگیری و ممنتوم نامیده می شوند، برابر با $\eta = 0.01$ و $\alpha = 0.95$ بودند. نتایج از فرایند آموزشی و پیش‌بینی در یک شکل از «نمودار حداقل مربعات در شکل (۳-۱۸) نشان داده شده‌اند. توجه کنید که شبکه می‌تواند خیلی خوب آموزش داده شود به طوری که توسط یک خطای خیلی کوچک برای مجموعه آموزشی ثابت شده. خطا همچنین برای مجموعه ای که متعلق به مجموعه آموزشی نبود، کوچک بود.

سپس ما یک مجموعه از داده‌های مقاومت ویژه صحرایی را وارد کردیم (که زودتر توسط سن و دیگران تفسیر شده بود و نیز مشخصه‌هایی مشابه با مجموعه داده با مجموعه داده‌های آموزشی دارد) و از شبکه آموزش یافته برای پیش‌بینی مدل استفاده کردیم. مدل پیش‌بینی شده به خوبی با مدل به دست آمده توسط یک وارونه‌سازی ذوب شبیه‌سازی شده مستقل مقایسه شد. (شکل ۳-۱۸) b. داده‌های مصنوعی برای مدل پیش‌بینی شده شبکه همچنین توافق قابل قبول با مشاهدات نشان داده شده‌اند (شکل ۳-۱۸) b. اگر چه یک مجموعه کوچک از الگوهای شناخته شده در فرایند آموزشی استفاده شده بودند، شبکه قادر به «تعمیم» و تولید مدل‌های قابل قبول بود.

۳-۳-۵- نتایج

ما امکان استفاده از سیستم‌های عصبی مصنوعی را در وارونه‌سازی داده‌های ژئوفیزیکی با کاربرد در داده‌های سونداژ مقاومت ویژه d.c. و گردآوری لرزه‌ای بررسی کردیم. نتایج حتی با استفاده از یک تعداد کوچک از مجموعه داده‌های آموزشی، دلگرم کننده است. قوت روش برای حضور نویز و آموزش برای نمونه‌های عمومی بیشتر به طور متداول بررسی شد. ما تنها نتایج را با استفاده از یک شبکه پیشخور ۳ لایه ارائه داریم که با قاعده پس انتشار آموزش یافته بود. کار بیشتر برای بررسی کارایی شبکه‌های دیگر و روش‌های مختلف آموزشی لازم است. کارایی شبکه، با

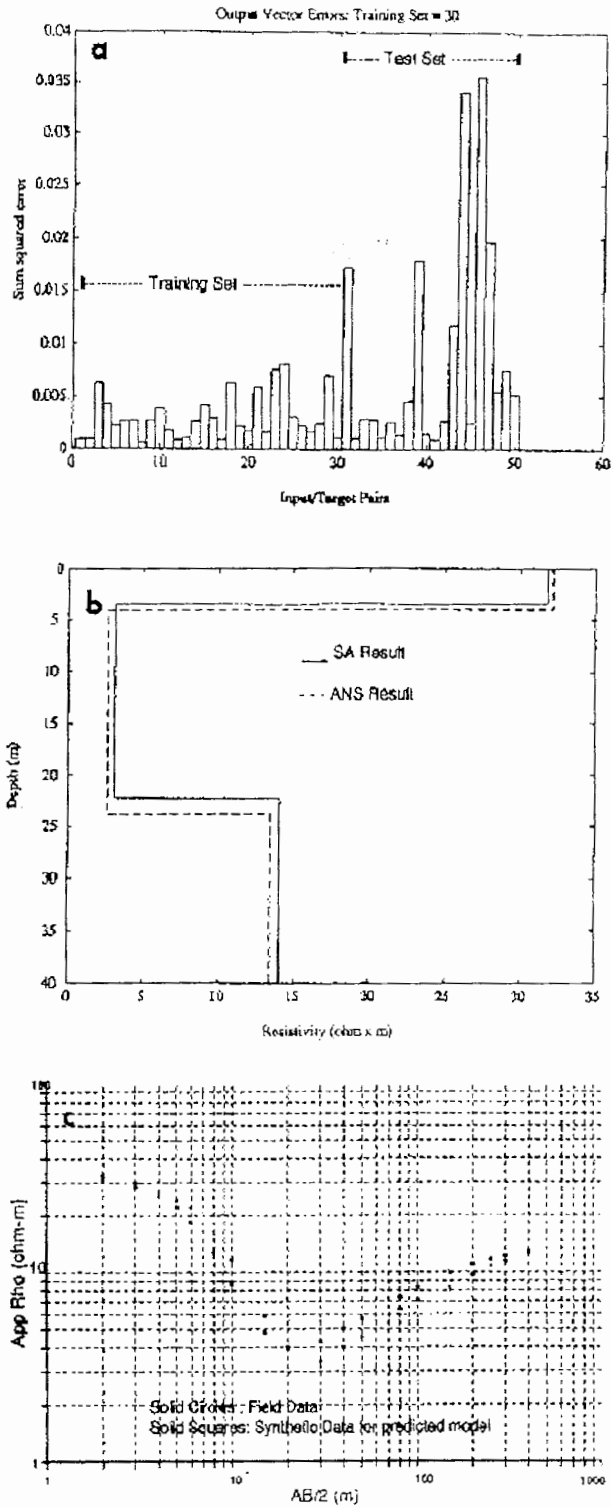
پارامترهایی مانند نرخ یادگیری، ممنتوم، اندازه مجموعه آموزشی و تعداد واحدهای میانی استفاده شده برای آموزش شبکه، تغییر می‌کند. تعداد نمونه‌های ورودی به نظر می‌رسد اهمیت عمده‌ای در وارونه‌سازی با استفاده از ANS داشته باشد. [۳۲]



شکل ۳-۱۷: یک شبکه پیشخور سه لایه‌ای نوعی

شکل ۳-۱۸: آموزش و پیش بینی ANS داده‌های سونداژ مقاومت ویژه :

(a). نمودار خطای مربع برای آموزش و مجموعه داده‌های آزمایشی. (b). نتیجه از وارونه‌سازی داده‌های صحرایی مقاومت ویژه. مدل توسط ANS (خط نقطه‌چین) و نتیجه SA (خط پر) پیش بینی شده است. (c) مقایسه بین داده‌های صحرایی و تئوری از ANS



۳-۴- مثالی دیگر

کالدرون - ماسیاس و دیگران (Calderon-Macias et al. 2000) یک شبکه عصبی ذوب شبیه سازی شده خیلی سریع (VFSA) استفاده شده برای وارونه سازی داده‌های مقاومت ویژه الکتریکی را شرح داد. داده‌های آموزشی از یک مدل مستقیم تولید شده بودند و داده‌های آزمایشی از مقالات انتشار یافته گرفته شده بود. یک روش سونداژ شلومبرژه برای آرایش الکترودها استفاده شده بود. ۲۵۰ منحنی برای مدل‌های زمین سه لایه‌ای تولید شده بود که در آن $p_1 > p_2 < p_3$ شده بود. مقاومت ویژه لایه بالایی برای همه مدلها در $1 \Omega m$ ثابت شده بود. ضخامت لایه اول بین ۱ و ۱۰ متر تغییر می‌کرد. مقاومت ویژه لایه دوم بین $0.3 \Omega m$ و $0.2 \Omega m$ تغییر می‌کرد. و مقاومت ویژه لایه سوم بین $0.15 \Omega m$ و $0.61 \Omega m$ بود. ضخامت لایه سوم در محدوده بین ۳ متر تا ۲۰ متر بود. ۲۰ فاصله الکترودی مختلف مدل‌سازی شده بود. یک لایه میانی با ۸ PES (نرون) برای بهینه شدن در نظر گرفته شد. مدل ارزیابی شده شبکه به عنوان یک مدل شروع برای یک وارونه سازی حداقل مربعات با روش نیوتن استفاده شد. در حالی که تخمین شبکه عصبی بر پایه داده‌های آزمایشی صحرائی، به مدل صحیح نزدیک بود، وارونه سازی حداقل مربعات دقت تخمین ضخامت لایه دوم را بهبود بخشید. وقتی یک مدل آغازین تصادفی، درون مرزهای پارامترهای آموزشی برای وارونه سازی حداقل مربعات استفاده شده بود، روش به مدل صحیح همگرا نشد. [۲۷]

فصل چهارم

کاربرد شبکه‌های عصبی در الکترومغناطیس

۴-۱- تشخیص الگوهای بیضی وار EM با شبکه‌های عصبی^۱

۴-۱-۱- خلاصه

اینجا ما یک مطالعه از قابلیت اجرایی شبکه‌های عصبی را برای حل برخی مسائل وارونه‌سازی ژئوفیزیکی شرح می‌دهیم. در این مطالعه، داده‌های الکترومغناطیس حوزه فرکانس نوع اسلینگرام جمع‌آوری شده بر روی استوانه‌های افقی هادی (لوله‌های زیر زمینی فولادی) برای آموزش یک شبکه عصبی برای تمیز دادن مخزنهای ذخیره‌ای زیرزمینی استفاده شده است. این داده‌ها شامل مؤلفه‌های هم فاز و تربیعی دوقطبی عمودی و افقی پاسخ الکترومغناطیس است. تفسیر اندازه‌گیری‌های بیضی وار الکترومغناطیسی سطح زمین به عنوان یک برنامه الگوشناسی در مقابل یک برنامه معکوس تلقی شده است. یک پردازش توزیع شده موازی یا روش شبکه عصبی برای الگوشناسی استفاده شده است. یک شبکه پیشخور، پس انتشار برای شناختن مکان فضایی یک آنومالی نمونه‌های بیضی وار داده شده از ۲۰ مدل نظری آموزش داده شده است.

۱ مطالب این بخش از منبع آدرده شده است.

۴-۱-۲- اندازه‌گیری‌های صحرائی

اخیراً دربارهٔ فواید پردازش توزیع شدهٔ موازی (PDP)، یا شبکه‌های عصبی، برای مسائل الگوشناسی زیاد نوشته شده است. یکی از اهداف این مقاله ارائهٔ برخی نتایج یک پروژه طراحی شده برای تعیین اینکه آیا روشهای PDP توانایی تهیهٔ سریع و تشخیص قابل اعتماد نمونه‌های EM سطحی در صحرا را دارند یا نه.

یکی از برنامه‌هایی که در اینجا مطالعه می‌گردد یک شبکهٔ "پیشخور، کاملاً مرتبط" است که در شکل ۴-۱ نشان داده شده است.

یک سیستم تصویرنگاری EM با قدرت تفکیک‌پذیری بالای منحصر به فرد (Stenberg et al. 1988 and 1989) برای بدست آوردن داده‌های صحرائی برای برنامهٔ آزمایشی عصبی استفاده شده است. این سیستم از یک سیم بلند که در یک خط روی زمین کشیده شده یا منبع حلقهٔ بزرگ استفاده می‌کند. فرکانس فرستنده از ۳۰ Hz تا ۳۰ kHz است. یک سیم پیچ دریافت‌کنندهٔ ۳ جزئی برای ضبط میدانهای مغناطیسی در طول خطوط کاوش، عمود بر منبع سیم بلند استفاده شده است. بیضی‌وار میدان مغناطیسی از این میدانهای ضبط شده تعیین می‌گردد. بیضی‌وار یک نشانگر حساس بخصوص از هادی‌های زیرزمینی دارد.

یکی از جنبه‌های منحصر به فرد این سیستم یک روش درجه‌بندی (کالیبره کردن) با دقت خیلی بالا است. برنامهٔ آزمایشی شبکهٔ عصبی برای بکار انداختن این اندازه‌گیری‌های با دقت بالا برای تولید یک تصویر زمینی با قدرت تفکیک‌پذیری بالا طراحی شده است.

۴-۱-۳- برنامهٔ آزمایشی شبکهٔ عصبی

تصاویر بیضی‌وار EM ترکیبی با استفاده از مدلسازی مستقیم عناصر محدود از یک هدف منبع خطی در یک دنباله از ۱۱ فرکانس بوجود آمده بودند. یک تصویر بیضی‌وار نوعی در شکل ۴-۲ نشان داده شده است. محور X فاصلهٔ افقی از منبع خطی تقسیم شده توسط یک میانگین جبران‌کننده (فاصلهٔ بین منبع و آنومالی) برای منطقهٔ مدنظر است. محور Y لگاریتم عمق پوسته‌ای است. هدف استفاده شده برای برنامهٔ مستقیم و آزمایش صحرائی شامل یک لولهٔ گالوانیزه شدهٔ ۵۰ cm D دفن شده در عمق ۲ m و جبران‌کنندهٔ ۱۰ متری از منبع است. چنین هدفی

می‌تواند به راحتی برای ارائه یک موقعیت صحرایی نوعی زیاد مقیاس شود. ۲۰ هدف نظری در موقعیتها و عمقهای گوناگون اطراف لوله و دارای ویژگیهای مشابه به عنوان لوله مدلسازی شده بودند (شکل ۴-۳ را ببینید). تصاویر بیضی‌وار به‌عنوان ورودی برای آموزش شبکه عصبی بوجود آمده بودند.

یک شبکه با ۹۰ عنصر ورودی، ۲۰ عنصر میانی و ۲ عنصر خروجی برای تشخیص بین ۲۰ نمونه بیضی‌وار و تعیین موقعیت جبران کننده و عمق مدل که نمونه را تولید می‌کند آموزش داده شده بود. توانایی شبکه برای عمومی‌سازی و تشخیص نمونه‌های جدید توسط ارائه ۱۱ نمونه جدید آزمایش شده بود. شامل شده در این مجموعه آزمایشی پاسخ تئوری تولید شده توسط لوله دفن شده واقعی بود. قدم بعدی ما استفاده از داده‌های صحرایی جمع‌آوری شده بر روی لوله در مکانی از داده‌های نظری خواهد بود.

شبکه‌های عصبی می‌توانند در یک تصویر کامل آموزش ببینند اما این به نتیجه در آموزش آرام گرایش پیدا می‌کند. بنابراین فشرده‌سازی تصاویر برای نگهداشتن تنها جنبه‌های نمونه‌های اصلی برای شناخت عصبی لازم است. چشم از جنبه‌های ویژه برای تشخیص چندین نمونه‌های زیاد استفاده می‌کند. برای یک امتحان فشرده‌سازی اولیه، یک آنالیز مولفه‌های اصلی برای تعیین انجام شده بود که ساختارهای نمونه اغلب برای همه مدلها تشخیصی بودند. مقدار پیک در بیضی‌وار در مقابل منحنی فاصله شامل ۳۱٪ از واریانس می‌شود. یک ترکیب از فاصله و دامنه تغار اول، پیک اول و تغار دوم برای هر عمق پوسته‌ای شامل ۹۶٪ از واریانس می‌شود بطوریکه این مقادیر به عنوان ورودی به شبکه آموزش استفاده شده بودند.

بعد از ۲۰۰۰۰۰ تکرار مدل آموزشی پس‌انتشار، شبکه عصبی قادر به تعیین مکان آنومالی با یک دقت میانگین ۹۹٪ در جبران کننده و عمق بود.

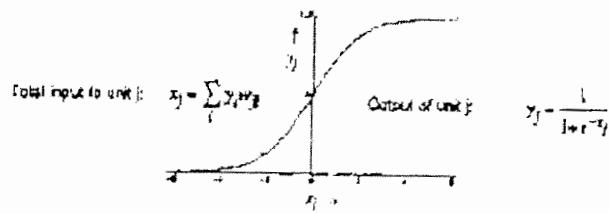
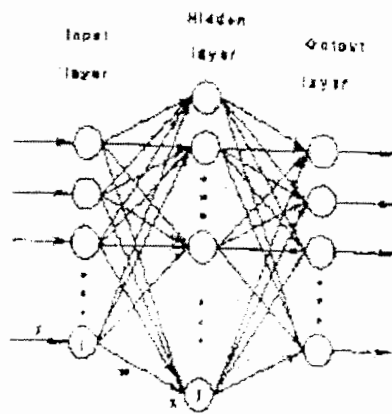
موقعیت‌های مدل‌های نظری استفاده شده برای آموزش در شکل ۵-۲ به شکل دایره‌های بی‌نشان نمایش داده شده‌اند. مدل‌های استفاده شده برای آزمایش توانایی شبکه برای عمومی‌سازی با A نشان داده شده‌اند و موقعیت‌های محاسبه شده توسط شبکه با C نشان داده شده‌اند. مدل‌های A6، A7، A10 و A11 برای آزمایش توانایی شبکه برای برون‌یابی آنسوی حدود مجموعه آموزش استفاده شده بودند. بطوریکه شکل ۴-۳ نشان می‌دهد، شبکه بطور صحیح عمقها را برای مدل‌های کم عمق A6 و A7 محاسبه کرده است اما برون‌یابی کافی دور برای جبران کننده موفق نشده است.

برای مدل‌های عمیق A10 و A11، شبکه برای قادر شدن به برون‌یابی عمق‌های عمیق کافی نمایان نیست؛ آن توسط دور کردن زیاد آنها از مبداء جریان می‌کند. مقایسه بردارهای ورودی برای مدل‌های عمومی‌ساز و مدل‌های آموزشی برای چیزی که شبکه بایستی به طور نزدیکی به آنها وابسته شده نشان می‌دهد که شبکه اغلب موقعیت‌های منطقی براساس اطلاعات ارائه شده به آن را محاسبه کرده است.

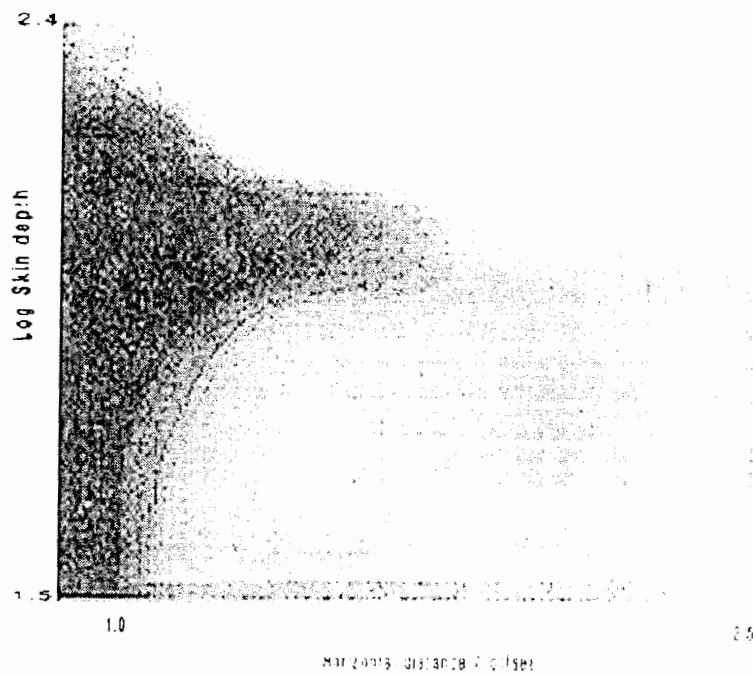
تجربه ما با آموزش این شبکه‌ها با گذشت زمان می‌تواند به صورت زیر خلاصه شود. آموزش شبکه با استفاده از تابع سیگموئید، معادله (*)، به طور ناچیزی انجام می‌دهد مگر اینکه بردارهای ورودی برای کاهش محدوده دینامیک برای یک دهه یا بیشتر میزان شوند. به عنوان تولید نقطه (X در معادله *) بزرگ می‌شود (مثبت و منفی)، خروجی تابع سیگموئید در صفر یا یک نتیجه در وقفه شبکه اشباع می‌شود. میزان‌دهی نمونه ورودی و کاهش مقادیر وزن اولیه برای کم کردن این برنامه‌ها به نظر می‌رسد.

$$f(x) = (1 + e^{-x})^{-1} \quad \text{معادله (*)}$$

شبکه‌ها اطلاعات را یاد می‌گیرند که خیلی خوب داده شده‌اند. اگر داده‌های ورودی ارائه نشوند، نتایج هم همینطور خواهند بود. بنابراین، دقت زیادی بایستی در انتخاب مدل‌های استفاده شده برای آموزش بکار برده شود. روشهای بهبود یافته فشرده‌سازی تصویر، برای بررسی پیدا کردن راههای کاهش اندازه بردارهای ورودی بدون کم کردن جزئیات نمونه اصلی لازم است. [۳۵]

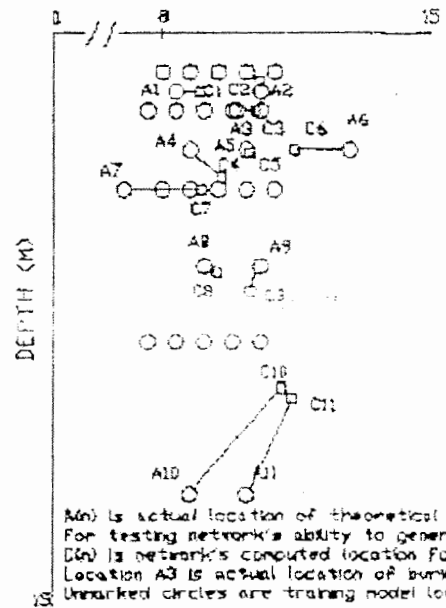


شکل ۴-۱. شبکه عصبی پس انتشار پیشخور



شکل ۴-۲. تصویر بیضی وار LAM نوعی ساه=بیضیت کم، سفید=بیضیت بالا.

HORIZONTAL OFFSET FROM TRANSMITTER (M)



شکل ۴-۳. موقعیت مدل‌های آزمایشی و آموزشی، و مکانهای اجسام نشان داده شده برای مدل‌های آزمایشی.

۴-۱-۴- ارزیابی داده‌های الکترومغناطیس با استفاده از شبکه عصبی^۲

۴-۱-۴-۱- روش بررسی

یک توده غیر هادی برای جمع‌آوری داده‌های الکترومغناطیس (EM) حوزه فرکانس نوع اسلینگرام در یک هندسه فضای آزاد استفاده شد که اجازه می‌داد بخش‌های تربیتی و زمین‌شناسی برای پاسخ ابزار برداشته شود. اهداف، شامل لوله‌های زیرزمینی فلزی، در زیر یک ابزار EM قرار گرفتند. قرائت‌ها در سه ارتفاع بالای اهداف انجام شد، با محورهای موازی و نرمال استوانه برای خط آرایه فرستنده/گیرنده. اندازه‌گیری‌ها در ۳ فوت فاصله روی یک "فاصله کاوش" از ۹۶ فوت انجام شد به طوری که اهداف زیر ابزار کشیده شده بود.

داده‌ها شامل مؤلفه‌های هم فاز و هدایت دوقطبی (ثابت X تربیتی) افقی و عمودی پاسخ الکترومغناطیسی است. روابط این داده‌ها مربوط به پارامترهای هدف و هندسه کاوش می‌شود که برای تعریف آنومالی الکترومغناطیسی مربوط به استوانه‌های هادی کافی است. مجموعه‌های آموزش برای یک شبکه عصبی از این داده‌ها بوجود می‌آیند.

یک مجموعه آموزش شامل اندازه‌گیری‌های تربیتی و هم فاز برای هر دو دوقطبی عمودی (VQ و VI) و دوقطبی افقی (HQ و HI) در سه فاصله اندازه‌گیری متوالی از مرکز آرایش ابزار است. هر مجموعه آموزشی نتیجتاً شامل دوازده اندازه‌گیری است. برای فاصله ۹۶ فوت ۴۲ مجموعه برای هر آرایش هندسی هدف/ابزار وجود دارد، وقتی که از یک "فاصله کاوش ۶ فوتی" استفاده می‌شود. برای توضیح، یک مجموعه آموزشی به صورت زیر شرح داده می‌شود:

-48 ft HQ HI VQ VI

-42 ft HQ HI VQ VI

-36 ft HQ HI VQ VI

مقادیر

-42 ft HQ HI VQ VI

-36 ft HQ HI VQ VI

-30 ft HQ HI VQ VI

یک مجموعه آموزشی بعدی (اما نه لزوماً متوالی) را شرح می‌دهد، که انجام می‌دهد:

-45 ft HQ HI VQ VI

-39 ft HQ HI VQ VI

^۲ مطالب این بخش از منبع ۴۳ آورده شده است.

دو جهت آرایش هدف/ ابزار برای جمع‌آوری داده‌های مجموعه آموزش استفاده شدند:

(۱) محور استوانه و موازی ابزار برای "جهت کاوش"، (۲) موازی محور استوانه، آرایش ابزار نرمال برای جهت کاوش. این در ۵۰۴ نقطه آموزش برای هر ارتفاع بالای هدف نتیجه می‌دهد.

در طول تابستان ۱۹۸۹، داده‌های الکترومغناطیسی هم فاز و تربیعی دوقطبی افقی و عمودی در هاوایی جمع‌آوری شد، برای تعیین اگر مخزنهای ذخیره زیرزمینی هنوز در مکان چندین ایستگاه قرار داشتند، و اگر در مکان، برای نگاشت آن موقعیتها. این داده‌ها به عنوان داده‌های آزمایشی صحرائی برای شبکه عصبی استفاده شده بودند. داده‌ها توسط برنامه شبکه عصبی توسعه یافته برای بررسی‌هایمان دوباره قالب‌بندی شده و پردازش شد. برنامه داده‌های صحرائی را در "windows" قابل مقایسه با داده‌های مجموعه آموزشی استفاده شده برای آموزش شبکه عصبی برای تفکیک مخزنهای ذخیره زیرزمینی جدا کرد. هر مقطع سپس توسط شبکه عصبی برای تعیین اینکه آیا این داده‌ها وابسته به یک آنومالی مدنظر هستند آزمایش می‌شود.

این فرآیند برای یک عدد تعیین شده دلخواه از تکرارها، تکرار می‌شود. شبکه اختصاص می‌دهد مقادیر نزدیکتر به ۱ را برای آن قطعاتی که تفکیک شده هستند به عنوان یک مؤلفه پاسخ الکترومغناطیسی موازی یا عمودی برای یک آنومالی مربوط به یک استوانه هادی. برنامه یک تصویر مکعبی با یک مقدار بعدی متناسب با وزن‌های اختصاص یافته توسط شبکه عصبی، اختصاص می‌دهد. به طوریکه با آموزش شبکه عصبی، این یک عملکرد تکراری تصادفی است و وزنهای اختصاص یافته به طور پیوسته با هر عبور از میان شبکه تعدیل می‌شوند.

در حال حاضر، عدد تکرارها، یا طول زمان شبکه عصبی برای "مطالعه" داده‌های بطور دلخواه انتخاب شده داده شده است. تصویرها برای تصویر ویدئویی کامپیوتر نگاشته شده هستند به عنوان بلوکهای سه بعدی میزان شده در رابطه با کاوش شبکه استفاده شده برای بدست آوردن داده‌های کاوش F.M. بلوکها یک نمایش بصری وزندار از پاسخهای آنومالی تهیه می‌کند. بلوکهای کوچکتر احتمال کمتری برای یک مخزن ذخیره‌ای زیرزمینی برای ارائه دادن نشان می‌دهند، در حالیکه بلوکهای بزرگتر یک احتمال بیشتر را نشان می‌دهند که یک مخزن ذخیره‌ای وجود دارد.

داده‌های هم فاز و تربیعی جمع‌آوری شده در Hickam در نقشه‌های خطوط میزان چند خطی نشان داده شده‌اند (شکل ۴-۴ و ۴-۵). با استفاده از این نقشه‌ها، یک متخصص ژئوفیزیک مجرب نبایستی زحمتی در تعیین حضور و موقعیت وجود مخزنهای ذخیره‌ای زیرزمینی داشته باشد. چگونه آیا کاربرد شبکه عصبی مقایسه می‌شود با فرایندها و محاسبات انجام شده توسط متخصصین ژئوفیزیک؟ تصویر تولید شده توسط برنامه شبکه عصبی در شکل ۴-۶ نشان داده شده است. این مقاله با تفسیرهای بصری ساخته شده از نقشه‌های خطوط میزان هم فاز و تربیعی موافق است.

مفاهیم هوش مصنوعی تعبیر و تفسیر آنومالی‌های ژئوفیزیکی را افزایش می‌دهد. شبکه عصبی آموزش یافته می‌تواند برای تفسیر آنومالی‌های داده‌های ژئوفیزیکی مربوط به ویژگیهای پاسخ یک نمونه اولیه استفاده گردد.

استفاده از یک شبکه عصبی یک فرایندی است که برای متخصصین ژئوفیزیک فواید زیادی دارد. آنها شامل هستند از:

شبکه عصبی می‌تواند مجموعه داده‌ها را در هر یک پاسخ صحیح گام به گام یا چندگانه بدون نیاز به یک اختلاط بصری کامل کند.

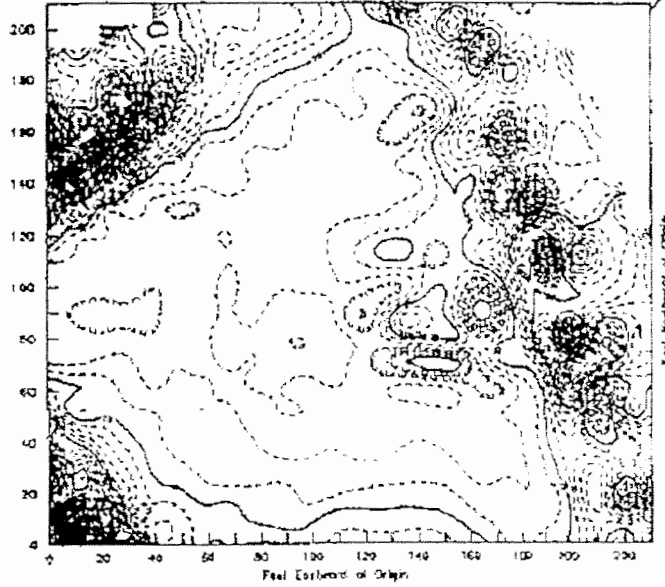
شبکه عصبی یک صرفه جویی کننده در زمان است.

شبکه عصبی یک فرایند بسط پذیر است. بعلاوه جداسازی پارامترهای پاسخ الکترومغناطیسی برای این بررسی استفاده شده، نسبتهای بین این داده‌ها، به خوبی انواع داده‌های دیگر، می‌تواند به مجموعه آموزش کامل شود.

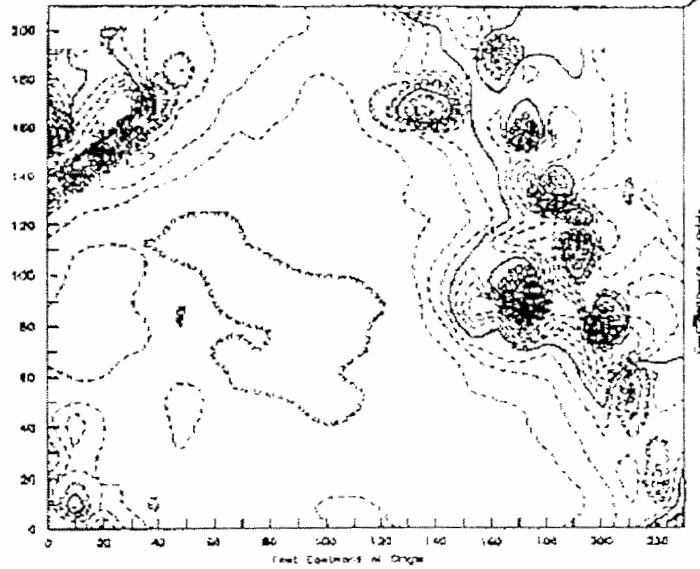
هدف ما آزمایش اثر شبکه عصبی، قوت، سرعت و تطبیق پذیری در حضور نویز بود. ما چیزهای زیادی برای یادگیری داریم، اما تجربه اخیر ما با این سیستمها یک امید و تکنیک تطبیق پذیر را نشان می‌دهد. [۳۳]

F.F. 150

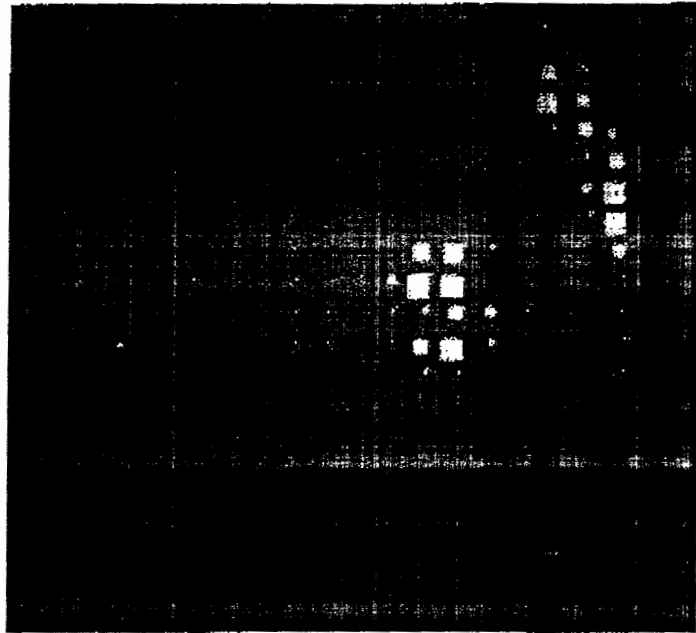
HICKAM AFB-BASEBALL FIELD-VERTICAL CONDUCTIVITY



HICKAM AFB-BASEBALL FIELD-VERTICAL INPHASE



F.F. 150



شکل ۴-۶

۴-۲- تعیین محل اهداف زیر زمین از داده‌های ژئوفیزیکی با استفاده

از شبکه‌های عصبی^۲

۴-۲-۱- خلاصه

شبکه‌های عصبی برای تخمین دور افتادگی، عمق و هدایت ناحیه حاصل از یک هدف هادی که یک تصویر بیضی‌وار الکترومغناطیسی از هدف را داده، مورد استفاده بودند. پنج نمونه شبکه عصبی متفاوت و ۵ نمایش متفاوت از تصویر بیضی‌وار مقایسه شده بودند، شبکه‌ها با تصاویر مصنوعی از هدف و آموزش داده شده بودند. و روی داده‌های صحرایی بیشتر داده‌های مصنوعی آزمایش شدند. قابلیت‌های برون یابی شبکه‌ها همچنین با داده‌های مصنوعی خارج از محدوده‌های فضایی مجموعه آموزشی، آزمایش شده بودند. ارائه داده‌ها شامل کل تصویر، تصویر زیر نمونه، راس و تغارهای مجاور، راس و مولفه‌ها از یک تبدیل فوریه سریع دو بعدی (2-D) می‌شود. نمونه‌های آزمایش شده، پس انتشار استاندارد، جستجوی تصادفی جهت‌دار، اتصال تابعی، بسط داده شده، و ترکیب نقشه خود سازمان دهی و پس انتشار بودند. برای نمونه‌های ورودی با کمتر از ۱۰۰ عنصر، جستجوی تصادفی هدایت شده و شبکه‌های اتصال تابعی، بهترین نتیجه را می‌دهند، برای نمونه‌های با بیشتر از ۱۰۰ عنصر، نقشه خود سازمان دهی به پس انتشار بیشتر دقیق بود. استفاده از کل تصویر بیضی‌وار نتایج دقیق بیشتری برای همه نمونه‌های شبکه می‌دهد. ارائه داده‌های تبدیل فوریه سریع همچنین نتایج خوبی را با یک مان محاسبه‌ای سریعتر فراهم می‌آورد. دقت‌های میانگین دور افتادگی، عمق و تولید هدایت ناحیه به اندازه ۹۷ درصد برای آزمایش و داده‌های صحرایی و ۸۸ درصد برای داده‌های برون یابی می‌توانست بدست آید.

۴-۲-۲- مقدمه

شبکه‌های عصبی یک رده وسیع از الگوریتم‌های محاسبه‌گر هستند که چندین نوع از مسائل شامل دسته‌بندی، تخمین پارامتر، پیش بینی پارامتر، تکمیل الگو، وابسته سازی الگو، فیلتر کردن و بهینه‌سازی را حل می‌کند. در این جا تمرکز ما روی شبکه‌های تخمین پارامتر خواهد بود.

۲- مطابق اس بخش ۱۲۶ آورده شده است.

شبکه‌های عصبی می‌تواند با سیستم‌های ماهر، منطق فازی و الگوریتم‌های ژنتیک برای به دست آوردن قابلیت حل مسائل به طور خیلی قوی، ترکیب شود. شبکه‌های عصبی یک علاج برای همه مسائل تشخیص الگو نیستند. رفتار آن‌ها برای فهمیدن همیشه آسان نیست. طراحی و آموزش یک شبکه هنوز بیشتر یک هنر است تا یک علم.

در این مقاله، ما توانایی چهار نمونه اضافی را برای تخمین دور افتادگی، عمق و هدایت تولید ناحیه از یک هدف دفن شده را بررسی می‌کنیم. ما علاقمند به این هستیم که بدانیم چطور شبکه‌های گوناگون با داده‌های آزمایشی داده‌های صحرایی جمع آوری شده بر روی هدف یکسان که برای آموزش استفاده شده اما با یک مدل ژئوالکتریکی زمینه مختلف، سر و کار دارد، و بالاخره، اینکه چطور شبکه‌ها می‌توانند خوب برون‌یابی کنند برای هر دو پارامترهای ورودی و خروجی.

۴-۲-۳- شبکه‌های عصبی

شبکه‌های عصبی به چندین رده بر پایه روش یادگیری، ساختار، و کنترل می‌تواند تقسیم‌بندی شوند. اغلب و تقسیم‌بندی اصلی میان شبکه‌ها، جهتی است که اطلاعات در آن جریان پیدا می‌کنند. در شبکه‌های پیشخور، عناصر پردازشی، ورودی رانها از لایه قبلی دریافت می‌کنند و خروجی را تنها به لایه بعدی ارسال می‌کنند. شبکه‌های پسخور می‌تواند یک خروجی عنصر پردازش PE را به ورودی خودش یا به یک ورودی لایه به عقب بفرستند.

برای یک تابع محرک که توسط مارن و دیگران به صورت زیر داده شده است:

$$e^{-\sum_{i=1}^n \frac{(1-c_i)^2}{2\sigma_i^2}} \quad \text{معادله (۴-۱)}$$

که c_i مرکز تابع پایه‌ای شعاعی برای لامین PE و σ_i پهنای تابع هستند.

یکی از بهترین جنبه‌های یک تابع انتقال استحکام و پایداری آن است. مشتق تابع انتقالی، نسبت خطای توزیع شده به هر اتصال را وقتی که وزنها تنظیم شده‌اند، تعیین می‌کند. تابع زیگموئید اغلب مورد استفاده قرار می‌گیرد زیرا مشتقش یک شکل دارد که به پایدارسازی شبکه کمک می‌کند.

همه شبکه‌های عصبی از یک قاعده یادگیری برای تعدیل وزنه‌های اتصالات در طول آموزش استفاده می‌کنند. شبکه پس انتشار از یک تغییر قاعده دلتا اصلاح شده برای استفاده با لایه‌های

میانی که قاعده دلتای تعمیم یافته نامیده می‌شود، استفاده می‌کند. بنابراین تغییرات وزن بین خروجی و لایه میانی توسط قانون دلتا داده می‌شود.

$$\Delta_p w_{ij} = \eta (t_{pj} - v_{pj}) i_p = \delta_{pj} \eta i_p \quad \text{معادله (۲-۴)}$$

که دلتا $\Delta_p w_{ij}$ تغییرات وزن است، η اندازه پله، t_{pj} خروجی هدف از نرون لایه خروجی j برای نمونه p است، U_{pj} خروجی محاسبه شده از نرون j برای الگوی p است، و i_p مقدار مولفه i از الگوی ورودی p است. قاعده دلتا یک تکنیک کاهش شیب است.

قاعده دلتا یک تشابه قوی با رگرسیون خطی مضاعف را در بر دارد. در رگرسیون خطی هدف پیش بینی یک متغیر وابسته از یک مجموعه متغیرهای مستقل بوسیله پیدا کردن یک مجموعه از ضرایب که خطای جمع مربعات را کمینه می‌کند، است. استون نشان داد که قاعده دلتا در حقیقت یک روش تکراری از محاسبه ضرایب رگرسیون خطی است.

تغییرات وزن بین لایه‌های میانی و ورودی بوسیله قاعده دلتای تعمیم یافته داده می‌شوند.

$$\delta_{pj} = f'(out_{pj}) \sum_k w_{kj} \delta_{pk} \quad \text{معادله (۳-۴)}$$

که $f'(out_{pj})$ مشتق تابع فعال ساز است. زیرنویس‌های j و k به ترتیب نرون‌ها در لایه‌های خروجی و میانی را نمایش می‌دهند.

یک جمله ممتنم اغلب به معادله (۲-۴) برای بیرون کشیدن شبکه از مینیمم موضعی، اضافه می‌گردد. در طول آموزش، استفاده کننده می‌تواند اندازه گام و ممتنم را تغییر دهد.

۴-۲-۴- شرح آزمایش

۴-۲-۴-۱- اکتساب داده‌ها

یک سیستم سونداژ الکترومغناطیس (EM) با تفکیک‌پذیری بالا، ایجاد شده در آزمایشگاه برای تصویرنگاری زیر سطح زمین پیشرفته (Laboratory for Advanced Subsurface Imaging) برای ثبت داده‌های صحرائی برای آزمایشات شبکه‌های عصبی، استفاده شده بود. سیستم EM با تفکیک‌پذیری بالای L-ASI، چندین خصیصه جدید، شامل درجه‌بندی همزمان با دقت بالا، و چرخش ریاضی میدانهای EM اندازه‌گیری شده سه بعدی (3-D) را ترکیب می‌کند. سیستم، برای ثبت بیضی‌وار میدان مغناطیسی در ایستگاه‌های دریافت کننده مستقر شده نزدیک

یک منبع ارسال کننده یک شکل موج فرکانس ممتد، طراحی شده بود. سیستم یک پهنای باند وسیع دارد و می‌تواند به سرعت وضعیت‌ها را در یک چگالی نمونه‌گیری فضایی بالا ثبت کند. یک توصیف دقیق از سیستم بیضی‌وار با تفکیک بالای LASI داده شده است. یک تعریف از سیستم اکتساب داده‌ها به خوبی اندازه‌گیری میدان در داده شده است.

بر پایی آزمایشات صحرائی در شکل (۴-۷) نشان داده شده است یک لوله گالوانیزه شده به ضخامت بیرونی 6CM(OD) با یک ضخامت دیوار ۰/۴۸ cm، ۶۰ m طول، در عمق ۱/۲ m و به فاصله لبه ای ۱۰m از یک فرستنده خطی طویل دفن شده بود. یک دریافت کننده، شامل سه سیم پیچ عمودی از سیمی به ضخامت تقریباً ۰/۷۵ متری با یک سیم‌پیچ درجه‌بندی در ۵۴/۷ برای هر سیم پیچ، برای اندازه‌گیری میدان مغناطیسی کلی در سه جهت در هر ایستگاه در طول یک پیمایش عرضی روی لوله، استفاده شده بود. یک دنباله از ۱۱ فرکانس‌های (6814 Hz, 4609 Hz, 3207 Hz, 2205 Hz, 1403 Hz, 1002 Hz, 601 Hz, 501 Hz, 351 Hz, 200 Hz, 150 Hz) در هر ایستگاه کاوش برای بدست آوردن یک سونداژ عمق، ثبت شده بود.

میدان مغناطیسی اندازه‌گیری شده به طور بیضی‌وار پلاریزه شده بود و این بیضی قطبش نمی‌توانست توسط مختصات کارتزین H_x ، H_y و H_z یا توسط پارامترهای بیضی e (بیضی‌وار) و α (زاویه شیب) شرح داده شده باشد.

$$e = \left| \frac{H_z}{H_1} \right| = \left| \frac{H_z \cos \alpha - H_r \sin \alpha}{H_z \sin \alpha + H_r \cos \alpha} \right| \quad \text{معادله (۴-۴)}$$

$$\tan(2\alpha) = \frac{2 \left| \frac{H_z}{H_1} \right| \cos(\phi_z - \phi_r)}{1 - \left(\frac{H_z}{H_r} \right)} \quad \text{معادله (۵-۴)}$$

که ϕ_z ، ϕ_r فازهای مولفه‌های عمودی و افقی میدان کلی هستند، و H_r مولفه میدان مغناطیسی شعاعی است. از چندین مطالعات نتیجه گرفته شده که بیضی‌وار یک توصیف‌گر حساس از هدایت زمین و عمق هدف است.

با استفاده از معادلات مقیاسی الکترومغناطیسی شرح داده شده توسط (Frischknecht 1987, Watts 1978 and Poulton 1990)، لوله دفن شده می‌توانست به عنوان یک جسم هادی بزرگتر، عمیق‌تر و با هدایت کمتر مدلسازی شود. مدل ژئوالکتریکی استفاده شده برای

آموزش شبکه عصبی در شکل ۲ نشان داده شده است. مقادیر مقیاس گذاری شده در پراترها نشان داده شده اند و این مدل از مدلسازی داده‌های میدان مقاومت ویژه دو قطبی - دو قطبی تعیین شده بود. مدل ساده شده است اما به طور کافی خصیصه‌های اساسی آنومالی را شرح می‌دهد. تفاوت اصلی بین داده‌های مصنوعی تولید شده برای مدل در شکل (۴-۷)، در دنباله‌های تصاویر اتفاق می‌افتد. دنباله تصویر میدان اندکی بالا می‌آید، در حالی که دنبال تصویر مصنوعی کاهش یافته. دنباله‌ها اثر کمی در تفسیر موقعیت هدف یا رسانایی دارد.

تصاویر بیضی‌وار بوسیله شبکه‌بندی داده‌ها با x ند در هر ایستگاه کاوش و y ند برای هر عمق پوسته‌ای ایجاد شده بودند. تصاویر آموزشی مصنوعی با یک برنامه عنصر محدود 2D ایجاد شده بودند. شکل (۴-۸)، تصاویر بیضی‌وار نمایشگر را به عنوان موقعیت و هدایت متغیرهای هدف، مقایسه می‌کند. شکل (۴-۹) موقعیت‌های فضای هدف را برای تصاویر مصنوعی که ایجاد شده بودند، نشان می‌دهد. بیست تصویر آموزشی و هشت تصویر آزمایشی برای هر سه رسانایی هدف مختلف ساخته شده بودند. مقادیر میزان شده از رسانایی $0.74s/m$ ، $1/11s/m$ و $2/22s/m$ بودند. داده‌های میدان یک تصویر آزمایشی مجزا را تشکیل داد. سه تا از هشت تصاویر آزمایشی مصنوعی (E، F و H در شکل ۴-۹) مشابه آنومالی‌هایی هستند که بیرون محدوده‌های فاصله‌ای مجموعه آموزشی قرار گرفتند. این تصاویر، توانایی شبکه را برای برون یابی آزمایش کرد.

۴-۲-۴-۲- طراحی شبکه عصبی

شبکه‌ها عهده‌دار تخمین دور افتادگی افقی، عمق و هدایت یک هدف داده شده از داده‌های بیضی‌وار چند فرکانسی، هستند ابتدا کل تصویر مصنوعی، با ۶۶۰ پیکسل از بیضی‌وار درصد، ورودی برای یک شبکه با ۶۶۰ ورودی PES بود. زمان آموزش ۱۰۰ دقیقه در هر ۱۰۰۰ تکرار بود، و بیشتر از ۲۰۰۰۰ تکرار برای همگرایی لازم بود. ما انتقاد داشتیم که کل تصویر شامل اطلاعات غیر ضروری می‌شد. بنابراین یک تلاشی ایجاد شد برای پیدا کردن آن بخش‌هایی از تصویر که شامل اطلاعات وابسته به مساله موقعیت‌یابی هدف می‌شد.

ما برای چهار توصیف‌گر تصویر تلاش کردیم:

(۱) ناحیه اطراف پیک در تصاویر مربوط به موقعیت هدف، به طوری که دامنه بیضی‌وار پیک و موقعیت افقی هر فرکانسی استخراج شده بودند.

۲) مقادیر پایین بیضی‌وار، دامنه، و موقعیت مجاور به پیک همچنین برای هر فرکانس استخراج شده بودند و متعاقباً ارجاع داده می‌شد به عنوان نمایشگر ناوه - پیک - ناوه.

۳) تصاویر همچنین توسط زیر نمونه‌سازی هر پیکسل دیگر در هر دو جهت افقی و عمودی مترکم شده بودند.

۴) نمایش داده‌های بهایی شامل بزرگی، فاز و فرکانس‌های بنیادی از یک تبدیل فویه سریع (2-D) (FFT) اقتباس شد.

نمایش FFT شامل چهار PES در لایه ورودی می‌شد؛ پیک ۳۰ نرون ورودی لازم داشت؛ ناوه - پیک - ناوه ۹۰ نرون ورودی دارد؛ زیر نمونه‌گیری ۱۷۶ نرون ورودی داشت؛ و کل تصویر ۶۶۰ نرون ورودی داشت. آن جا یک کاهش نمایی در زمانهای آموزشی با کاهش اندازه الگو وجود داشت. شبکه‌ها به وسیله سعی و خطا برای هر نمایش داده برای تولید کمترین خطاها برای آزمایش آموزش و داده‌های صحرایی ایجاد شده بودند. اندازه‌های لایه میانی برای هر شبکه و نمایش داده‌ها در جدول ۱ همراه با تعداد تکرارهای آموزشی و تکرارها در هر ثانیه در یک کارگاه، فهرست شده‌اند. شبکه‌ها ارزیابی می‌شوند بر روی اینکه چطور به طور دقیق آنها عمق، دورافتادگی و رسانایی را از داده‌های آزمایشی، مدل‌های برون یابی، و داده‌های صحرایی پیش بینی کنند.

پس انتشار، جستجوی تصادفی هدایت شده، و شبکه‌های بسط داده شد، همه دو لایه میانی در اغلب موارد برای تخمین سه پارامتر لازم داشتند. اتصال تابعی و شبکه‌های نگاشت خود سازمان ده، اساساً دو لایه میانی دارند اگر لایه‌های اتصال تابعی و کوهن شامل شده باشند. یک لایه میانی کافی برای تخمین دورافتادگی و عمق پیدا کرد. یک مقایسه از نتایج، با استفاده از یک تانژانت هیپربولیک در مقابل تابع فعال‌ساز زیگنویید برای شبکه پس انتشار و نمایش داده‌های پیک، ایجاد شده بود. تانژانت هیپربولیک، نتایج دقیق بیشتری را با مقادیر آموزشی یکسان تولید کرد. برای یک شبکه پس انتشار با استفاده از نمایش داده‌های پیک، تانژانت هیپربولیک، روی داده‌های آزمایشی دقت‌های ۹۵/۲ درصد برای دورافتادگی، ۹۷/۶ درصد برای عمق، و ۸۵/۵ درصد برای رسانایی داد. شبکه‌های یکسان. با یک تابع زیگموتیید به ترتیب دقت‌های ۸۸/۲، ۹۹/۷ و ۳۲ درصد برای دورافتادگی، عمق، و رسانایی داد. تانژانت هیپربولیک برای شبکه‌های FT، پیک و ناوه - پیک ناوه در این مطالعه استفاده شده بود. تابع زیگموتیید برای شبکه‌های زیر نمونه‌گیری شده و کل تصویر بدلائل ارائه شده در بخش نتایج استفاده شده بود.

شبکه‌های عصبی روی داده‌هایی از مدل ژئوالکتریک ساده در شکل (۴-۷) آموزش داده شده بودند اگر چه ما اعتقاد داریم مدل ژئوالکتریک واقعی ممکن است خیلی پیچیده باشد. به طوری که قبلاً بحث شد، تفاوت‌ها در مدل‌های ژئوالکتریک زمینه، به طور عمده در دنباله‌های تصاویر بیضی وار آشکار شده‌اند. بنابراین، نمایش‌هایی که از داده‌های دنباله استفاده می‌کند (e.g. زیر نمونه‌گیری شده و کل تصویر) انتظار می‌رود که دقت پایین‌تری روی داده‌های صحرائی داشته باشند. دقت‌های میانگین روی داده‌های آزمایشی برای نمایش‌های داده‌های گوناگون و شبکه‌ها در جدول (۴-۲) فهرست شده‌اند؛ جدول (۴-۳) شامل دقت‌ها برای داده‌های صحرائی است؛ جدول (۴-۴) شامل دقت‌های میانگین برای مدل‌های برون‌یابی است. به طوری که انتظار می‌رفت، ناوه - پیک - ناوه، نمایش‌های زیر نمونه‌گیری شده و کل تصویر، خیلی روی داده‌های صحرائی خیلی ضعیفتر از نمایش‌های FFT و پیک به سبب در برگیری اطلاعات از دنباله‌های منحنی‌های بیضی‌وار، انجام شدند.

اغلب شبکه‌ها برخی دشواری‌های تخمین هدایت را تجربه می‌کنند. استفاده از تنهاسه رسانایی مختلف، برای آموزش مناسب برای بیشتر شبکه‌ها کافی نبود. آنجا هیچ گرایش سازگاری از دقت در مقابل اندازه شبکه بر روی داده‌های آزمایشی موجود نبود. برای تعیین نمایش‌های داده‌هایی که بهتر بودند، جمع خطاها برای پارامترهای خروجی، از میان همه شبکه‌ها و برای مدل‌های آزمایشی، صحرائی، و برون‌یابی، میانگین‌گیری شده بود. کمترین خطای کل میانگین، بهترین نمایش. نمایش کل تصویر ۲۷/۲۱ درصد حساب شد، FT ۳۱/۵۱ درصد حساب شد، پیک ۳۵/۷۴ درصد حساب شد، زیر نمونه‌گیری ۵۷/۸۸ درصد حساب شد، و ناوه - پیک - ناوه ۶۹/۷ درصد حساب شد. نمایش کل تصویر رتبه اول را روی داده‌های آزمایشی، سوم روی داده‌های صحرائی، اول در داده‌های برون‌یابی برای محاسبه کمترین خطای سرتاسری را کسب کرد. نمایش FFT رتبه دوم برای مدل‌های آزمایشی، صحرائی، و برون‌یابی کسب کرد. نمایش پیک رتبه اول را توسط یک تفاوت وسیع روی داده‌های صحرائی، پنجم روی داده‌های آزمایشی و سوم روی داده‌های برون‌یابی کسب کرد. نمایش ناوه - پیک - ناوه رتبه سوم را روی داده‌های آزمایشی، پنجم روی داده‌های صحرائی، و چهارم روی داده‌های برون‌یابی کسب کرد. نمایش زیر

نمونه‌گیری چهارم روی داده‌های آزمایشی و صحرایی و پنجم روی داده‌های برون‌یابی محاسبه شده. نمایش کل تصویر به طور پیوسته تخمین زیادی از مقادیر پارامترهای خروجی داشت؛ نمایش زیر نمونه‌گیری شده به طور پیوسته تخمین کمی از مقادیر داشت. نمایش پیک به طور پیوسته تخمین خوبی از دور افتادگی و عمق داشت اما نتایج ترکیبی برای تولید هدایت ناحیه داد. نمایشات FFT و ناود پیک ناود هم تخمین خوب و هم بد برای مقادیر داشت.

شبکه‌های زیر نمونه‌گیری شده و کل تصویر در طول ده تکرار اول در فرآیند آموزشی هنگامی که از تانژانت هیپربولیک استفاده شده بود، از کار افتاد. یک آزمایش از رسم خروجی نرون در مقابل فعال‌سازی نشان داد که تانژانت هیپربولیک در خروجی کوچکتری از تابع زیگموتید اشباع شد. با تعداد زیادی از نرونهای اتصالات در شبکه‌های زیر نمونه‌گیری شده و کل تصویر، خروجی‌های P11 به اندازه کافی بزرگ بودند برای اشباع برای اتفاق افتادن محدوده مقادیر وزن تصویر پیوسته. مقادیر قرار گیرنده از $[-0.1, 0.1]$ به $[0.001, -0.001]$ ، و تصاویر نرخ یادگیری و موثر بودن برای احتساب از کار افتادگی به دو نیم شدند.

سه مدل برون‌یابی در شکل (۹-۴)، F_1 ، F_2 و F_3 هستند. مدل F_1 ، دو فاصله‌گذاری مدل آنسوی لبه مجموعه آموزشی است؛ مدل F_2 ، یک فاصله مدل آنسوی لبه است و مدل H یک و نیم فاصله مدل آنسوی لبه است. دقت میانگین برای همه پارامترهای خروجی شبکه‌ها برای مدل E ، $79/25$ درصد، برای مدل F_1 ، $73/9$ درصد، و برای مدل F_2 ، $84/3$ درصد بود و با وجود این حقیقت که مدل F_1 جلوتر از لبه مجموعه آموزشی نسبت به مدل F_2 است، دقت بالاتر بود. مدل F نزدیکتر به فرستنده نسبت به مدل F_2 است. و مقادیر بیضی‌وار بیشتر از هر مدل دیگر در مجموعه آموزشی هستند. ولو اینکه مدل‌های F_1 و H آن سوی لبه مجموعه آموزشی قرار گرفته‌اند، مقادیر بیضی‌وار در محدوده مقادیر پیدا شده در مجموعه آموزشی هستند. شبکه‌ها می‌توانند خیلی آسانتر با پارامترهای خروجی که بیرون محدوده آموزشی قرار گرفته‌اند نسبت به پارامترهای ورودی که انجام می‌دهد، سر و کار داشته باشند.

طراحی شبکه. در این مطالعه، توافقی بین داده‌های آزمایشی، داده‌های صحرایی، و قابلیت برون‌یابی دارد. بین سه ستیاز وزن برابر در رابطه بندی طراحی‌های شبکه گوناگون داده بودند. جمع حضاها برای سه پارامتر خروجی، برای همه نمایشات داده‌ها که میانگین گیری شده بود چون شبکه اتصال ثابتی نمی‌توانست برای نمایشات زیر نمونه‌گیری شده و کل شبکه آموزش داده شده

باشد، رتبه‌بندی‌ها مجبور شد به تقسیم شدن به دو گروه؛ یکی برای شبکه‌ها با کمتر از ۱۰۰ نرون ورودی و دیگری برای شبکه‌ها با نرونهای ورودی بیشتر. کمترین خطای کل میانگین، بهترین شبکه، برای نمایشات داده‌های کوچک، جستجوی تصادفی هدایت شده ۲۹/۵۶ درصد محاسبه شد، اتصال تابعی ۳۳/۹۱ درصد، توسعه یافته ۳۸/۷۵ درصد، نگاشت خود سازمان‌ده به پس انتشار ۴۰/۳۲ درصد، پس انتشار ۳۵/۷۷ درصد، توسعه یافته ۴۱/۳۱ درصد، به پس انتشار ۳۵/۷۷ درصد، و جستجوی تصادفی جهت‌دار ۴۸/۶۸ درصد محاسبه شد.

شبکه جستجوی تصادفی جهت‌دار، به هر حال، مقیاس‌بندی خیلی دشواری به اندازه‌های بردار الگوی بزرگتر نسبت به هر شبکه دیگر دارد. خطاهای میانگین سرتاسری از الگوهای ورودی کوچک به بزرگ نسبتاً خوب سازگار بودند برای نگاشت خود سازمان‌ده به پس انتشار، پس انتشار استاندارد بسط یافته.

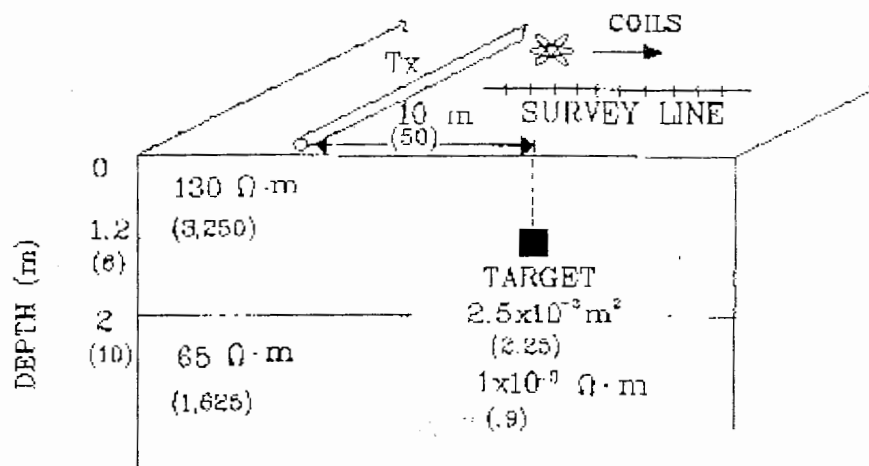
۴-۲-۶- خلاصه و نتایج

برای بردارهای الگو متجاوز از ۱۰۰ عنصر، شبکه هیبرید از نگاشت خود سازمان‌ده به پس انتشار اغلب صحیح بود. شبکه خود سازمان‌ده هیبرید می‌توانست هر دو الگوهای کوچک و بزرگ به کار برد و سوم سریعترین زمان محاسبه‌ای را داشت. در مقایسه، شبکه جستجوی تصادفی جهت‌دار مقیاس‌بندی خوب به الگوهای ورودی بزرگ انجام نداد، و شبکه اتصال تابعی به تعداد زیادی نرون برای ساخته شدن برای الگوهای بزرگ نیاز داشت. روی هم رفته، شبکه نگاشت خود سازمان‌ده هیبرید، بیشترین تطبیق پذیری را برای این آزمایش اثبات کرد.

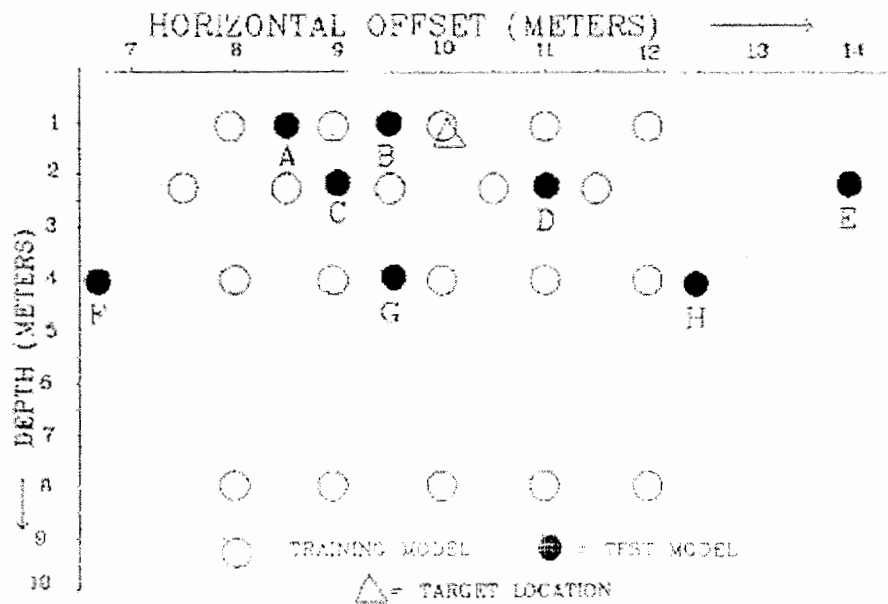
جستجوی تصادفی جهت‌دار، اتصال تابعی، پس انتشار به قرار دادن تنها دو متغیر آموزشی احتیاج داشت به طوری که آن‌ها برای برپایی آسان بودند.

نمایشات پیک و FFT زمانهای محاسبه‌ای خیلی سریعتر و دقت‌های نزدیک به آن‌ها از کل تصویر دارد. نمایش FFT اغلب برای داده‌های برون‌یابی، صحرایی و آزمایشی سازگار بود. روشهایی که تفکیک پذیری FFT را بهبود می‌بخشند مانند ماکزیمم ممکن است، شاید، بیشتر دقت نمایش را بهبود بخشد.

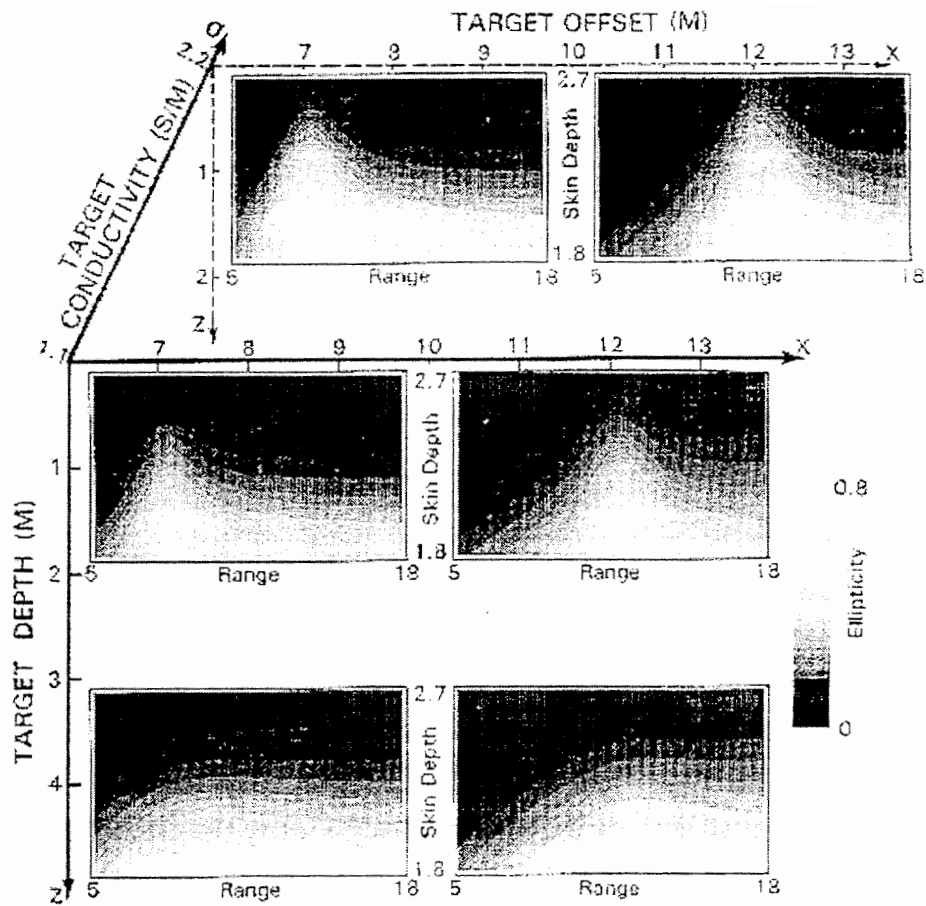
تکنیک‌های شبکه عصبی و عده‌ای برای تشخیص الگو در داده‌های ژئوفیزیکی را نشان می‌دهد. موقعیت‌یابی یک هدف در یک زمینه ثابت، نسبتاً یک مساله آسان است. کار زیادی لازم است که انجام شود برای فمیدن نیازها، محدودیت‌ها، و توانایی‌ها داده‌ها از نمونه‌های گوناگون علاوه بر اینکه چطور شبکه‌ها ممکن است در بیشتر مسائل پیچیده به کار برده شود. ترکیب هیبرید از نمونه‌ها در داخل یک سیستم شبکه عصبی می‌تواند توانایی تشخیص الگو خیلی قوی را پیشنهاد کند. توانایی برای ترکیب کردن سریع اطلاعات جدید در بیشتر الگوها وجود ندارد و آن یک محدودیت مهم از تفکیک است. محدودیت‌ها، خوشبختانه، یک پدیده پایدار در مهندسی نیستند. [۲۶]



شکل ۴-۷: مدل ژئوالکتریک ساده تعیین شده توسط مدلسازی مقاومت ویژه. مقادیر مدل اندازه گیری شده در پارانتزها نشان داده شده اند. این مدل برای آموزش شبکه عصبی استفاده شده بود.



شکل ۴-۹: توزیع مدل‌های آموزشی و مدل‌های آزمایشی برای آزمایش شبکه عصبی. مدل‌های E, F و H مربوط می‌شود به مدل‌های بیرون یابی.



شکل ۴-۸: بهم پیوستن تصاویر بیضی وار برای دو خمیدگی، دو عمق و دو هادی. محور افقی تصاویر، محدوده فرستنده است؛ محور عمودی ۱۰ الگوریتم مبنای عمق پوسته‌ای برای یک مقاومت ویژه ثابت از $130 \Omega m$ است. مرکز هر تصویر در محل هدف مربوط قرار گرفته است. با حرکت چپ به راست از بالا به پایین، موقعیت‌های هدف و مقدار هادی عبارتند از:

$$x = 8m, z = 1m, \sigma = 2.2s/m; x = 12m, z = 1m, \sigma = 2.2s/m; x = 8m, z = 1m, \sigma = 1.1s/m;$$

$$x = 12m, z = 1m, \sigma = 1.1s/m; x = 8m, z = 4m, \sigma = 1.1s/m; x = 12m, z = 4m, \sigma = 1.1s/m$$

جدول ۴-۱: مقایسه اندازه‌های لایه میانی، سرعت محاسباتی، و تعداد تکرارهای آموزش.

Data Representation		Network				
		BP ^a	DRS ^b	FL ^c	DBD ^d	SOM ^e
FFT	Hidden layer sizes (PEs)	15 × 5	15 × 5	6 ^f × 8	15 × 5	45 ^g × 5
	Computational speed (it./s)	270	256	354	233	229
	No. of training iterations	10 K	60 K	36 K	31 K	18 K
PEAK	Hidden layer sizes (PEs)	20 × 15	20 × 15	435 ^h × 10	10 × 5	25 ^g × 5
	Computational speed (it./s)	110	240	6	60	207
	No. of training iterations	20 K	18 K	6 K	30 K	24 K
TPT ⁱ	Hidden layer sizes (PEs)	30 × 10	30 × 10	4005 ^j × 10	15 × 7.5	60 ^g × 5
	Computational speed (it./s)	68	135	2	55	57
	No. of training iterations	120 K	12 K	9 K	42 K	32 K
SUB ^k	Hidden layer sizes (PEs)	20 × 3	20	—	20	60 ^g × 5
	Computational speed (it./s)	61	127	—	26	32
	No. of training iterations	10 K	14 K	—	18 K	20 K
WHOLE	Hidden layer sizes (PEs)	100 × 12	100 × 15	—	100 × 15	60 ^g × 5
	Computational speed (it./s)	4	10	—	2	10
	No. of training iterations	20 K	46 K	—	45 K	42 K

- ^a standard back propagation
^b directed random search
^c functional link
^d extended delta bar delta
^e self-organizing map to back propagation
^f number of PEs in functional link layer
^g number of PEs in Kohonen layer
^h trough-peak-trough
ⁱ subsampled

جدول ۴-۲: خلاصه دقت‌های درصد میانگین بر روی داده‌های آزمایشی برای ارائه داده‌های استفاده شده در آموزش شبکه

Data Representation		Network				
		BP ^a	DRS ^b	FL ^c	DBD ^d	SOM ^e
FFT	Offset	98.88	96.05	98.12	99.00	98.31
	Depth	99.23	98.90	98.51	98.93	99.30
	Cond.	85.62	89.60	98.20	96.02	95.16
PEAK	Offset	85.19	87.23	85.42	87.82	88.55
	Depth	96.39	97.81	98.15	96.58	96.21
	Cond.	88.18	87.02	85.86	96.16	92.21
TPT ^f	Offset	94.53	96.75	96.86	98.89	95.91
	Depth	88.99	97.24	97.10	99.29	99.82
	Cond.	88.51	98.61	91.34	99.03	87.36
SUB ^g	Offset	96.51	96.91	—	95.88	99.28
	Depth	99.47	97.33	—	96.53	96.42
	Cond.	84.48	78.50	—	93.76	85.72
WHOLE	Offset	97.25	99.01	—	99.79	99.96
	Depth	99.73	99.86	—	99.80	99.31
	Cond.	96.01	94.73	—	99.32	95.85

- ^a standard back propagation
^b directed random search
^c functional link
^d extended delta bar delta
^e self-organizing map to back propagation
^f trough-peak-trough
^g subsampled

جدول ۳-۴: خلاصه درصد دقتها بر روی داده‌های صحرائی برای ارائه داده‌های استفاده شده در آموزش شبکه

Data Representation	Network					
	BP ^a	DRS ^b	FL ^c	DBD ^d	SOM ^e	
FFT	Offset	95.04	74.11	83.04	98.52	95.49
	Depth	98.72	99.92	98.14	99.44	98.47
	Cond.	96.51	99.89	71.22	80.05	96.47
PEAK	Offset	95.40	98.43	96.71	97.69	95.97
	Depth	99.13	99.56	97.31	99.46	94.83
	Cond.	97.80	97.39	98.57	90.91	99.64
TPT ^f	Offset	86.96	98.37 ^g	77.45	95.80	74.40
	Depth	98.60	98.18	96.62	94.72	97.76
	Cond.	05.00	95.64	67.47	42.86	13.6
SUB ^h	Offset	97.85	79.23	—	81.33	92.59
	Depth	74.32	63.44	—	89.80	94.34
	Cond.	62.79	87.30	—	71.43	83.33
WHOLE	Offset	89.05	85.56	—	92.71	92.67
	Depth	92.78	96.30	—	95.32	99.68
	Cond.	89.09	86.68	—	44.51	81.36

^astandard back propagation

^bdirected random search

^cfunctional link

^dextended delta bar delta

^eself-organizing map to back propagation

^ftrough-peak-trough

^gsubsampled

جدول ۴-۴: خلاصه درصد میانگین دقتها بر روی داده‌های برون‌یابی شده برای نمایش داده‌های استفاده شده در آموزش شبکه

Data Representation	Network					
	BP ^a	DRS ^b	FL ^c	DBD ^d	SOM ^e	
FFT	Offset	66.70	71.95	79.69	86.81	78.84
	Depth	95.67	96.50	97.37	93.99	97.59
	cond.	69.14	78.13	83.58	57.00	76.79
PEAK	Offset	79.75	80.62	80.87	76.76	77.58
	Depth	86.62	87.59	88.84	89.06	87.18
	Cond.	64.15	75.80	76.80	76.53	64.49
TPT ^f	Offset	70.72	66.89	85.82	69.76	79.47
	Depth	77.36	92.12	92.24	81.07	92.25
	cond.	80.22	73.69	73.56	69.27	83.06 ^g
SUB ^h	Offset	82.97	67.36	—	88.97	69.57
	Depth	83.45	76.40	—	88.77	85.03
	Cond.	44.52	48.23	—	48.08	53.14
WHOLE	Offset	81.70	77.55	—	88.84	78.65
	Depth	92.61	95.83	—	93.18	94.98
	Cond.	80.09	78.23	—	82.13	82.99

^astandard back propagation

^bdirected random search

^cfunctional link

^dextended delta bar delta

^eself-organizing map to back propagation

^ftrough-peak-trough

^gsubsampled

فصل پنجم

کاربرد شبکه‌های عصبی در لرزه شناسی

۵-۱- شبکه‌های عصبی برای تخمین و پیش بینی^۱

۵-۱-۱- وارونه‌سازی داده‌های لرزه‌ای توسط FNNها

یادگیری و تعمیم توانایی‌های FNNها آن‌ها را یک ابزار جالب برای خودکار سازی تخمین سرعتها از داده‌های لرزه‌ای ساخته است. FNNها در تحلیل سرعت برای انجام تصحیح اختلافهای زمانهای ورود امواج مستقیم (NMO) از داده‌های انعکاسی لرزه‌ای مورد استفاده قرار گرفتند. روت و تارانولا (Roth and Tarantola 1994) ابتدا مساله وارونه‌سازی شکل موج لرزه‌ای را با استفاده از FNNهای آموزش یافته با روش پس انتشار مطالعه کرد. اینجا، ما امکان تخمین سرعتهای فاصله از داده‌های انعکاسی لرزه‌ای با FNNهای آموزش یافته با یک برنامه آموزشی هیبرید را مطالعه می‌کنیم.

۵-۱-۱-۱- آزمایشات مصنوعی

ابتدا لایه‌ای افقی همگی را فرض می‌کنیم. لرزه نگارهای ترکیبی استفاده شده برای آموزش شبکه تنها شامل انعکاسات اولیه موج P هستند برای اینکه مسیر پرتوها به سادگی توسط قانون

^۱ - مطالب این بخش از منبع [۱۸] آورده شده است.

اسنل قابل محاسبه هستند. مقصود اصلی آموزش یک FNN برای تخمین سرعتها از زمانهای سیر با استفاده از شکل موجهای لرزه‌ای است. از آن جا که زمانهای سر مورد توجه هستند، هر انعکاسی، یک دامنه واحد را تعیین می‌کند. در عمل، این با استفاده از ردهای لرزه‌ای بعد از یک تصحیح شبکه خودکار به کار برده شده، انجام گرفته شده است.

به طوری که در بخش قبل گفته شد، دو مجموعه از مدلها به طور تصادفی از محدوده‌های سرعت قبلا تعریف شده، انتخاب شدند. مجموعه اول از مثالها برای آموزش شبکه استفاده شد، در حالی که مجموعه دوم برای آزمایش این که آیا شبکه می‌تواند مدل‌های سرعت را از داده‌های جدید با استفاده از وزنهای منتج شده بدست آورد، استفاده شد. مدلها برای تولید داده‌های ترکیبی شامل ۶ لایه با سرعتها و ضخامت‌های شرح داده شده در جدول (۵-۱) استفاده شد. پنج پارامتر مدل، سرعت صوتی لایه‌های ۲-۶ (ضخامت‌ها ثابت هستند)، خروجی شبکه را تعریف می‌کند. محدوده‌های سرعت با عمق افزایش می‌یابد اما وارونه‌سازی محلی سرعتها ممکن است در مدلها اتفاق بیافتد.

داده‌های ترکیبی محاسبه شده از این شبیه‌سازی مدلها C MP جمع‌آوری می‌کند با داشتن هر ۱۶ رد با یک فاصله ۰/۱ کیلومتری دریافت کننده‌ها و یک مبدا نخستین ۰/۱۵ کیلومتری، سرعت نمونه‌گیری ۰/۰۱ ثانیه است و هر رد شامل ورودی‌های بین ۰/۳۴ ثانیه و ۱/۵۸ ثانیه برای کل ۱۲۸ نمونه در هر رد می‌شود. با دادن ۲۰۴۸ نمونه در هر جمع‌آوری، موجک لرزه‌ای در هم پیچیده شده مشابه یک موجک با یک فرکانس مرکزی ۲۲ Hz است. پالس‌ها تنها یک قطبش دارند، بطوری که شامل اطلاعات دامنه در آزمایش نیست. شکل (۵-۱) در جمع‌آوری ترکیبی دلخواه که از دو مجموعه آموزشی با هم انتخاب شده را با توزیع سرعت با عمق برای هر مدل نشان می‌دهد.

پیرو یک تحلیل مشابه به چیزی که در بخش قبلی برای انتخاب اندازه شبکه و مجموعه آموزشی مناسب شرح داده شد، یک FNN تعریف شده توسط ۱۴ نرون در لایه میانی، با ۴۰۰ نمونه از داده‌ها آموزش داده شده بود. این حقیقت که بیشتر نرونها و مثالهای آموزشی نیازمند هستند برای پیش بینی دقیق در مقایسه با مساله وارونه مقاومت ویژه، با اعداد بزرگتری از داده‌های استفاده شده برای آموزش ارتباط دارد. تبعاً، همچنین آموزش مناسبتر انتظار می‌رود. ما از روش

آموزشی هیبرید (چندگانه) شرح داده شده در بالا برای بدست آوردن وزنه‌های بهینه شبکه استفاده کردیم. بیشینه خطای به دست آمده با روش هیبرید در شکل (۵-۲) نشان داده شده است.

VFSA در ۵۰۰ تکرار اول، برای بدست آورد یک نقطه آغازین برای روش پس انتشار و در دو مورد بیشتر برای آشفته‌سازی راه حل بدست آمده توسط روش گرادیان مورد استفاده قرار گرفته بود. هدف روش VFSA گرادیان ترکیب شده، بدست آوردن یک مجموعه از وزنه‌ها (یا بیشتر) به همان اندازه قابل قبول از آن‌هایی که با پس انتشار ساده با انجام محاسبات کمتر بدست آمده، است. شکل (۵-۱) دو مثال از مجموعه CMP به طور دلخواه انتخاب شده از مجموعه آموزشی با ارزیابی سرعت شبکه را نشان می‌دهد. از این شکل واضح است که بعد از آموزش، شبکه تقریباً به طول کامل سرعت‌های درست هر لایه را انطباق داده است.

ابعاد ماتریس لایه‌های میانی توسط تعداد نمونه‌های یک مجموعه ساده و تعداد نرونهای در لایه میانی داده می‌شود، در حالی که ابعاد ماتریس وزن خروجی توسط تعداد نرونهای در لایه میانی و تعداد لایه‌های سرعت استفاده شده در وارونه‌سازی، داده می‌شود. شکل‌های (۵-۳ a و b) به ترتیب قطعه‌های زنگی ماتریس‌های وزن لایه میانی و خروجی شبکه آموزش یافته را نشان می‌دهد. قطعات رنگی برای هر وزن از لایه میانی (شکل ۵-۳ a) ویژگی‌های مشخصی از داده‌های آموزشی (ارتباط کل مسیرهای شبه هندلولی). جالب توجه است که در اکثر این قطعات، بیشتر حادثه‌های یکسان در زمانهای مختلف می‌شود. توزیع فراوانی ساده قبل تشخیص هستند. این می‌تواند این حقیقت را شرح دهد که داده‌های آموزشی شامل منحنی‌های شبه هندلولی برای حادثه‌های یکسان در زمان‌های مختلف می‌شود. توزیع فراوانی برای ماتریس وزن لایه خروجی (شکل ۵-۳ b) مسیر ناگهانی را با وزنه‌ها با افزایش بزرگی لایه‌های عمیق‌تر را آشکار می‌سازد.

شکل (۵-۴ a) نتایج پیش‌بینی را از یک مجموعه ترکیبی انتخاب شده از مجموعه پیش‌بینی نشان می‌دهد، که شامل ۱۰۰ مثال می‌شود. ارزیابی‌ها برای هر لایه در شکلی از خطوط مستقیم منطبق شده از نقشه‌های متقاطع از سرعت صحیح در مقابل سرعت ارزیابی شده در شکل (۵-۴ b) نشان داده شده‌اند. ارزیابی‌های خطا، که مشابه با انحراف معیار میانگین برای هر کدام از لایه‌ها است، با استفاده از میله‌های عمودی رسم شده‌اند، لایه‌های ۲ و ۳ بهتر از لایه‌های زیرین ۴ تا ۶ پیش‌بینی شده بودند. (تقریباً ۰/۰۲ km/s برای لایه‌های ۲ تا ۳ در مقابل ۰/۱۷ km/s برای لایه‌های ۴ تا ۶).

شبکه آموزش یافته همچنین با استفاده از داده‌های با ۱۰٪ نویز افزایشی ناهمبسته اضافه شده، تست شد. نتایج پیش‌بینی در شکل (۵-۵) نشان داده شده‌اند. خطاهای سرعت تقریباً دو برابر آنهایی هست که برای مورد داده‌های بدون نویز بدست آمده است (تقریباً ۰/۰۵ و ۰/۲۹ k m/s به ترتیب برای لایه‌های ۲-۳ و ۴-۶). انحرافهای خطوط مستقیم منطبق شده شباهت به آنهایی دارد که برای پیش‌بینی‌های بدون نویز به دست آمده، برای بهبود این نتایج، یک FNN جدید با استفاده از داده‌های با ۱۰٪ نویز افزایشی ناهمبسته آموزش داده شده. نتایج پیش‌بینی در شکل (۶-۵) نشان داده شده‌اند. به طوری که مشاهده شد، پیش‌بینی‌ها کمی بهتر از نتایج قبلی هستند (تقریباً ۰/۰۴ و ۰/۲۷ km/s به ترتیب برای لایه‌های ۲-۳ و ۴-۶).

این مثالها نشان می‌دهند که یک FNN آموزش یافته می‌تواند برای پیش‌بینی تغییرات در سرعت تراکمی از داده‌های شکل موج لرزه‌ای، استفاده شده باشد. مشابه مساله مقاومت ویژه، یک برنامه وارونه‌سازی تراکمی می‌تواند برای بهبود تخمین‌های شبکه مورد استفاده قرار گیرد. هر چند، پیچیدگی برای تولید دوباره داده‌های لرزه‌ای ترکیبی که به طور نزدیکی مشابه است. روش ارائه شده در این بخش برای یک مطالعه وارونه‌سازی جهت یافته هدفمند در جایی که یک اطلاعات اولیه ویژه بدست آمده از نقاط اندازه‌گیری انتخاب شده می‌تواند برای آموزش شبکه مورد استفاده قرار گیرد، خیلی مناسب است، و شبکه آموزش یافته می‌تواند سپس برای پیش‌بینی مدل‌های زمین در نقاط مشاهده‌ای دیگر مورد استفاده قرار گیرد.

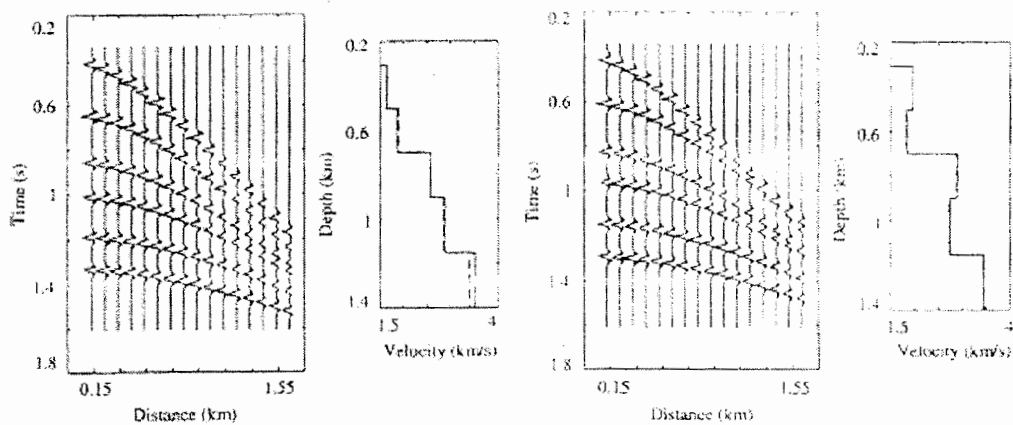
۵-۱-۲- بحث و نتیجه گیری

در کاربرد لرزه‌ای، یک FNN برای به دست آوردن مدل‌های سرعت صوتی ID از داده‌های لرزه‌ای ترکیبی استفاده شد. نتایج نشان دادند که شبکه‌های آموزش یافته می‌توانند رابطه داده‌های لرزه‌ای شکل موج با سرعت واسطه ای را یاد بگیرند. پیش‌بینی‌های سرعت برای لایه‌های بالایی (لایه‌های ۲-۳ خیلی دقیق‌تر از آن‌ها برای لایه‌های زیرین لایه‌های ۴-۶) بودند. این حقیقتی را بیان می‌کند که لایه‌های کم عمق بیشتری از لایه‌های عمیق دارد. نتایج همین طور نشان می‌دهند که یک FNN آموزش یافته با استاندارد از داده‌های با ۱۰٪ ناهمبسته اضافه شده به طور مرزی کارایی بهتری از یک شبکه آموزش یافته با داده‌های بدون نویز دارد.

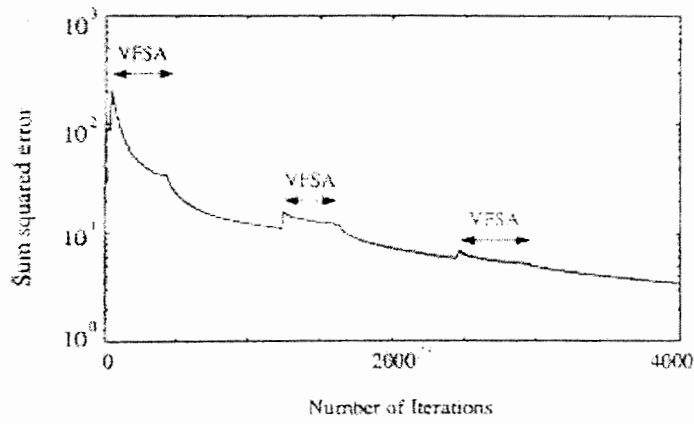
در هر دو کاربرد مقاومت‌ویژه و لرزه‌ای، نقشه‌های رنگی توزیع دانه ماتریس وزن لایه میانی برخی تشابهات با داده‌های آموزش را نشان می‌دهد. از آزمایشات انجام شده ملخص نیست که چطور سازماندهی وزن‌ها برای فرایند یادگیری FNN شرکت داده می‌شود. [۸]

جدول ۵-۱- حدود سرعت و ضخامتهای مدلهای فرض شده برای آموزش شبکه

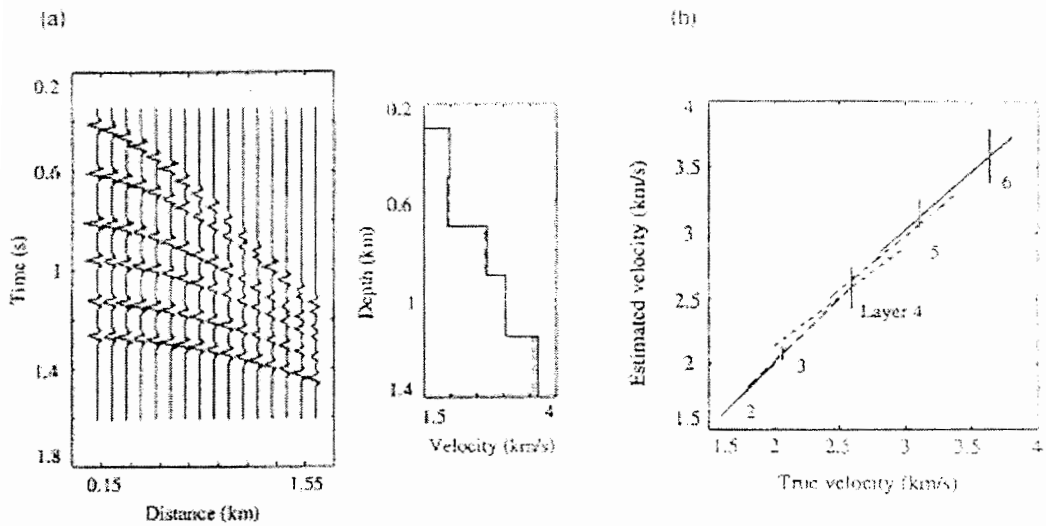
No. of layer	v_{min} (km/s)	v_{max} (km/s)	h (km)
1	1.5	1.5	0.30
2	1.6	2.0	0.20
3	1.8	2.6	0.20
4	2.0	3.0	0.20
5	2.4	3.4	0.25
6	2.8	3.8	0.25



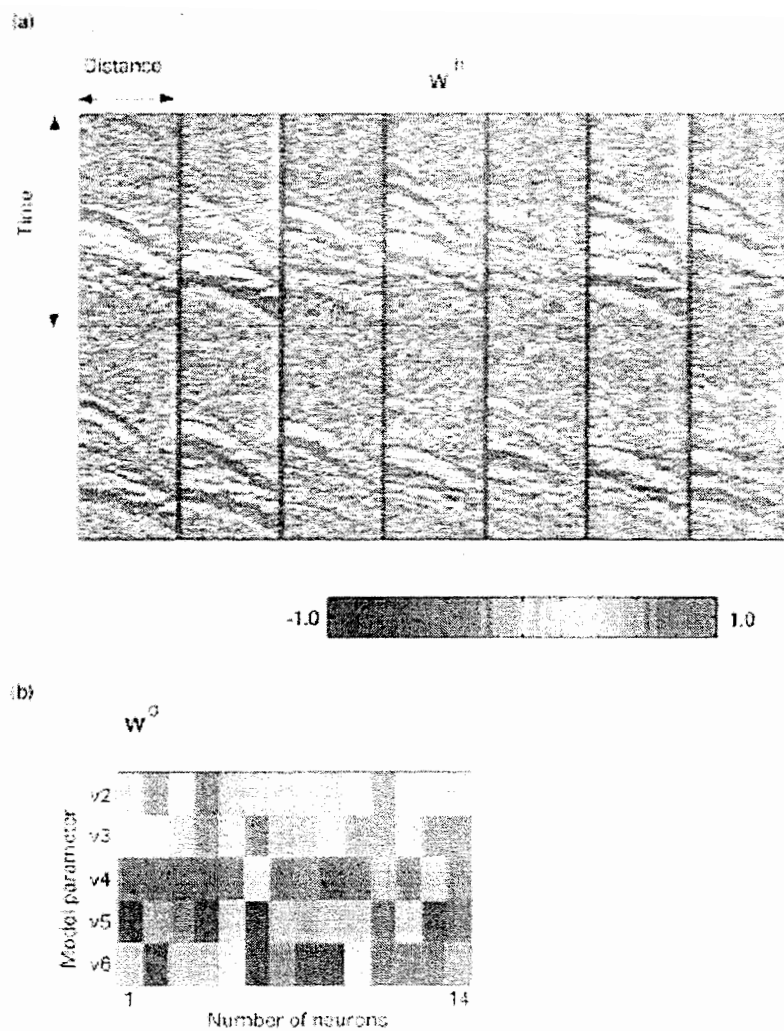
شکل ۵-۱: دو مجموع CMP مصنوعی و مدلهای سرعت متناظر بدلیخواه انتخاب شده از نمونه‌های آموزشی. سرعت‌های حقیقی و تخمین زده شده دست‌آمده از شبکه آموزش یافته بترتیب با خطوط نقطه چین و پر کنار داده‌های مصنوعی رسم شده‌اند.



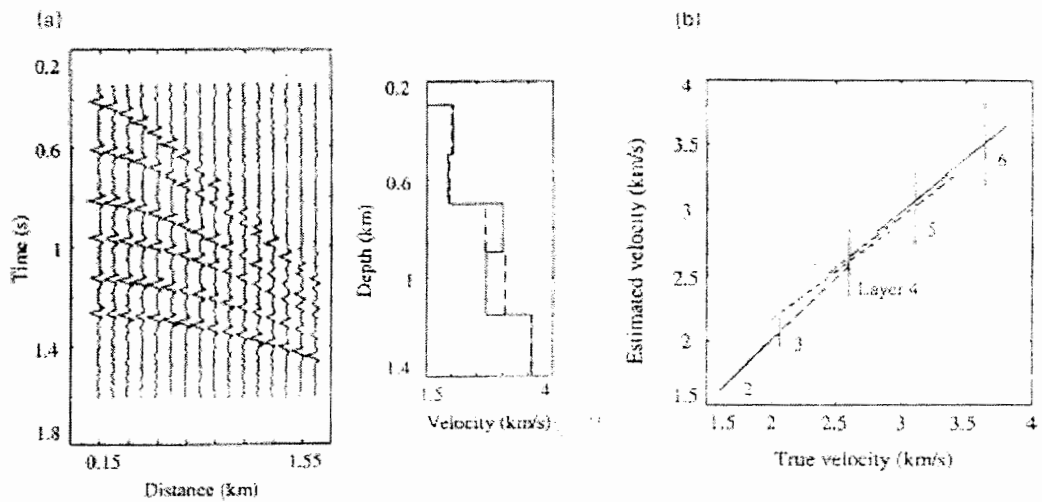
شکل ۵-۲: پیشینه خطای آموزشی بدست آمده با روش آموزشی هیبرید



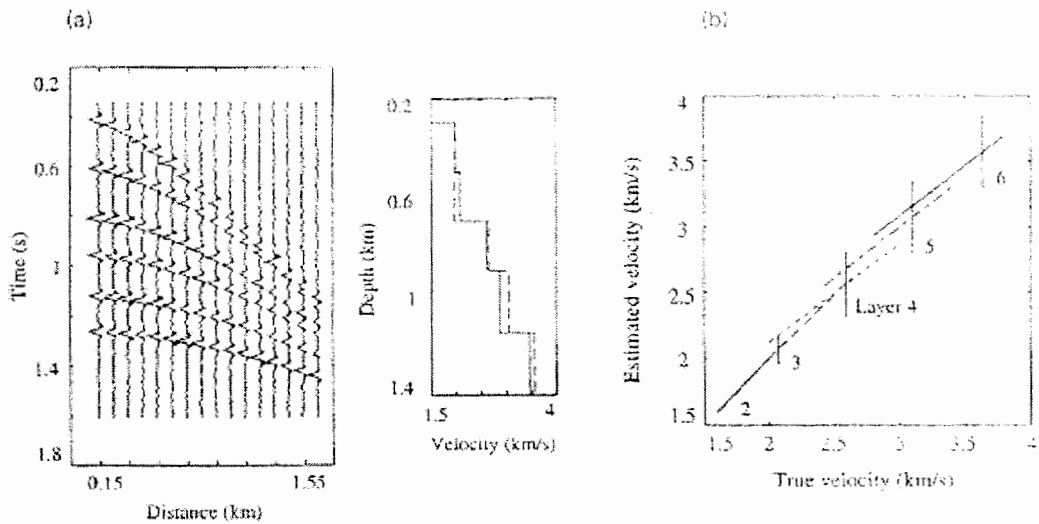
شکل ۵-۴: (a) نتایج پیش بینی بدست آمده بعد از آموزش یک FNN ۱۴ نرونی. یک مجموعه عکس نوعی از نمونه‌های پیش بینی و تخمین‌های سرعت متناظر. مدل واقعی توسط یک خط نقطه‌چین نشان داده شده است. (b) رسم مقاطع سرعت واقعی در مقابل سرعت تخمینی برای هر لایه. خطوط مستقیم از سرعت‌های پیش بینی شده برای ۱۰۰ نمونه پیش بینی تشکیل یافته‌اند. تخمین‌های خطا برای هر لایه با میل‌ها نشان داده شده‌اند.



شکل ۵-۳: (a) قطعه‌های رنگی توزیع دامنه ماتریس لایه میانی را نشان می‌دهد. (b) قطعه رنگی توزیع دامنه ماتریس لایه خروجی را نشان می‌دهد.



شکل ۵-۵: (a) نتایج پیش بینی بدست آمده بعد از اضافه کردن ۱۰٪ نویز به نمونه داده‌های پیش بینی از شکل (۴-۵). داده‌های آموزشی بدون نویز هستند. جمع CMP نوعی از نمونه‌های پیش بینی و تقریب‌های سرعت متناظر. مدل حقیقی توسط خط تیره نشان داده شده است. (b) رسم مقاطع سرعت واقعی در برابر سرعت تخمینی برای هر لایه. تقریب‌های خطی برای هر لایه با میل‌ها رسم شده‌اند.



شکل ۵-۶: (a) نتایج پیش بینی بدست آمده بعد از اضافه کردن ۱۰٪ نویز به داده‌های آموزشی به خوبی داده‌های پیش بینی. این نتایج قابل مقایسه با نتایج در شکل‌های (۴-۵) و (۵-۵) هستند. (b) رسم مقاطع سرعت واقعی در برابر سرعت تخمینی برای هر لایه.

۵-۲- تنظیم رد لرزه‌ای^۲

برای نشان دادن اینکه شبکه‌های عصبی می‌توانند برای چالش‌های بیشتر به کار برده شوند، مساله دنیای حقیقی از نمونه نموداری طرح‌ریزی شده و ساده، ما خروجی یک مدل اولیه شبکه پس انتشار طراحی شده را برای نشان دادن ردهای نوپزدار در مجموعه داده‌های واقعی لرزه‌ای نشان می‌دهیم (شکل ۵-۷). ساختار شبکه باری این دایراشگر رد ۵۱۷ رد ورودی تماماً متصل شده به دو ند خروجی. هیچ نرون لایه میانی در اینجا وجود ندارد. ندهای ورودی شامل ۵۱۲ نمونه برای طیف دامنه تبدیل فوریه سریع رد ارزیابی شده و ۵ پارامتر آماری (فرکانس رد میانگین، دامنه مطلق میانگین، انرژی رد میانگین، مقدار همبستگی متقابل ماکزیمم رد تحلیل شده بار ردهای مجاور، و میزان کاهش انرژی). دو نرون خروجی برای تولید یکی از دو حالت زیر آموزش یافته‌اند: (۵ و ۱) اگر رد «خوب» باشد، و (۱ و ۰) اگر «بد» باشد استفاده کننده چندین نمونه رد خوب (G) و بد (B) برای مجموع داده‌ها تنظیم شده برای آموزش شبکه انتخاب می‌کند. بعد از کامل شدن آموزش، شبکه عصبی ردهای باقیمانده در این نگاشت را با استفاده از مورد پایین تر g و b برای نشان دادن ویرایش‌های رد بد و خوب ویرایش می‌کند.

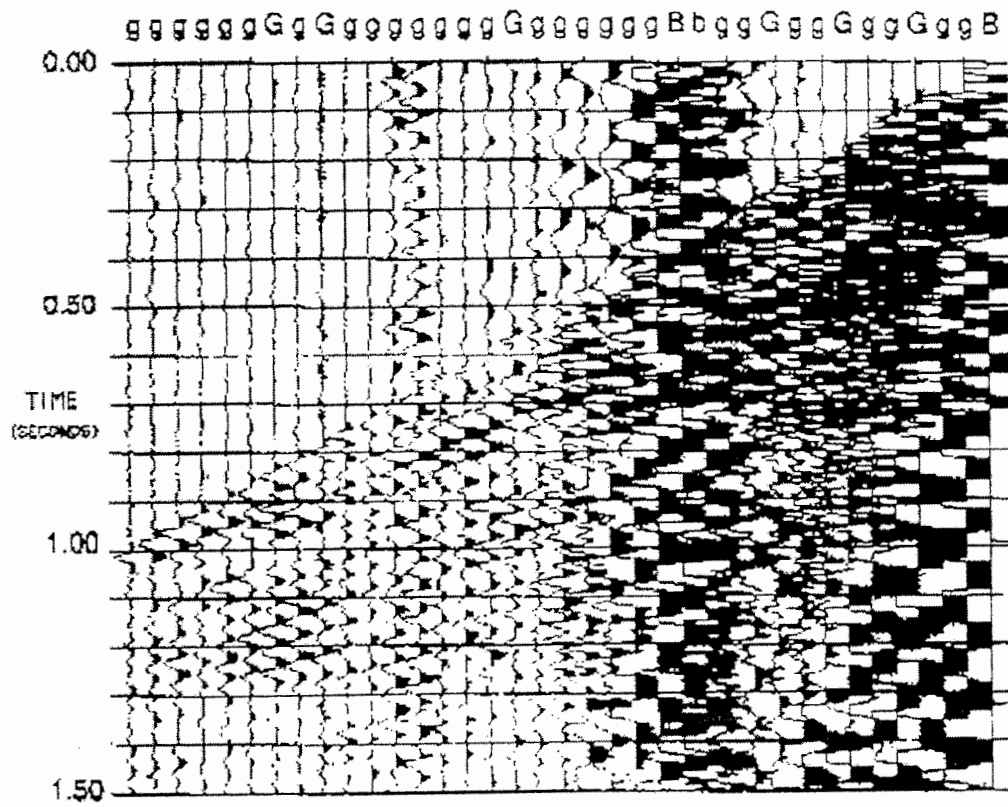
اجرای این مدل اولیه سیستم خوب بود. ویرایش‌های رد شبکه عصبی رضایت استفاده کننده را در بیشتر از ۸۰ درصد موارد جلب کرده است. در مواردی که رضایت مندی حاصل نشده بود، شبکه معمولاً رد مشکوک را «بد» دسته‌بندی کرده در حالی که استفاده کننده آن را «خوب» نامیده است.

۵-۲-۱- کاربردهای دیگر

شبکه‌های عصبی یک شکل از بازشناسی نمونه خودکار هستند در چیزی که یک مجموعه از نمونه‌های ورودی که مرتبط است با یک خروجی توسط یک تبدیل کدبندی شده در وزنهای شبکه. ما می‌توانیم همه را برای شناختن یک حادثه انکساری اولین شکست روی یک لرزه‌نگار یاد بگیریم، اما آن برای توضیح الگوریتمی اینکه چگونه ما این را انجام می‌دهیم بی‌نهایت مشکل است. بعلاوه برای پیک اولین شکست و ویرایش رد لرزه‌ای قبلاً بحث شده، شبکه‌های عصبی می‌توانند برای

۲- مطالب این بخش از منبع [۱۸] آورده شده است.

انجام تحلیل‌های سرعت، کنترل کیفیت پردازش لرزه‌ای، حادثه‌های انعکاسی لرزه‌ای متوالی از میان بلوک‌های گسل، تحلیل رنگ نگارهای گاز، و تصویر سائیدگی کم خطری آموزش یابد. البته اینها فقط اندکی از کاربردهای ژئوفیزیکی پتانسیلی تکنولوژی شبکه عصبی بررسی شده در صنعت لابراتورهای R & D هستند. [۱۸]



شکل ۵-۷: یک تبت تصویر لرزه‌ای ویرایش شده با ردهای خوب (G) و بد (B) استفاده شده برای آموزش شبکه عصبی. ویرایش‌های رد خوب و بد انجام شده توسط شبکه عصبی بعد از آموزش بترتیب توسط g و b نشان داده شده‌اند.

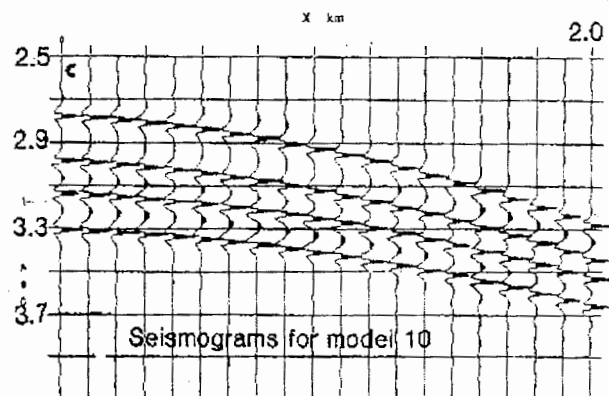
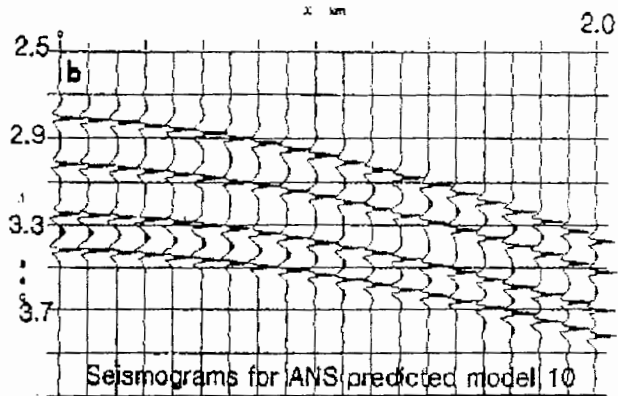
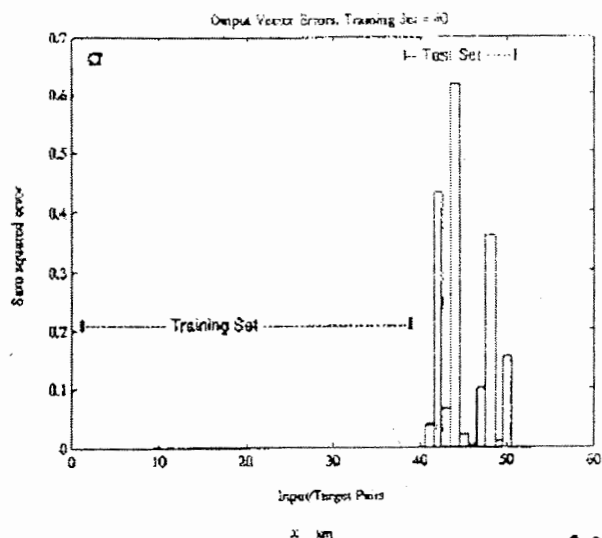
۵-۳- کاربرد شبکه‌های عصبی برای داده‌های لرزه‌ای^۳

اگر چه به طور ایده‌آل ما وارونه‌سازی برای خواص کشسان (الاستیک) از داده‌های انعکاسی لرزه‌ای با زمان سیر، اطلاعات فاز و دامنه، دوست داریم، در کار حال حاضر ما خودمان را محدود می‌کنیم به مساله تخمین سرعت موج تراکمی با استفاده از اطلاعات زمان سیر از لرزه نگارهای انعکاسی. برای اجتناب از زمان سیر و در زمان یکسان نگهداری الگوریتم آماده برای یک مساله خیلی عمومی و مشکل ما ردهای لرزه‌ای مصنوعی برای ۵۰ مدل ID هر کدام با ۵ لایه تولید کردیم به طوری که لایه بالایی یک لایه آب با ضخامت ۲km و $V_p = 1/5 \text{ km/s}$ است. سه لایه بیشتر هر کدام با ضخامت ۲۰۰ متر با سرعتهای صوت تولید شده به طور تصادفی در محدوده $1/5 - 2/8 \text{ km/s}$ در مدل استفاده شده بودند. بیست لرزه‌نگار مصنوعی با فاصله کدگذاری ۱۰۰ متر برای هر مدل با ۲۵۶ نمونه زمانی هر کدام، با ساختن $20 \times 256 = 5120$ نمونه در هر بردار ورودی، محاسبه شده بودند. نظر به اینکه ما تنها علاقمند به زمان سیر هستیم، هر انعکاس یک دامنه واحد علامت‌گذاری شده بود و یک تابع منبع در محدوده فرکانسی ۱۰-۶۰ Hz در ترکیب استفاده شده بود. یک مجموعه از ۴۰ الگو برای آموزش شبکه استفاده شده بود، و ده تای دیگر برای آزمایش شبکه (شکل ۵-۸ a). بر این شبکه، ۲۰ واحد میانی در ۵۰۰۰ تکرار با استفاده از یک نرخ یادگیری $\eta = 0.005$ و یک ممنتوم $\alpha = 0.95$ ، آموزش داده شده بودند. مجموعه آموزشی کاملاً دوباره بازسازی شده بود، اما می‌تواند به نظر برسد که یک تفاوت قابل ملاحظه بین مجموعه آزمایشی داده شوند. جدول (۵-۲) یک مقایسه بین ۱۰ مدل سرعت آزمایشی پیش بینی شده توسط شبکه و مدل‌های سرعت صحیح را نشان می‌دهد. شکل (۵-۸ b) یک مقایسه از لرزه‌نگارهای مصنوعی برای مدل و مدل پیش بینی شده ۱۰ (فهرست شده در جدول ۵-۲) از مجموعه آموزشی را نشان می‌دهد. به طوری که می‌تواند به نظر رسیده باشد، شبکه یک مدل محاسبه کرد که لرزه‌نگارهایی تولید می‌کند که چندین عدم مطابقت در زمان سیر در مقایسه با آنهایی که با مدل صحیح محاسبه شده‌اند، دارد. این علت خطا در پیش بینی برای این مجموعه داده‌ها است که به کلی بزرگ بود. [۳۲]

۳- مطالب این بخش از منبع [۳۲] آورده شده است.

جدول ۵-۲: سرعت‌های پیش‌بینی شده و واقعی برای مجموعه داده‌های آزمایشی استفاده شده در تخمین ANS لرزه‌نگارها

	* True Velocity (Km/s)				
	** Computed Velocity (Km/s)				
		Layer 1	Layer 2	Layer 3	Layer 4
Model 1	*	1.5	2.4	1.66	2.46
	**	1.54	2.42	1.92	2.39
Model 2	*	1.5	2.14	1.6	1.71
	**	1.6	1.5	2.92	1.96
Model 3	*	1.5	2.47	2.06	1.5
	**	1.52	2.55	1.97	2.05
Model 4	*	1.5	1.58	2.2	2
	**	1.47	2.66	1.58	3.8
Model 5	*	1.5	2.31	2.35	2.42
	**	1.46	2.18	2.19	2.3
Model 6	*	1.5	2.56	1.98	2.39
	**	1.44	2.61	2.03	2.33
Model 7	*	1.5	1.63	2.17	2.5
	**	1.57	2.24	2.72	2.6
Model 8	*	1.5	1.59	1.7	2.44
	**	1.49	2.75	2.01	3.53
Model 9	*	1.5	2.12	2.09	2.17
	**	1.49	2.1	2.15	2.31
Model 10	*	1.5	1.87	2.58	2.4
	**	1.4	1.87	1.76	2.3



شکل ۵-۸: آموزش و پیش‌بینی ANS لرزه‌نگارها:

(a) نمودار پیشینه مربع خطای برای مجموعه داده‌های آموزشی و آزمایشی، (b) لرزه‌نگارهای مصنوعی برای مدل پیش‌بینی شده ANS،

(c) لرزه‌نگارهای مصنوعی برای مدل حقیقی ۱۰ از مجموعه آزمایشی.

۵-۴- استفاده از شبکه‌های عصبی برای وارونه سازی داده‌های

لرزه‌ای

۵-۴-۱- خلاصه

در یک مثال اولیه، یک شبکه عصبی با یک تکنیک بهینه‌سازی مونت کارلو وابسته به عمق از یک لرزه‌ای و سرعت لایه از زمان سرهای ثبت شده در سطح زمین، آموزش یافته است. فرایندهای یادگیری، خیلی سریع همگرا می‌شوند و همچنین شبکه‌های عصبی آموزش یافته توانایی‌های درون‌یابی را خیلی خوب برای زمانهای مسیر شناخته شده، نشان می‌دهد. با تکنیک مشابه، شبکه دیگری برای محاسبه یک پروفیل سرعت - عمق پیوسته از یک لرزه‌نگار رد سازه، آموزش یافته می‌یابد. مدل زمین یک پروفیل عمق - سرعت یک بعدی بود، شامل دو در عمق‌های ثابت برای همه نمونه‌ها اما هر سه سرعت‌های لایه تصادفی بودند. فرایندهای یادگیری بر پایه ۱۴ نمونه انجام شده شبکه آموزش یافته یک مقاومت ویژه خیلی خوب را در مقابل لرزه‌نگارهای نویزدار نشان می‌دهد.

۵-۴-۲- مقدمه

داده‌های لرزه‌ای ممکن است برای تفسیر و پردازش پیچیده و مشکل باشند. مساله وارونه‌سازی لرزه‌ای یکی از مشکل‌ترین تجزیه و تحلیل‌ها است. الگوریتم‌های وارونه‌سازی، برای کار کردن خوب به نظر می‌رسند و نتایج خوبی را برای تباین‌های مقاومت ظاهری نتیجه می‌دهد، اگر مدل سرعت آغازین یک تخمین خوب از ساختار زیرزمینی حقیقی باشد. تکنیک‌های زیادی برای بازیافت این مدل سرعت زمینه شناخته شده‌اند، اما همه آن‌ها به یک زمان محاسبه ای مهم، یا مقدار زیادی از کار تفسیر نیاز دارند.

شبکه‌های عصبی یک انتخاب خوب برای تکنیک‌های وارونه‌سازی کلاسیک به نظر می‌رسد. ما سعی می‌کنیم که بررسی کنیم چطور شبکه‌های عصبی می‌توانند برای پذیرش لرزه‌نگارها به عنوان ورودی و برای محاسبه مدل زمینی متناظر به عنوان خروجی آموزش یابند. در طول فرایند یادگیری الگوی ورودی و خروجی (به ترتیب لرزه نگارها و مدل‌های زمین) به طور همزمان به شبکه

نشان داده می‌شوند. شبکه وزنه‌های را در اتصالاتش برای مرتبط ساختن مدل زمینی مطلوب به لرزه‌نگار داده شده، اصلاح می‌کند. یادگیری به این صورت انجام می‌گیرد که شبکه خروجی مطلوب را برای هر لرزه‌نگار از مجموعه آموزشی را محاسبه می‌کند، که این معنی را می‌دهد که آن یک مجموعه از وزنه‌هایی را پیدا می‌کند که هر الگوی ورودی را به مدل زمینی مربوطه نگاشت می‌کند. حال شبکه تنها نمونه‌های «یادگیری شده» را به طور صحیح تشخیص نمی‌دهد بلکه همچنین به طور آنی مدل‌های زمینی را برای لرزه‌نگارهای ناشناخته درون‌یابی خواهد کرد (e.g. داده‌های غیر از مجموعه آموزشی). به هر حال، فرایند یادگیری ممکن است به یک زمان محاسبه‌ای مهم و کامپیوترهای قوی نیاز داشته باشد.

۵-۴-۳- شرح شبکه عصبی مان

ما در شبیه‌سای مان از شبکه‌های پیشخور چند لایه، با اتصالات جزئی استفاده می‌کنیم. ساختار و تابع خروجی در طول فرایند آموزشی تغییر نمی‌کنند. بنابراین فرایند یادگیری شامل می‌شود در میزان کردن مجموعه کاملی از وزنه‌ها تنها پارامترهای آزاد از شبکه وقتی شبکه یکبار آموزش می‌یابد، ما یک مجموعه از وزنه‌های w را پیدا می‌کنیم که خطا را بین خروجی O^α محاسبه شده و خروجی مطلوب O_{des}^α را برای همه نمونه‌های α کمینه می‌کند. ما برای کمینه‌سازی ضابطه حداقل مربعات را انتخاب کردیم.

$$M(W) = \sum_{\alpha} \sum_I (O^{\alpha I} - O_{des}^{\alpha I})^2 \quad (1-5)$$

که شاخص α روی تعداد نمونه‌ها اجرا می‌شود، I روی تعداد نمونه‌های خروجی (i.e. روی تعداد نرونها در لایه خروجی)، در این روش ما می‌توانیم مساله مان را به عنوان یک مساله بهینه‌سازی در جهت حداقل مربعات فرمول‌بندی کنیم و تکنیکهای بهینه‌سازی متداول می‌تواند استفاده شوند، برای مثال کاهش شیب، مونت کارلو غیره الگوریتم پس انتشار بسط یافته برای یک الگوریتم شیب، مشتقات جزئی تابع خطا (معادله ۵-۱)، را نسبت به وزنه‌ها محاسبه می‌کند.

متأسفانه رفتار کلی شبکه‌های عصبی هنوز کاملاً شناخته نشده است و تعدادی از پارامترها برای این منظور خاص مجبور هستند که ثابت و تنظیم شوند. بنابراین استفاده از الگوریتم پس

انتشار برای وارونه‌سازی داده‌های لرزه‌ای یک آموزش موفق را تضمین نمی‌کند (i.e.) کمینه‌سازی معادله (۱-۵).

ما تصمیم گرفتیم از تکنیک مونت کارلو برای آموزش شبکه استفاده کنیم. در مرحله اول همه وزن‌ها ارزش صفر داده شدند. سپس وزن‌ها یکی یکی تغییر می‌یابند و بعد از هر تغییر ما مقدار خطای جدید داده شده توسط معادله (۱-۵) را دوباره محاسبه می‌کنیم. اگر این مقدار کوچکتر شود ما تغییر را می‌پذیریم، و گرنه آن رد می‌شود و وزن بدون تغییر باقی می‌ماند. اشکال این الگوریتم این است که آن گران است و آن همگرایی در جهت مینیمم مطلق را تضمین نمی‌کند.

۵-۴-۴- مثالهای عددی

۵-۴-۴-۱- وارونه‌سازی زمان‌های سیر

یک مدل زمینی را با یک سطح منعکس کننده در ته آن را فرض کنید. از موقعیت‌های نسبی داده شده از ۲۴ دریافت کننده و موقعیت منبع، ما می‌توانیم زمانهای سیر را برای هر دریافت کننده برای عمقهای متفاوت منعکس کننده، و سرعتهای مختلف لایه (شکل ۵-۹) محاسبه کنیم. ما ۲۰۰ مدل مختلف را توسط تغییر دادن سرعت در فاصله $2/3 - 3/8$ m/sec k و عمق منعکس کننده در فاصله از $0/4$ km، تولید کردیم. چنانچه ما سرعتها را به یک مقدار ماکزیمم ۱ هنجار کردیم، سپس فاصله از $0/6 - 1$ km/s شد. ما سپس برای همه مدلها زمانهای سیر متناظر برای هر ۲۴ دریافت کننده را محاسبه کردیم (شکل ۵-۱۰).

به شبکه‌ها ۲۴ زمان سیر به عنوان ورودی شان برای محاسبه سرعت لایه متناظر و عمق منعکس کننده داده شد. بنابراین شبکه شامل ۲۴ نرون در لایه ورودی است، و لایه خروجی از ۲ نرون تشکیل شد. یک لایه میانی با ۱۰ نرون به طور دلخواه انتخاب شده بود.

با استفاده از روش مونت کارلو شرح داده شده در بالا برای آموزش شبکه‌مان، مقدار خطای

$M(W)$ به صورت زیر محاسبه شد:

$$M(W) = \sum_{\alpha} (vel_{comp}^{\alpha} - vel_{des}^{\alpha})^2 + (depth_{comp}^{\alpha} - depth_{des}^{\alpha})^2 \quad (2-5)$$

که α روی تعداد نمونه‌ها اجرا می‌شود و شاخص‌های d es و $comp$ به ترتیب نشان دهنده مقادیر مطلوب و محاسبه شده است.

انحرافات برای وزنها به طور تصادفی در فاصله $[-1, +1]$ انتخاب شدند.

در یک اجرای اولیه ما شبکه‌مان را برای یادگیری همه ۲۰۰ نمونه آموزش دادیم. (شکل ۵-۱۱) بعد از چندین ساعت در یک CI، الگوریتم به یک مقدار خطای $10^{-2} \times 1/19$ همگرا شد. فرایند یادگیری بجز برای یک انحراف اندک برای مقادیر سرعت بالاتر که احتمالاً مربوط به تابع خروجی زیگموئید است، موفق بود.

در اجرای دوم ما شبکه‌مان را تنها روی بر پایه ۲۰ نمونه آموزش دادیم (شکل ۵-۱۲) خطا بعد از ۵۰ تکرار به $M = 6/83 \times 10^{-4}$ در چند دقیقه، تقلیل یافت.

اکنون دیدن تعمیم توانایی‌ها برای نمونه‌های ورودی، ناشناخته جالب بود (i.e.) که چطور خروجی برای زمان‌های سیر، که خارج مجموعه آموزشی بودند، محاسبه خواهد شد.

در شکل (۵-۱۳) ما خروجی محاسبه شده برای همه ۲۰۰ زمان سیر را رسم کردیم. مقادیر سرعت بجز برای یک انحراف برای مقادیر بالا، تقریباً خوب منطبق شد، در حالی که مقادیر عمق محاسبه شده کاملاً منطبق شدند ظاهراً درونیابی برای مقادیر عمق خیلی سریعتر از آن برای سرعتها است.

۵-۴-۲- مرتب‌سازی یک لرزه‌نگار به یک پروفیل عمق - سرعت

در مثالهای بعدی ما شبکه‌مان را برای پذیرش یک رد لرزه‌ای به عنوان ورودی و برای محاسبه یک پروفیل عمق سرعت پیوسته به عنوان خروجی، آموزش دادیم.

ما ۱۴ مدل تولید کردیم؛ i.e. لرزه‌نگارها و پروفیل‌های عمق سرعت، برای فرایند یادگیری. سه لایه با عمق‌های لایه‌ای ثابت در ۲۰۰ و ۴۰۰ متر برای همه نمونه‌ها، اما با متغیر بودن سرعت لایه‌ای بای هر نمونه مدل ساخته شده، برای محاسبه لرزه‌نگارهای مصنوعی سرعت لایه اول 1200 m/s با یک تغییر تصادفی $100 \text{ m/s} +$ است. همه سرعتهای دیگر تصادفی هستند، اما افزایش میانگین سرعت روی همه مدلها 450 m/s است. مدل توسط ۶۲ نقطه در عمق و لرزه‌نگار توسط ۲۴۹ نرون و لایه خروجی شامل ۶۲ نرون است. یک لایه میانی با ۵ نرون به دلخواه انتخاب شده بود.

بعد از ۱۰۰۰۰ تکرار (i.e.) یک تکرار پایان می‌یابد وقتی که همه وزنها یک بار تغییر می‌کنند (مقدار خطا به $0/2$ کاهش یافت (شکل ۵-۱۵)).

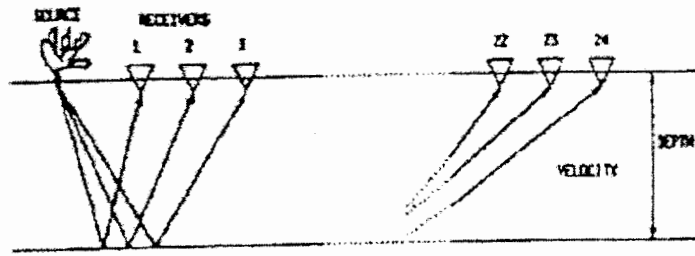
شکل (۵-۱۵) برخی پروفیل‌های زمین و همگرایی آن‌ها را در مقابل مدل مطلوب نشان می‌دهد. شبکه فورا عمق‌های را یاد می‌گیرد، به طوری که آن‌ها برای همه مدل‌ها ثابت هستند. سرعت لایه اول و دوم خیلی مختصرا منطبق شدند. این به دامنه مهم سیگنال اول در لرزه‌نگار مربوط می‌شود که شامل اطلاعات مهمی درباره سرعت لایه اول و دوم است.

آموزش برای تشخیص سرعت لایه بعدی عمدتا تکرارهای بیشتری را لازم دارد. به طور بدیهی برای شبکه بازیافت مقدار مطلوب با یک دقت بالا مشکل است. دلیل برای این امر ممکن است این باشد که دامنه سیگنال مربوط در اغلب موارد در مقایسه با دامنه اولیه کوچک است. بنابراین شبکه ما برای استخراج اطلاعات کافی برای دادن مقدار سرعت صحیح برای لایه بعدی مشکلاتی دارد.

شبکه یک مقاومت ویژه نويز خیلی خوب را نشان می‌دهد ما به شبکه لرزه‌نگارهای یکسان را در طول فرایند یادگیری، ارائه می‌دهیم، اما ما در آن ۱۰٪ نويز قرار می‌دهیم. خروجی محاسبه شده از شبکه به خروجی مطلوب برای سرعت لایه اول و دوم خوب همبسته شده است. سرعت لایه سوم در اغلب موارد اشتباه است و بنابراین هیچ اطلاعاتی نمی‌تواند از لرزه نگارها بازیافت کند. (شکل ۵-۱۶)

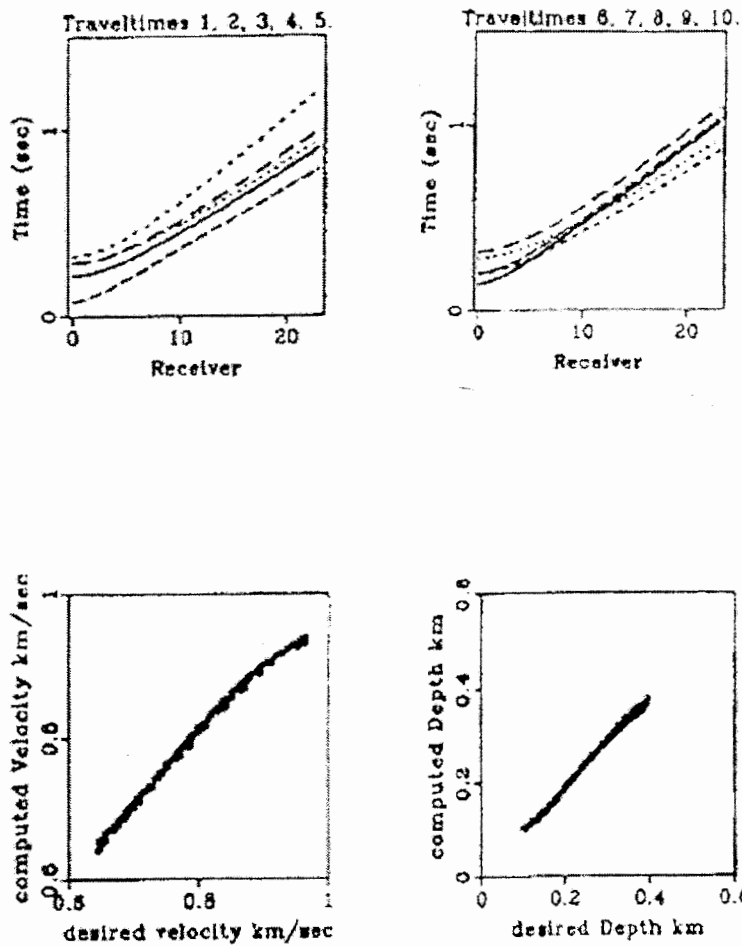
۵-۴-۵- نتایج

ما نشان دادیم که شبکه‌های عصبی می‌توانند برای وارونه‌سازی داده‌های لرزه‌ای ساده مورد استفاده قرار گیرند و اینکه یک الگوریتم مونت کارلو برای فرایند آموزشی مناسب است. شبکه‌های عصبی اغلب استفاده شده برای مسائل دسته‌بندی، می‌توانند روی مقادیر خروجی پیوسته حتی با یک تعداد بزرگی از نمونه‌ها آموزش یابند.



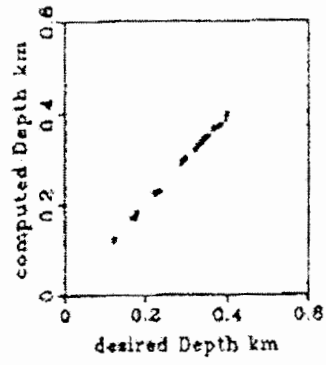
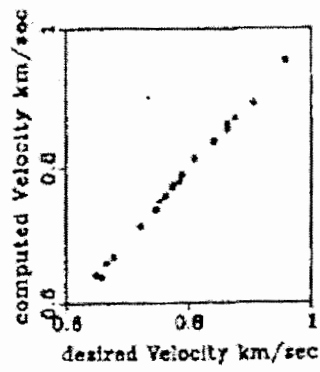
شکل ۵-۹: مدل ساده زمین، با یک منعکس کننده مسطح در کف لایه اول. موقعیت گیرنده‌ها و منبع نشان داده شده‌اند.

شکل ۵-۱۰: ۱۰ متحنی زمان سیر داده شده به شبکه به عنوان الگوهای ورودی

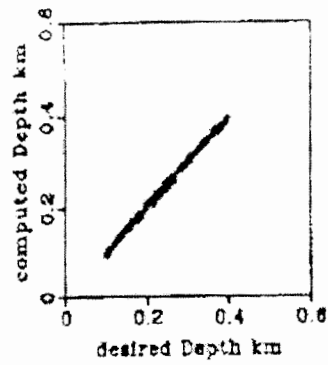
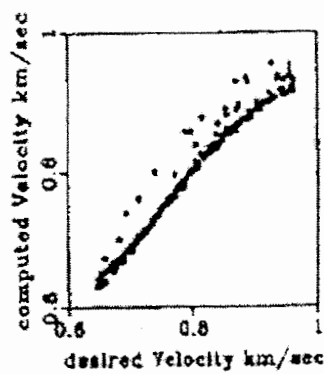


شکل ۵-۱۱: نمودار سرعت مطلوب در برابر مقدار سرعت محاسبه شده (چپ)، و عمق مطلوب در برابر مقدار عمق محاسبه شده (راست)

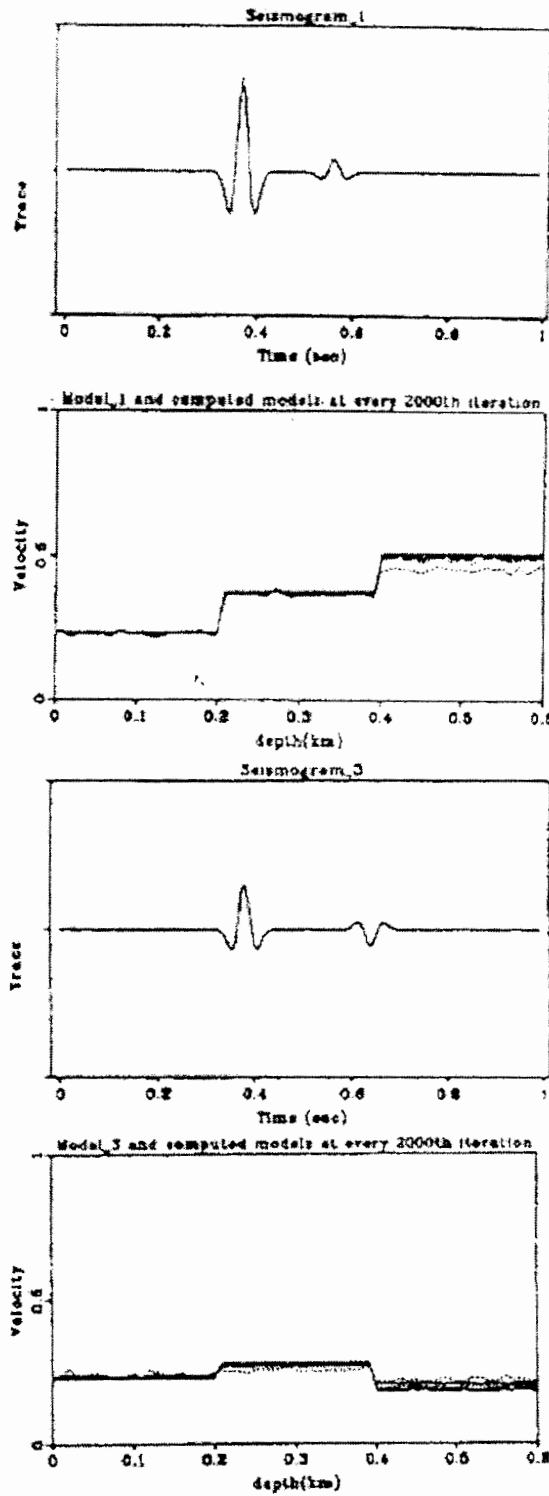
برای شبکه آموزش دیده‌مان با ۲۰۰ نمونه.



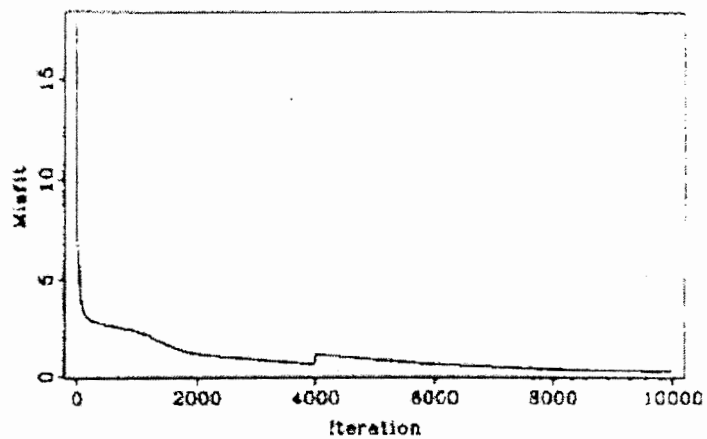
شکل ۵-۱۲: نمودار سرعت مطلوب در برابر مقدار سرعت محاسبه شده (چپ)، و عمق مطلوب در برابر مقدار عمق محاسبه شده (راست) برای شبکه آموزش دیده‌مان با ۲۰ نمونه.



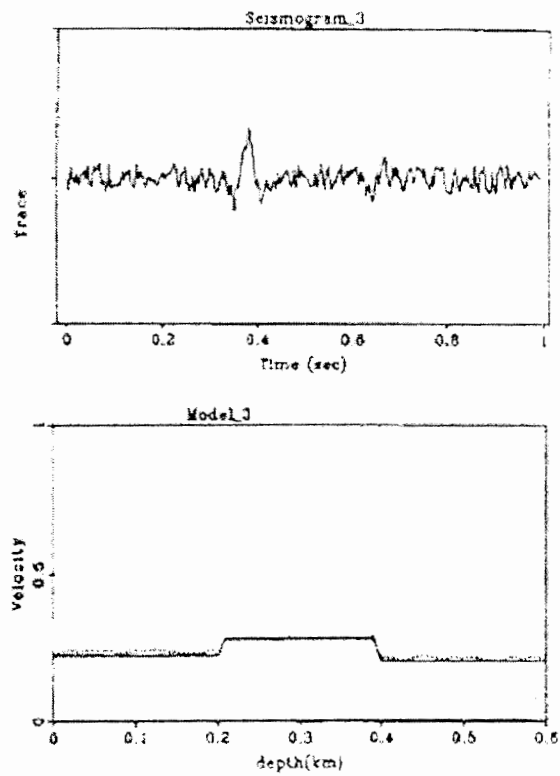
شکل ۵-۱۳: نمودار سرعت مطلوب در برابر مقدار سرعت محاسبه شده (چپ)، و عمق مطلوب در برابر مقدار عمق محاسبه شده (راست) برای همه ۲۰۰ نمونه. برای شبکه آموزش دیده‌مان با ۲۰ نمونه، برخی مقادیر محاسبه شده صحیح نیستند و انحراف برای مقادیر سرعت بالا خیلی مهم می‌شود.



شکل ۳. ۱۴. ۲ از ۱۴ نمونه مدل (پایین) و لرزه‌نگارها (بالا) استفاده شده برای آموزش شبکه‌عصبی مان. مدل مطلوب و واقعی توسط یک خط بر نشان داده شده است. اضافه بر خطوط شکسته با نقطه چین، بترتیب ۵ مدل محاسبه شده بعد از هر ۲۰۰۰ تکرار وجود دارند.



شکل ۵-۱۵: کاهش تابع عدم تطبیق برای شبکه آموزش یافته با لرزه نگارها. مقدار عدم تطبیق نهایی حدود ۰/۲ است.



شکل ۵-۱۶: تصویر مشابه (۵-۱۵)، اما اینجا لرزه نگارها را با ۱۰٪ نویز (بالا) و مدل زمینی مطلوب متناظر (پایین) را توسط یک خط پر، اضافه بر این مدل محاسبه شده را توسط یک خط نقطه چین نشان می دهیم.

فصل ششم

کاربرد شبکه‌های عصبی در ژئومغناطیس و گرانی‌سنجی

۱-۶- تاریخچه

پیرسن و دیگران (pearson et al. 1990) از داده‌های مغناطیس هوایی با قدرت تفکیک پذیری بالا برای دسته‌بندی آنومالی‌ها به عنوان بالای پی‌سنگ (uprabasement) و داخل پی‌سنگ (intrabasement) در حوضه رود دنور- جولدسبرگ شمالی (Denver-julesberg) استفاده کرد. برخی مخازن Permo-pennsylvanian در ساختارهای بلندی‌های Paleotopographic ذخیره شده‌اند که در ارتباط با بلندی‌های پی‌سنگ هستند. بلندی سنگ بستر آنومالی مغناطیسی ظریف تولید می‌کند که می‌تواند در پروفیل‌ها توسط یک مفسر تشخیص داده شود. یک MLP با یادگیری پس انتشار ۱۰ ورودی مربوط به داده‌های مغناطیسی و تبدیلات گوناگون داده شده بودند، مانند گرادیانهای عمودی و افقی. شبکه از دو خروجی عنصر پردازشی (PE) برای دسته‌بندی علائم به عنوان خارج پی‌سنگ و داخل پی‌سنگ استفاده کرد. مجموعه آموزشی از هر دو داده‌های صحرایی و مدل‌های ترکیبی برای تهیه یک نوع از آنومالی‌ها استفاده کرد. سپس شبکه روی داده‌های صحرایی و داده‌های ترکیبی بیشتری آزمایش شد. آنومالی‌های تشخیص داده شده توسط شبکه برای تأیید با داده‌های لرزه‌ای و چاه‌پیمایی مقایسه شدند. شبکه ۸۰٪ از آنومالی‌های ساختاری در داده‌های صحرایی و ۹۵٪ از ساختارها در داده‌های ترکیبی را موقعیت‌یابی کرد.

ژائو و دیگران (Guo et al. 1992) و کارتابیا و دیگران (Cartabia et al. 1994) راههای متفاوتی از استخراج اطلاعات خطی از داده‌های مغناطیسی ارائه دادند. ژائو و دیگران (۱۹۹۲) برای دسته‌بندی داده‌ها به هشت جهت قطبنمایی (NS, NE, NW و غیره) احتیاج داشتند. یک شبکه پس‌انتشار جدا برای هر جهت کمپاس درست شد. شبکه‌ها با پنجره‌های مدل 7×7 پیکسل آموزش یافتند. داده‌های صحرائی سپس به شبکه‌ها در پنجره‌های مؤثر 7×7 وارد شدند و شبکه با بزرگترین خروجی برای آن پنجره، جهت فرض شده بود.

کارتابیا و دیگران (۱۹۹۴) از یک ساختار ماشین بولتزمن (Boltzmann machine) استفاده کردند، شبیه به روش ذوب شبیه‌سازی شده سریع ارائه شده توسط سن و استوفا (Sen and Stoffa 1995)، برای پیش‌بینی پیوستگی به پیکسل‌های مشخص شده توسط یک الگوریتم آشکارسازی لبه‌ای با استفاده از داده‌های گرانی‌سنجی. الگوریتم آشکارسازی لبه‌ای، متصل‌شدگی یا نازک‌شدگی پیکسل‌های لبه‌ای را پیش‌بینی نمی‌کند که برای نگاشت خطوط لازم است. با به کار بردن یک شبکه بهینه، مانند ماشین بولتزمن برای پیکسل‌های لبه‌ای، یک نگاشت خطی می‌تواند به طور خودکار تولید شده باشد که با نگاشتی که توسط یک مفسر ماهر تولید شده مطابقت کند.

تیلور و واسکو (Taylor and Vasco 1990) داده‌های گرادیومتری گرانی‌سنجی را با یک الگوریتم یادگیری پس‌انتشار وارونه کردند. مدل‌های ترکیبی از یک سنگ بستر با چگالی بالا و اندکی یک ته‌نشست سطحی با چگالی کمتر تولید شده بودند. مدل‌ها به ۱۸ سلول تقسیم شده بودند و شبکه برای تخمین عمق به فاصله میانی در هر سلول مورد نیاز بود. عمق میانگین برای فاصله میانی ۱ کیلومتر بود. مجموعه آموزشی بوسیله انتخاب تصادفی عمقها برای فاصله میانی و محاسبه شیب گرانی برای مدل تصادفی ایجاد شده بود. از شبکه انتظار می‌رفت که عمق داده شده از داده‌های شیب را تخمین بزنند. شبکه بر روی یک مدل ترکیبی جدید که شامل یک مرزبندی با جهت شمالی - جنوبی قرار گرفته بر روی بستر افقی در عمق ۱۰ کیلومتری بود آزمایش شد. شبکه قادر به تولید دوباره به اندازه کافی مدل آزمایشی تنها با خطاهای کوچک در عمق‌ها در هر محل سلول بود.

سالم و دیگران (Salem et al. 2000) یک سیستم شناسایی شبکه عصبی دقیق و سریع برای آشکارسازی ساختارهای استوانه‌ای شکل فلزی مدفون با داده‌های مغناطیسی را توسعه داد. قرائت داده‌ها از ۲۱ ایستگاه در هر ۱ متر در طول یک پروفیل به عنوان ورودی استفاده شدند.

خروجی شامل دو PES بود که عمق و فاصله افقی در طول پروفیل را برای یک شیء دفن شده را تخمین می‌زند. برای شبیه‌سازی اثر یک ساختار استوانه‌ای شکل فلزی، محاسبات مدل مستقیم انجام شده بودند که بر پایه یک منبع دو قطبی هم ارز بود. ساختار استوانه‌ای شکل در عمق‌های از ۲ متر تا ۶ متر در موقعیت‌های گوناگون در طول پروفیل مدلسازی شد. یک مجموعه کامل از ۷۵ پاسخهای مدل برای مجموعه آموزشی محاسبه شد. نویز به داده‌ها اضافه شد بوسیله شبیه‌سازی یک لحظه مغناطیسی قرار گرفته در ۱۰ متری کنار خط پروفیل در یک عمق ۲.۱ متری. محدوده نویزی از ۱۰٪ تا ۴۰٪ به داده‌ها اضافه شد. ارزیابی‌های شبکه از موقعیت ساختار استوانه‌ای شکل با نویزی بالای ۲۰٪ قابل قبول بود. داده‌های از ۱۰ پروفیل در ژئومتریک‌های EG&G ایستگاه آزمایشی دانشگاه استنفورد برای آزمایش شبکه استاده شد. به طور میانگین، عمق‌های این ساختارها با ۵٪ متر ارزیابی شدند. ارزیابی‌های موقعیت‌های کناری از دقت کمتری برخوردار بودند اما در بیشتر موارد درون یک بعد استوانه‌ای از موقعیت صحیح بودند (استوانه‌ها با ضخامت ۰/۵۹ متر و ارتفاع ۰/۹۸ بودند). [۲۷]

۲-۶- تشخیص ساختار از میدانهای پتانسیلی با استفاده از شبکه‌های

عصبی

۲-۶-۱- خلاصه

تعبیر و تفسیر میدانهای پتانسیل بینش مهمی به تحلیل ساختمانی و نگاشت ساختار شناسی و سنگ شناسی می‌دهد. چنین تفسیری کار ذهنی و وقت گیر است. در این جا یک روش جدید را شرح می‌دهیم که می‌تواند برای استخراج مشخصه‌های ساختاری و اطلاعات ساختارشناسی از میدانهای پتانسیل به طور خودکار و سریع و با استفاده از شبکه‌های عصبی پس انتشار، استفاده شود. نتایج آزمایشات با داده‌های ترکیبی و حقیقی نشان می‌دهد که اگر ما بتوانیم پارامترهای مناسب برای شبکه‌های عصبی انتخاب کنیم و نمونه‌های مدل معقول ایجاد کنیم، ما نقشه‌های رده‌بندی ساختارشناسی و مشخصه بدست خواهیم آورد که برای تحلیل ساختاری و نقشه‌های زمین‌شناسی مفید هستند. این تکنیک به طور موفقی برای تفسیر داده‌های مغناطیس هوایی به کار برده می‌شود. بعلاوه برای مشخصه‌های شناخته شده که مشابه ساختارهای شناخته شده در نقشه

زمین‌شناسی است، برخی خصیصه‌ها تشخیص داده شده بودند که اعتقاد بر این داشتند که مربوط به گسل‌های دفن شده می‌شدند. نتایج معقول رده‌بندی ساختارشناسی همچنین بدست آمده بودند. [۳۷]

۶-۲-۲- مقدمه

استخراج الگوها و خصوصیات از میدانهای پتانسیلی با استفاده از روشهای دسته‌بندی آماری مشکل است، هر چند شبکه‌های عصبی پتانسیل بزرگی برای تشخیص نمونه نشان می‌دهد. یکی از فواید شبکه‌های عصبی روی روشهای رده‌بندی آماری این است که شبکه‌های عصبی توزیع آزاد هستند، بنابراین هیچ دانش قبلی درباره توزیع‌های آماری دسته‌ها لازم نمی‌شود. مزیت دیگر روش شبکه عصبی این است که می‌تواند اطلاعاتی درباره مفهوم مؤلفه‌های بردار خصیصه ورودی تهیه کند، نظر به اینکه دسته‌بندی‌های غیر پارامتری اطلاعاتی درباره این نوع حال نمی‌دهد.

مقالات کمی درباره کاربرد شبکه‌های عصبی برای میدانهای پتانسیل وجود دارد. بیشترین دلیل اهمیت آن این است که تشخیص خصوصیات از میدانهای پتانسیل با استفاده از شبکه‌های عصبی خیلی مشکل‌تر از داده‌های دریافتی پرت یا دیگر داده‌های زمین‌شناسی است. یکی از اهداف این تحقیق توسعه تکنیکی است که بتواند برای استخراج خودکار خصیصه‌ها (تشخیص مشخصه و دسته‌بندی ساختاری) از میدانهای پتانسیلی با استفاده از شبکه‌های عصبی پس انتشار استفاده شود. تنها داده‌های مغناطیسی در این تحقیق استفاده شوند، اما روش مشابه می‌تواند برای داده‌های گرانی نیز به کار برده شود. [۳۷]

۶-۲-۳- تشخیص مشخصه

تحلیل ساختاری، یک بخش مهم از تعبیر و تفسیر میدانهای پتانسیل است. هر چند، شکل‌های ساختاری در میدانهای پتانسیل خیلی پیچیده هستند که ما نمی‌توانیم نمونه‌های مدل را برای ارائه رفتارهایشان به طور دقیق برای همه نمونه‌های ساختارها پیدا کنیم. خوشبختانه، شکل‌های ساختاری اصلی میدان پتانسیل شامل مشخصه‌هایی است که می‌تواند برای آنومالی‌های دو بعدی در

نظر گرفته شده باشد؛ دیگر شکلها می‌تواند برای ساخته شدن از یک تعداد مشخصه در نظر گرفته شود.

لزوم این که مقادیر ورودی BP از صفر تا یک تغییر کند ما را به ساختن نمونه‌های مدل در بیشترین روش ساده و خواهد داشت، برای مثال، نمونه‌های مدل شامل تنها ۰ و ۱ می‌شوند. اندازه‌های نمونه‌های مدل بر پایه ویژگی عمومی آنومالی‌های خطی از میدانهای پتانسیل تعیین شده بودند. سرانجام، ۸ نمونه مدل در اندازه 7×7 ساخته شدند، هر یک با داشتن یک مشخصه ساده به ترتیب در طول جهت. (شکل ۶-۱). شبکه‌های عصبی برای تشخیص مشخصه، با استفاده از این نمونه‌های مدل آموزش یافتند. تعداد ندهای خروجی به ترتیب 49 و 1 بودند. توسط روش سعی و خطا، ما یافتیم که می‌توانیم نتایج بهتری بدست آوریم اگر ثابت تابع زیگموئید، برابر با ۰/۵ قرار داده شود و تعداد ندها در لایه میانی ۱۵ باشد.

شکل (۶-۲) نمودار قواعد کلی تشخیص مشخه با استفاده از BP است. در این شکل، شبکه آموزش یافته یا وزنهای آموزش یافته ۸ نمونه مشخصه مدل با مربع‌ها نمایش داده شده‌اند. وقتی که داده‌های حقیقی برای تشخیص مشخصه پردازش شدند، آن‌ها به طور متوالی برای تطبیق مجموعه داده‌های حقیقی در یک پنجره متحرک فراخوانده می‌شوند (پنجره متحرک). به عبارت دیگر، داده‌های حقیقی در پنجره به ترتیب به ۸ شبکه آموزش یافته داده شدند، سپس ۸ خروجی واقعی شبکه به دست آمدند. فرض کنید آن جا یک مشخصه با حالت NNE در پنجره وجود دارد؛ سپس پاسخ شبکه از نمونه مدل NEE بایستی ماکزیمم باشد. اگر ما این ماکزیمم را انتخاب کنیم، ما می‌توانیم اطلاعاتی درباره جهت مشخصه در پنجره بدست آوریم. عموماً، مشخصه در مجموعه داده‌های حقیقی به یک نمونه مدل نزدیکتر است اگر مقدار تطبیقی بیشتر باشد، ماکزیمم بهترین تطبیق را بین مشخصه‌ها در داده‌های حقیقی و در یک نمونه مدل نشان می‌دهد. سرانجام وقتی که پنجره از میان کل مجموعه داده‌ها حرکت داده شد، ما یک نقشه مشخصه بدست آوردیم. [۳۷]

۴-۲-۶ - رده‌بندی ساختار شناسی

خوب است بدانید که برخی رفتارهای مغناطیسی مجزا برای انواع مختلف سنگها به سبب تفاوت‌های قابلیت نفوذپذیری و دیگر صفات وجود دارند که امکان رده‌بندی ساختارشناسی با استفاده

از شبکه‌های عصبی را فراهم می‌آورد. البته، دسته‌بندی کردن سنگ‌شناسی با استفاده از تنها داده‌های مغناطیسی مشکل است زیرا میدان مغناطیسی تنها توسط اجسام زمین‌شناسی مغناطیسی نتیجه می‌دهد. در حقیقت، رده‌بندی ساختارشناسی از میدانهای مغناطیسی با استفاده از شبکه‌های عصبی در حقیقت یک رده‌بندی از آنومالی‌های مغناطیسی است.

ایده مبنایی دسته‌بندی ساختارشناسی با استفاده از شبکه‌های عصبی، آموزش $B-P$ با برخی نمونه‌های مدل است که در رابطه با ساختارهای مختلف هستند، برای مثال، آموزش شبکه‌هایی برای یادگیری تشخیص دادن ویژگیهای آنومالی‌هایی که ما می‌خواهیم. این شبکه‌های آموزش یافته می‌توانند سپس برای دسته‌بندی کل مجموعه داده‌ها استفاده شود برای دسته‌بندی ساختاری با استفاده از شبکه‌های عصبی، اختلاف اصلی از تشخیص مشخصه این است که ما نمی‌توانیم نمونه‌های مدل ترکیب مشابه با همه وضعیت‌ها ایجاد کنیم. بنابراین، نمونه‌های مدل بایستی از داده‌های حقیقی اقتباس شوند.

شکل (۳-۶) یک نمودار از قواعد کلی رده‌بندی ساختاری با استفاده از $B-P$ است. عموماً، ما بایستی ابتدا نمونه‌های مدل را از داده‌های حقیقی برداشت کنیم، سپس $B-P$ را با هر نمونه مدل آموزش دهیم و وزنهای شبکه‌های عصبی آموزش یافته را بدست آوریم. در آخر، ما $B-P$ آموزش یافته را در یک برای تحلیل مجموعه داده‌های دست نخورده فرا می‌خوانیم. اندازه نمونه مدل 3×3 انتخاب شده است زیرا ما امید داریم برای تشخیص یک رده در ناحیه واحد مینیمم. به عبارت دیگر، ما امید داریم برای رده‌بندی میدان مغناطیسی در دقیق‌ترین حالت ممکن. بایستی توجه شود که دقت رده‌بندی مستقیماً بستگی به دقت نمونه‌های مدل دارد. بنابراین، ما بایستی نمونه‌های مدل را به دقت، هنگام برداشت آن‌ها از داده‌های حقیقی، انتخاب کنیم. قواعد عملی اینها هستند که اندازه داده‌های نمونه‌های مدل بایستی در بزرگترین حالت ممکن باشد و بایستی شامل تنها یک نوع از ساختار باشد. معمولاً، تصویر برجسته سایه‌دار میدان مغناطیسی برای انتخاب نمونه‌های مدل از داده‌های حقیقی مفید است. [۳۷]

۶-۲-۵- تشخیص خصیصه از میدان مغناطیسی هوابرد

برای ارزیابی قابلیت این تکنیک برای تشخیص مشخصه و ساختارشناسی یا رده بندی متنی، داده‌های مغناطیس هوابرد، برای تشخیص مشخصه و رده بندی ساختاری پردازش شد. دلیل برای انتخاب این ناحیه این بود که داده‌های مغناطیس هواپردی با کیفیت بالا فراهم بود و زمین شناسی این منطقه خوب شناخته شده بود.

شکل (۴-۶) نقشه زمین شناسی (Pueblo Quadrangle, Colorado) را نشان می‌دهد. منطقه کاوش مستطیلی با پهناهای ۲۰۰ km در ۱۲۰ km است، که در زیرنویس نقشه در گوشه پایین سمت راست شکل (۴-۶) نمایش داده شده است. کوههای ناحیه به طور عمده از سنگهای متبلور پراکندگی کامبرین تشکیل یافته، و تپه‌های دامنه کوه و دشت‌ها اغلب توسط سنگهای رسوبی پانروزویک لایه بندی شدند. محل اتصال بین آن‌ها یک رشته از گسل‌های معکوس با زاویه زیاد در طول Front Range از Golden تا Colorado Springs است. [۳۷]

۶-۲-۵-۱- تشخیص مشخصه

شکل (۵-۶) نقشه‌های مشخصه را بعد از تشخیص با استفاده از BBPN نشان می‌دهد. در مقایسه با نقشه زمین شناسی، ما می‌توانیم ببینیم که یک تعداد از مشخصه‌ها به خوبی با نقشه‌های زمین شناسی تطبیق دارند، برای مثال ساختار بیضی شکل اطراف و جلوی کوه، مرز بین سنگهای آتش نشانی و سنگهای گرانیتی و ناحیه گسل شمال غربی در سنگهای دگردیسی پرامبرین و دایک‌های دگردیسی در سنگ‌های رسوبی ینسیلوانیایی.

علاوه بر مشخصه‌های توضیح داده شده در بالا، برخی مشخصه‌ها وجود دارد که نمی‌توان بر روی نقشه زمین شناسی یافت. آن معتقد بر این است که این مشخصه‌ها در ارتباط با گسل‌های پوشیده است، برای مثال، یک ساختار خطی شمال شرقی ناپیوسته (۷) در گوشه پایین سمت چپ از شکل (۵-۶)، و حداقل دو مشخصه وازی (۱۷ و ۷۱۱) با جهت‌های شرقی - غربی در شمال خلیج. [۳۷]

۶-۲-۵-۲- رده بندی ساختاری

شکل (۶-۶) نتایج رده بندی را شرح می دهد. این نقشه رده بندی نظری می افکند بر برخی چیزهای نمونه شده و نه کاملاً رضایتبخش؛ برای مثال رده ۴ و رده ۵ نمی توانند از یکدیگر تمیز داده شوند زیرا ویژگیهای مغناطیسی نمونه های مدل ۴ و ۵ خیلی نزدیک به هم هستند. به هر حال، ما هنوز می توانیم به راحتی سنگهای دگرذیسی پرکامبرین، سنگهای آذرین، رسوبات دوران سوم و رسوبات کرناسه بر طبق سطح خاکستری و نمونه هایشان تمیز دهیم. [۳۷]

۶-۲-۶- نتیجه

نتایج یک تعداد از آزمایشات با استفاده از داده های مصنوعی و داده های حقیقی در این مقاله نشان می دهد که شبکه های عصبی مصنوعی می توانند برای استخراج اطلاعات به طور خودکار درباره مشخصه ها و متن های مغناطیسی از میدانهای پتانسیلی به کار برده شوند.

شبکه های عصبی برای تشخیص ساختارها از میدانهای پتانسیلی مورد استفاده قرار گرفت (عموماً، همانطور در دیگر داده های ژئوفیزیکی) منوط به:

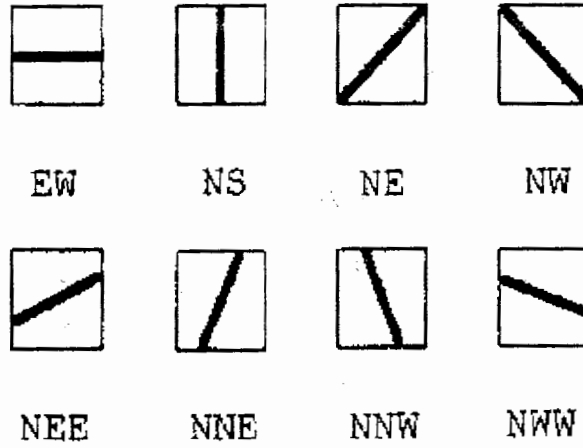
۱- انتخاب پارامترهای شبکه های عصبی. برای مثال، پارامترهای BP: تعداد ندها در لایه های میانی، یادگیری میزان و شیب تابع محرک زیگوئید.

۲- طراحی و ساخت نمونه های مدل مناسب

۳- ساخت یک ارائه قابل قبول

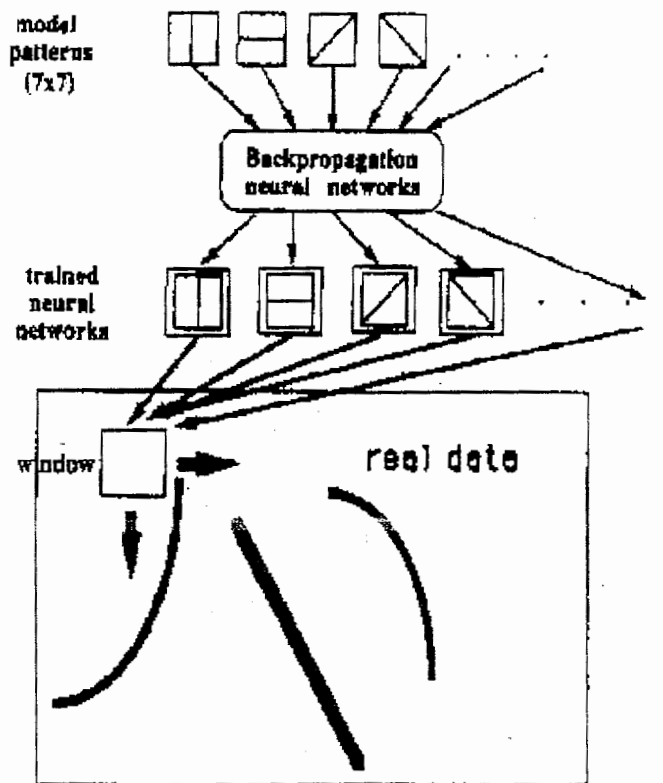
این تکنیک می تواند برای هر نوعی از داده های شبکه بندی شده به کار برده شود. تشخیص خصیصه با چندین انواع داده های ژئوفیزیکی استفاده از شبکه های عصبی می تواند مورد انتظار باشد برای به دست آوردن نتایج قابل قبول زمین شناسی. [۳۷]

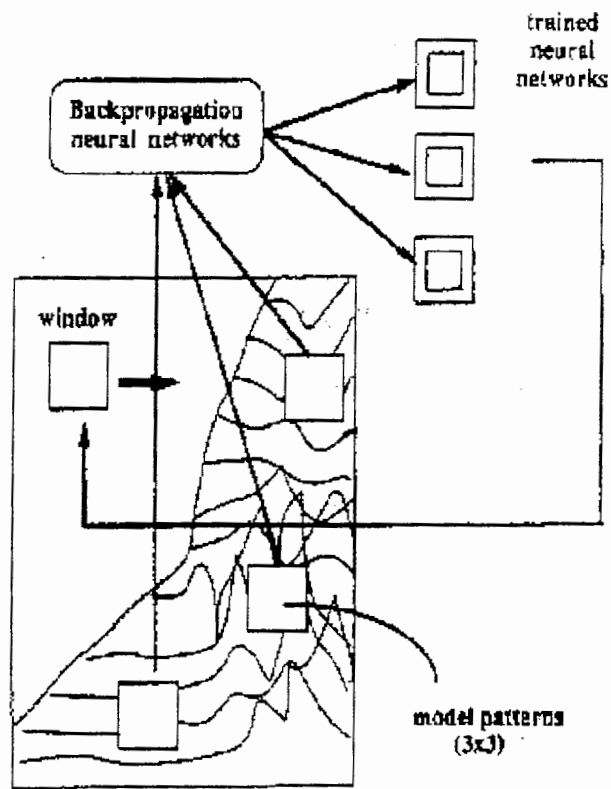
PATTERN SIZE 7X7



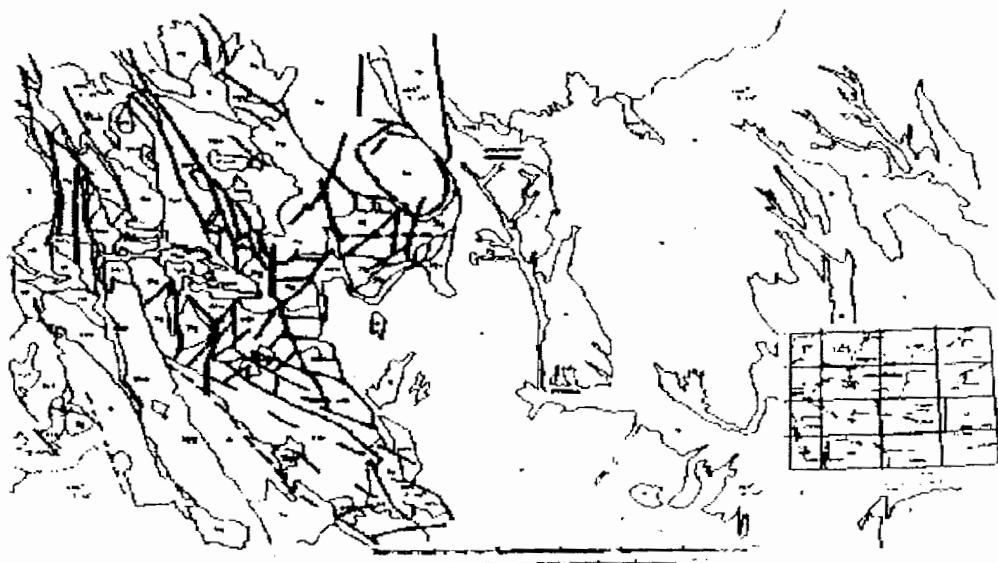
شکل ۱-۶: الگوهای مدل برای تشخیص خطوط

شکل ۲-۶: نمودار قواعد کلی تشخیص مشخصه با استفاده از BPNN

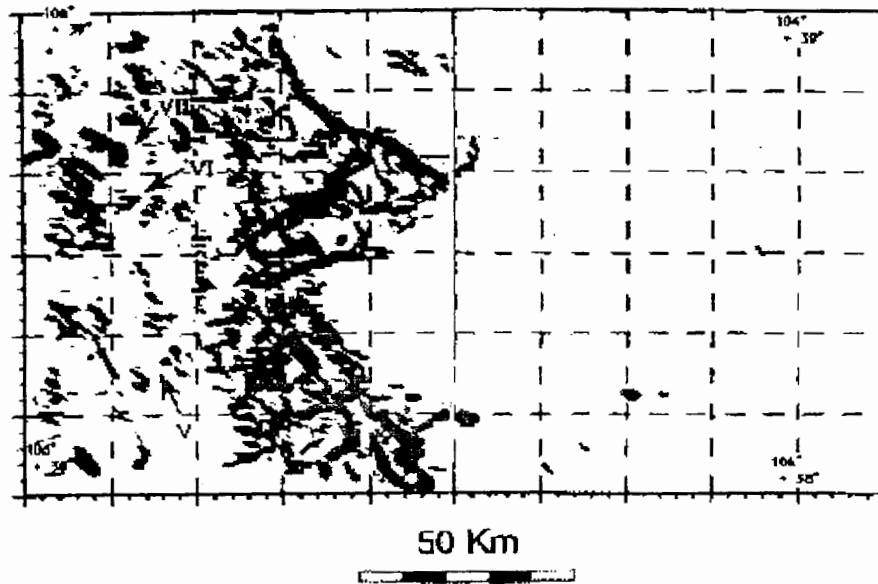




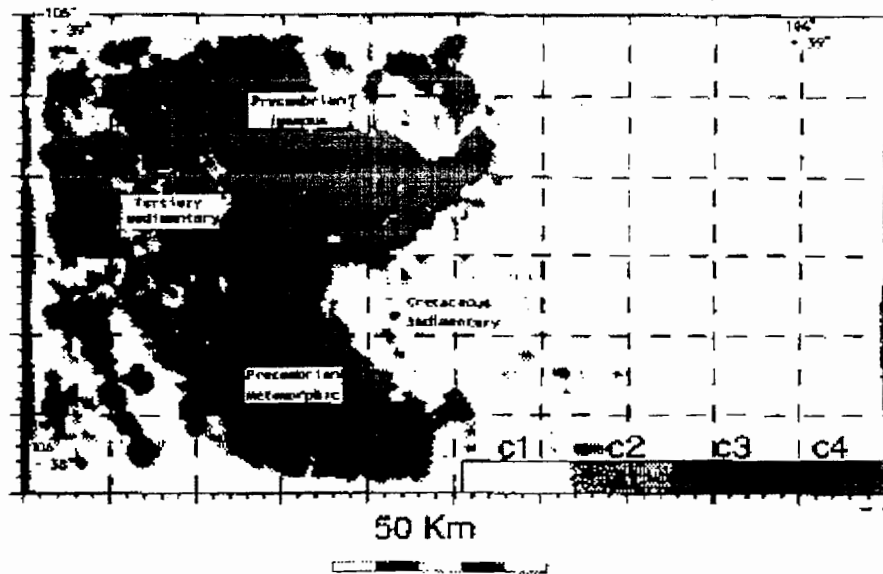
شکل ۶-۳: نمودار قواعد کلی دسته‌بندی سنگ‌شناسی



شکل ۶-۴: نقشه زمین‌شناسی Pueblo Quadrangle, Colorado



شکل ۵-۶: نقشه مشخصه مرکب از میدان مغناطیسی هوایی Pueblo Quadrangle, Colorado با استفاده از BPNN



شکل ۶-۶: نقشه دسته‌بندی سنگ‌شناسی از داده‌های مغناطیسی هوایی Pueblo Quadrangle, Colorado با استفاده از BPNN

۶-۳- تحلیل مشخصه‌ها برای داده‌های میدانهای پتانسیلی با استفاده

از شبکه‌های عصبی^۱

۶-۳-۱- خلاصه

ما یک روش جدید، با استفاده از شبکه‌های عصبی متقارن، برای تحلیل مشخصه‌ها در نگاره‌های میدانهای پتانسیل ارائه می‌دهیم.

روش دو مرحله است. اول، یک مجموعه از داده‌های محلی شامل اطلاعات لبه‌ای، با استفاده از مدل حقیقی استخراج می‌شود. این اطلاعات سپس با اطلاع از ساختار برای تحلیل شدن، کامل می‌شوند (کاراکترهای ساختاری) برای تعیین یک مجموعه سازگار از مشخصه‌ها. این توسط طرح مساله به عنوان یکی از بهینه‌سازی ترکیبی انجام می‌شود. یک تعریف از مشخصه بنا نهاده شده بر: متمرکز کردن (موضعی کردن) دقت، پیوستگی، سستی فرموله شود. تابع عمومی بوسیله تعیین اتصالات و آستانه‌های یک شبکه عصبی متقارن از واحدهای آستانه مضاعف تعریف می‌شود. شبکه، با شروع از یک حالت اولیه، به یک شکل سیستم با مقادیر بالاتر تابع عمومی استنتاج می‌کند. مسیر تکاملی شبکه می‌تواند قطعی یا احتمالی باشد.

آزمایشات روی هر دو داده‌های حقیقی و مصنوعی با نتایج خیلی خوب انجام شدند.

۶-۳-۲- آشکارسازی اولیه

یک تعداد از عملگرهای «شانگر لبه» برای تحلیل تغییرات شدت در نگاره‌های میدانهای پتانسیل پیشنهاد شد (Roest et alii, 1992; Blakely and Simpson, 1986). ما مدل هموار هارالیک را به کار گرفتیم، زیرا این مدل با داده‌های مشخص شده توسط تغییرات فضایی هموار شده مانند دانه‌های میدان پتانسیل، خوب کار می‌کند.

در مدل هموار، برای هر پیکسل نگاره یک سطح دو مکعبی، $f(r, c)$ منطبق با مجاورت پیکسل می‌شود با استفاده از:

$$f(r, c) = k_1 + k_2 r + k_3 c + k_4 r^2 + k_5 rc + k_6 c^2 + k_7 r^3 + k_8 r^2 c + k_9 rc^2 + k_{10} c^3$$

۱- مطالب این بخش از منبع [۳۶] آورده شده است.

جهت α از شیب توسط زیر داده می‌شود:

$$\alpha = \tan^{-1} \frac{k_2}{k_1}$$

در این جهت:

$$f_\alpha(\rho) = A\rho^3 + B\rho^2 + C\rho + D$$

اگر، برای برخی ρ ، $|\rho| < \rho_0$ که اندکی کوچکتر از طول ضلع یک پیکسل است،

$$f_\alpha''(\rho) = 0 \quad (1-6)$$

$$|f_\alpha'(\rho)| > \gamma \quad (2-6)$$

ما پیکسل مرکزی مجاورتی را به عنوان یک پیکسل لبه‌ای علامت می‌زنیم.

ما می‌توانیم از مساله تطبیق پارامتر γ اجتناب کنیم، به پیروی از روش دسته‌بندی پیشنهاد شده

توسط (1). در این مورد ما معادله (2-6) را توسط زیر جایگزین می‌کنیم:

$$\text{sign}(f_\alpha'''(\rho))\text{sign}(f_\alpha'(\rho)) < 0$$

مدل همواره اجازه می‌دهد به ما برای آشکارسازی لبه‌ها در مقیاس‌های مختلف با استفاده از

اندازه‌های مجاورتی مختلف (5x5 و 7x7 و ...).

مجموعه پیکسل‌های اقتباس شده، یک مدل ضعیف در روابط پیوستگی و سستی است. ما در بدست

آوردن نتایج بهتر توسط تغییر مقدار آستانه ناتوان هستیم.

اگر چه خروجی مدل برای تعیین مشخصه‌ها ناکافی است، ما از اطلاعات ρ استفاده می‌کنیم، از

مرکز پیکسل به محل تقاطع صفر مشتق جهتی دوم و جهت شیب، برای هر پیکسل نگاره.

۳-۳-۶ - شبکه‌های عصبی موازی (متقارن)

یک شبکه عصبی موازی از مدل واحدهای آستانه‌ای دوتایی یک سیستم زمانی گسسته

است که می‌تواند توسط یک گراف موزون شده و جهت دار نشده نمایش داده شود (Hopfield, 1988; Bruck and Goodman, 1982).

شبکه عصبی به طور یکتا توسط $\langle w, T \rangle$ تعریف می‌شود

که:

W - یک ماتریس متقارن $n \times n$ است، W_{ij} وزن متعلق به لبه (j, i) است.

T - یک بردار آستانه‌ای است، T_i به معنای آستانه نام است.

حالت شبکه در زمان t ، توسط حالت نرونها تعیین می‌شود، که می‌تواند ۰ یا ۱ باشد.

$$V(t) = \langle V_1(t), V_2(t), \dots, V_n(t) \rangle$$

شبکه عصبی یک سیستم پیوسته است: در هر زمان گسسته، یک مجموعه از نرونهای خودش بهنگام می‌شوند. شبکه به صورت ترتیبی عمل می‌کند، کاملاً موازی یا موازی اگر به ترتیب $|s|=1$ ، $|s|=n$ یا $1 < |s| < n$.

نرونها با یک تابع به هنگام می‌شوند که، اگر سیستم قطعی باشد، به صورت زیر است:

$$V_i(t+1) = \begin{cases} 1 & \text{if } H_i(t) > 0 \\ 0 & \text{if } H_i(t) \leq 0 \end{cases}$$

که

$$H_i(t) = \sum_{j=1}^n W_{ij} V_j(t) + T_i$$

اگر سیستم اتفاقی باشد، به هنگام شدن از توزیع احتمال بولتزمن پیروی می‌کند:

$$P[V_i(t) = 1] = \frac{1}{1 + e^{-\beta H_i(t)}}$$

که دمای β در طول تغییر شکل کاهش می‌یابد (Aarts and Korst, 1987).

یک ویژگی مهم مدل قطعی، این است که سیستم همیشه همگرا به یک حالت پایدار باشد، هنگام عمل در روش ترتیبی، و به یک دور از طول ماکزیمم ۲ وقتی که آن در روش کاملاً موازی عمل می‌کند. این ویژگی، مبنای قابل اجرا بودن مدل در بهینه‌سازی ترکیبی است. در حقیقت محقق شده است که شبکه توسط بهینه‌سازی توابع ویژه‌ای که توابع لیاپانف نامیده می‌شوند استنتاج می‌شود که هنگام عمل در روش ترتیبی به صورت زیر داده می‌شود:

$$E(t) = \frac{1}{2} \sum_i \sum_j W_{ij} V_i(t) V_j(t) + \sum_i T_i V_i(t) \quad (3-6)$$

تغییر شکل قطعی ممکن است در مینیمم موضعی به تله بیفتد، مینیمم مطلق می‌تواند در سیر تکامل اتفاقی با استفاده از یک اجرای موازی از ذوب شیشه‌سازی شده به دست آید.

۶-۳-۴- تحلیل مشخصه با استفاده از شبکه‌های عصبی

برای هر پیکسل از نگاره میدانهای پتانسیلی، مایک نرون شبکه عصبی را شرکت می‌دهیم. حالت هر نرون در شبکه یک پیکسل از نگاره را دسته بندی می‌کند. بنابراین حالت شبکه یک راه حل کاندید است. تابع اجمال (عمومی) (۶-۳)، با اختصاص دادن یک مقدار حقیقی برای هر راه حل، توسط مقادیر اتصالات و آستانه‌ها تعیین می‌شوند. این مقادیر توسط آزمایش خواص محلی مشخصه‌ها تعیین می‌گردند.

- پیوستگی

این خاصیت متلاشی سازی را توسط اتصال پیکسل‌ها در جهت عمودی به شیب کاهش می‌دهد.

- محلی کردن دقت

این خاصیت آستانه‌ای شبکه عصبی را تعریف می‌کند. این مقادیر، برای هر پیکسل، به وسیله در نظر گرفتن فواصل از مرکز پیکسل تا محل تقاطع صفر مشتق جهتی دوم در جهت شیب، تعیین می‌شوند.

- طول

بوسیله تعیین مقادیر غیر مثبت برای آستانه‌ها، طول مینیمم یک مشخصه ۲ است. در حقیقت هیچ نرون مجزا شده‌ای نمی‌تواند فعال باشد.

- سستی

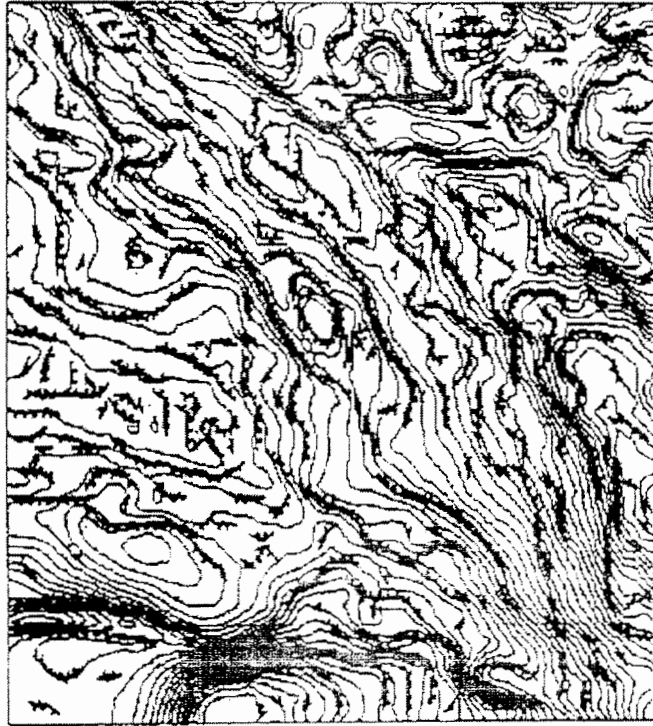
سستی یک رابطه تعریف شده بر روی نقاط سه‌گانه در مجاورت 3×3 است. به ویژه ما یک مقدار منفی برای هر نقاط سه‌گانه در پنجره 3×3 با تعیین یک ترکیب ضخیم مشخص می‌کنیم. این خاصیت، در شبکه توسط فرض یک تعمیم مدل شبکه عصبی کدبندی می‌شود. این بوسیله معرفی ارتباطها بین سه جزئی‌های نرونها انجام می‌یابد.

۶-۳-۵- نتایج آزمایش

نتایج آزمایشات روی نگاره‌های میدانهای پتانسیلی توسط مثالهای شکل (۶-۷) و (۶-۸) شرح داده شدند. عملگر با اندازه‌های مجاورتی مختلف ابتدا برای استخراج اطلاعات لبه‌ها به کار برده شد. شبکه عصبی سپس در روش قطعی و عمل ترتیبی به کار بده شد. شبکه، از یک حالت اولیه با نرونهای فاصله ۲۰/۵، به یک حالت پایدار بعد از چندین تکرار (۱۵-۱۰)، همگرا شد. به طوری که در مثالها نشان داده شد، نتایج آخری خیلی خوب بودند.

شبکه همچنین در تغییر شکل اتفاقی بدون هیچ اصلاح مدل مهمی آزمایش شد. در حقیقت در این روش، «نویز گرمایی» سیستم را از حالت اولیه خودش با کاهش زیاد میزان همگرایی، رها ساخت. [۳۶]

شکل ۶-۷: تصویر گرانی سنجی

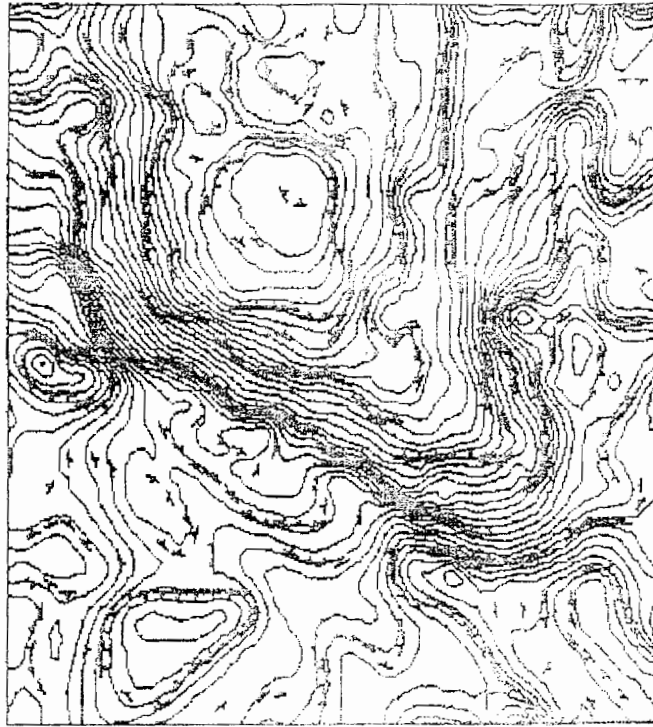


پوشش همبستگی هارالیک 5×5 Haralick

$E_0 = 160$ وفق اولیه N. N.

$E_f = 320$ وفق نهایی

$N = 12$ تعداد تکرارها



سنگ آهک شور کربناتی

پوسن همبستگی هارالیک Haralick ۵.۵

$E_0 = 160$ وفق اولیه N. N.

$E_r = 320$ وفق نهایی

$N = 12$ تعداد تکرارها

۴-۶- تشخیص مشخصه‌ها برای تصاویر میدانهای پتانسیلی با

استفاده از یک الگوریتم یادگیری برای ماشینهای بولتزمن^۲

۴-۶-۱- خلاصه

تفسیر داده‌ای میدان پتانسیلی به طور متداول دارای سه مرحله است. اولاً، یک مرحله تفسیر کیفی با استفاده از ترمیم و ترکیب داده‌ها برای به نقشه درآوردن ساختارهای اصلی و جهت‌های ساختاری ارائه شده. دوماً، یک مرحله افزایش داده‌ها با بکارگیری کاربرد فیلترها و استفاده از پردازش تصویر برای ساختن ساختارهای اصلی واضح‌تر و آسانتر برای تفسیر. و در آخر، یک مرحله تفسیر کمی از مدلسازی و وارونه‌سازی آنومالی‌های انتخاب شده بر پایه نتایج دو مرحله اول. تلاشهای اخیر در جهت گسترش روشهای خودکار تفسیر کیفی بر پایه تکنیکهای تخمین تصویر و تشخیص نمونه، پیش رفتند. در این مقاله، ما یک روش با استفاده از نمونه شبکه عصبی برای مشخصه‌های تشخیص داده شده در نگاره‌های میدانهای پتانسیلی پیشنهاد می‌کنیم. NN یک مجموعه از داده‌ها شامل اطلاعات لبه‌ای استخراج نشده با استفاده از مدل هموار با آگاهی از ساختارهای محلی برای تشخیص داده شدن برای ساختن یک مجموعه مرکب از مشخصه‌ها را جمع‌آوری می‌کند. این با استفاده از یک روش بهینه‌ساز ترکیبی توسط بهینه‌سازی یک تابع هزینه تعریف شده به عنوان یک ترکیب خطی از توابع ساخته شده بر روی خواص مشخصه قطعی، انجام می‌شود. یک الگوریتم یادگیری NN برای بهینه کردن ضرایب ترکیب خطی با استفاده از یک مجموعه از نمونه‌ها بسط داده شد.

۴-۶-۲- استخراج اطلاعات لبه‌ای

برای استخراج اطلاعات لبه‌ای ما از مدل هموار استفاده می‌کنیم. این الگوریتم با داده‌های مشخص شده توسط تغییرات فضایی هموار شده مانده داده‌های میدانهای پتانسیلی خیلی خوب انجام شد. ما فرض کردیم که آرایه مجزا از مقادیر (تصویر) می‌تواند به صورت یک ترکیب خطی از

۲- مطالب این بخش از منبع [۳۸] آورده شده است.

حاصل ضرب‌های تانسور از چند جمله‌ای‌های متعامد مجزا، بیان شود. برای هر پیکسل از تصویر یک تابع دو مکعبی $f(r,c)$ وجود دارد که با همسایگی پیکسل منطبق می‌شود و با استفاده از:

$$f(r,c) = k_1 + k_2r + k_3c + k_4r^2 + k_5rc + k_6c^2 + k_7r^3 + k_8r^2c + k_9rc^2 + k_{10}c^3$$

یک پیکسل به عنوان یک پیکسل لبه‌ای علامت‌گذاری می‌شود اگر برخی نقاط، در ناحیه پیکسل با داشتن یک محل تقاطع صفر از مشتق جهتی دوم که در جهت یک شیب غیر صفر در مرکز پیکسل گرفته شده، وجود داشته باشد.

گرادینان برای پیکسل مرکزی توسط فرمول زیر داده می‌شود:

$$\nabla f(0,0) = [k_2, k_3]$$

اگر V یک Versor در جهت $\nabla f(0,0)$ باشد آنگاه

$$g(\rho) = f(\rho v) = A\rho^3 + B\rho^2 + C\rho + D$$

که ρ = فاصله از پیکسل مرکزی.

اگر برای برخی ρ ، $|\rho| < \rho_0$ که اندکی کوچکتر از طول ضلع یک پیکسل است

$$g''(\rho) = 0$$

$$\text{sign}(g'''(\rho))\text{sign}(g'(\rho)) < 0$$

ما پیکسل مرکزی را به عنوان یک پیکسل لبه‌ای نشان می‌کنیم. این مدل اجازه می‌دهد برای نشان دادن لبه‌ها در مقیاس‌های مختلف با استفاده از یک پنجره مجاورتی متفاوت.

خروجی عملگر برای تعیین مشخصه‌ها کافی نیست زیرا تصمیم بر روی وجود یک پیکسل لبه‌ای، بدون فرض ساختار لبه‌ای محلی در همسایگی پیکسل گرفته می‌شود. در حقیقت، تعریف یک مشخصه بر مبنای ویژگی‌های انتخاب شده که مشخصه بایستی داشته باشد، است.

۶-۴-۳- یادگیری بهینه

روش تعریف مشخصه پیشنهاد شده در این جا بر پایه پیدا کردن راه حل تقریباً بهینه از

یک تابع مقداردهی شده حقیقی است:

$$V(\underline{x}) = V(x_1, x_2, \dots, x_n) \quad x_i \in \{0,1\}$$

ماشین بولترمن یک الگوریتم بهینه‌سازی احتمالی است که بر قصد غلبه بر مساله V که

افتادن از تله مینیمم موضعی است، قرار گرفته است. (Aarst and Korst . 1989)

بر پایه استراتژی شناخته شده (Metropolis et al., 1953)، آن "0,1" کشف گردید توسط میانگین یک مجموع استاندارد Gibbs متناسب با $\exp(-\beta V(\underline{x}))$ برای یک β نامنفی داده شده. BM یک NN حالت‌های دوتایی با نرونهای فعال تصادفی است.

$$p_{x,x'} = P\{X(t) = x' | X(t-1) = x\}$$

$$p_{x,x'} = \begin{cases} G' A'(x) & \text{se } x' = I, x \\ 1 - \sum_j p_{x,x'} & x = x' \\ 0 & \text{otherwise} \end{cases}$$

اگر V یک تابع از برخی پارامترهای جریمه باشد:

$$V = V(\underline{x}, \underline{\lambda}) \quad \underline{\lambda} \in \mathbb{R}^n$$

که $\underline{\lambda}$ از عمل روی \underline{x} می‌آید. عموماً این متغیرها به صورت پارامترهای آزاد در نظر گرفته می‌شوند که مجبور شدند با یک دسته از نمونه‌های مساله میزان شوند.

اینجا ما برای یادگیری زیر نمادین مقادیر مناسب $\underline{\lambda}$ از یک مجموعه نمونه‌ها را پیشنهاد می‌کنیم. BM یک ابزار خیلی موثر در انجام این کار است، زیرا توانایی هر دو توابع بهینه‌ساز و یادگیری توزیع‌های متغیر تصادفی را دارد. ترکیب مناسب این دو قابلیت اجازه می‌دهد برای ساختن یک ماشین که یاد می‌گیرد از نمونه‌ها به بهینه‌سازی V . ما این الگوریتم را یادگیری بهینه می‌نامیم.

۶-۴-۴ - تشخیص مشخصه‌ها

برای هر پیکسل نگاره، ما یک نرون NN شرکت می‌دهیم. حالت هر نرون یک پیکسل از نگاره را طبقه‌بندی می‌کند به طوری که وضع کلی مطلق شبکه یک راه حل کاندید مساله است. تابع ارزش توسط اتصالات و مقادیر آستانه تعیین می‌شود. این مقادیر توسط آزمایش خواص محلی مشخصه‌ها تعیین می‌شود. این خواص به صورت زیر تعریف می‌شود:

- پیوستگی

- موضعی کردن دقت

- سستی

به طوریکه توسط روزنبرگ نشان داده شده، ما می‌توانیم به آسانی هر تابع را در یک تابع درجه دوم تعریف شده روی $\{0, 1\}^n$ تفسیر کنیم به طوری که قابلیت بهینه‌سازی ما می‌تواند به همه توابع گسسته بسط داده شود. هر خاصیت یک تابع برای بهینه‌سازی تولید می‌کند. مساله تعیین مقادیر بهینه برای ضرایب ترکیب خطی است.

ما الگوریتم OPTILEARNING را با یک مجموعه آموزشی ساخته شده از دو جفت ورودی / خروجی به کار بردیم: نمونه‌های بهینه‌سازی، راه حل‌های دقیق، تولید شده توسط کارشناسان ژئوفیزیکی.

۶-۴-۵- نتایج

ما از چندین مثال برای آزمایش الگوریتم آموزشی خودمان استفاده کردیم. برای به دست آوردن اجرایی خوب، لازم بود که از یک برنامه ذوب آرام در طول یادگیری استفاده کنیم. نمونه‌های استفاده شده در طول مرحله یادگیری، برای آزمایش انجام یادگیری با یک BM با پارامترهای متصل شده در وزنهای آموزش دیده، بهینه‌سازی شده بودند. BM بهینه‌سازی شده روی چندین نگاره استفاده شده در طول آموزش با نتایج خیلی خوب، آزمایش شدند. [۳۸]

مثال A: تصویر گرانی سنجی



(1) مشخصه‌های تشخیص داده شده توسط یک متخصص ژئوفیزیک



(2) مشخصه‌های تشخیص داده شده توسط الگوریتم BM

۶-۵- وارونه‌سازی داده‌های گرادایومتری گرانی‌سنجی با استفاده از

یک شبکه عصبی

۶-۵-۱- خلاصه

یک شبکه عصبی پس انتشار مصنوعی برای وارونه‌سازی داده‌های گرادایومتر گرانی‌سنجی برای انحراف (آشفستگی‌ها) از یک حذف‌اصل تشکیل شده از ۱۸ عنصر مرزی مجزا استفاده می‌کند که یک نیم فضای با چگالی بالا را از یک لایه سطحی با چگالی کم‌تر (تباين چگالی ۰/۱ cc / g) جدا می‌سازد. وارونه‌سازی داده‌های شیب گرانی برای یک چنین مدل چگالی ساده یک راه حل منحصر به فرد دارد. در طول آموزش، شبکه به ۱۰٪ مینیمم خطا بعد از ۱۵۰ تکرار آموزشی رسید. هدف از این بررسی، گسترش یک الگوریتمی است که بتواند به طور موفقی یک علامت شیب مشاهده شده را به ساختار سنگ کف مربوط به آن علامت نگاشت کند. با تجهیز به یک چنین الگوریتمی ما بر آن می‌شویم که شبکه عصبی بر پایه وارونه‌سازی را با یک وارونه‌سازی مرسوم قبلی از داده‌های گرادایومتری گرانی مقایسه کنیم (Vasco and Taylor, 1990). این مجموعه داده‌ها در طول ۷ شیار شمالی جنوی در مرکز جنوب جمع‌آوری شده است. ریزانی‌های گسل از چندین کیلومتر، یک علامت شیب گرانی را به یک مقدار بیشتر از ۵۸ نشان می‌دهد.

۶-۵-۲- مقدمه

برای این مطالعه ما از الگوریتم پس انتشار استفاده کردیم. ما در اینجا همچنین ما توانایی یک شبکه عصبی پس انتشار را برای وارونه‌سازی داده‌های گرانی مصنوعی برای ساختار چگالی ارزیابی می‌کنیم.

۶-۵-۳- طراحی آزمایشی

در یک بررسی پیشین، مولف وارونه‌سازی داده‌های ثبت شده توسط سیستم کاوش گرادایومتری گرانی را در نظر گرفت. ما از ۱۲۶ اندازه‌گیری شیب سه مولفه‌ای برای حل برای عمق‌ها

به ۹۶ عنصر شامل یک جدایی یک پیچیدگی زیرزمینی با چگالی بالاتر از یک لایه سطحی با چگالی کمتر استفاده کردیم.

در اندازه‌گیری شیب گرادیان، همه ۶ مولفه تا نسور گرادیان دوباره بدست می‌آیند. هر مولفه داده شده، توسط تفاضل علامتهای ثبت شده توسط دو ثقل سنج با فاصله ۱۰ سانتیمتر تعیین می‌شوند. ثقل سنج‌های نادرست به علت تکانهای حمل کننده به طور برابر توسط هر دو ثقل سنج آزمایش خواهند شد و بوسیله تفاضل رفع می‌شوند. به هر حال در حضور یک میدان گرانی ثقل‌های ثبت شده توسط دو ثقل سنج می‌تواند برای بدست آوردن شیب در جهت خط مستقیم با دو ثقل سنج متفاوت باشد.

داده‌ها بر روی حوزه رودخانه از اوکلاهما مرکزی جنوبی گردآوری شده است، در جایی که توپوگرافی نسبتاً ثابت است، که آن جا سیگنال‌های گرانی بزرگی مربوط به خرابی‌های گسل خیلی بزرگ در هم تافته‌های زیرزمینی وجود دارد، و جایی که سیگنال توسط اندازه‌گیری‌های گرانی سطحی بهتر مستند می‌شود (شکل ۶-۹).

در این آزمایش زمین فرضی ما شامل یک نیم فضای پرچگال پوشیده شده توسط یک لایه با یک چگالی 0.1 g/cc کمتر از نیم فضای آن، می‌شود. جدایی میانی لایه از جسم لایه زیرین به یک آرایه 3×6 از عناصر مرزی تقسیم شده است که از یک عمق میانی میانگین ۱ کیلومتری آشفته شده‌اند. هدف از وارونه‌سازی، مشخص کردن عمق‌ها از سطح برای ۱۸ عنصر مرزی با استفاده از تنها آنومالی‌های شیب گرانی مصنوعی است.

مجموعه آموزشی ما شامل ۶۰۰ جفت آموزشی است. ما هر بردار هدف را توسط برداشت تصادفی عمق‌ها به ۱۸ عنصر مرزی، بدست آوردیم. بردار ورودی متناظر سپس توسط در نظر گرفتن مساله مستقیم برای محاسبه شیب گرانی عمودی در ۱۸ نقطه مشاهده‌ای روی هر ۱۸ عنصر مرزی، محاسبه شده بود. شبکه سپس برای نگاشت شیب‌های گرانی محاسبه شده به توزیع چگالی از آن‌هایی که شیب‌های گرانی محاسبه شده بودند، آموزش داده شده بود.

شبکه برای ۴۶۰۰ تکرار برای ۴ ساعت و ۴۰ دقیقه از زمان C PU روی یک 8650 VAX دیجیتال آموزش یافت. شکل (۶-۱۰)، خطای باقیمانده جمع شده بر روی کل مجموعه آموزشی در مقابل ۲۰۰۰ تکرار اولیه رسم شده است. ما از این نمودار نتیجه می‌گیریم که خطا در تقریباً ۱۳۰۰ تکرار کمینه شده بود، (۱ ساعت و ۲۰ دقیقه از C PU)، اگر چه خطا خیلی نزدیک به مینیمم بعد از ۱۵۰ تکرار بود (۹ دقیقه CPU).

برای آزمایش این که چطور شبکه ما خوب یادگیری داشته، ما یک ساختار میانی انتخاب کردیم که بخشی از مجموعه آموزشی نبود. ساختاری که ما انتخاب کردیم یک افقی بود در یک عمق ۱۰ کیلومتری پخش شده توسط یک برجستگی لبه‌ای کم عمق کشیده شده از شرق در یک عمق صفر کیلومتری به طوری که در پلن اولی از شکل (۶-۱۱) نشان داده شده است. شیب‌های گرانی عمودی، سپس از این مدل چگالی در ۱۸ نقطه هر یک عنصر میانی پوشیده محاسبه شده بودند. این مقادیر شیب، عناصر بردار ورودی را در بر داشت و برای شبکه آموزش یافته به کار گرفته شده بودند.

نتیجه بردار خروجی در پلان دوم از شکل (۶-۱۰) نمایش داده شده است. ساختار اولیه در مدل توسط شبکه خوب تصویر شده بود به طوری که می‌توانست توسط درجه بالایی از تشابه بین دو نقشه سطحی، به نظر برسد. سومین پلان شکل‌ها سطح خطای به دست آمده به وسیله تفریق بردار خروجی از بردار هدف را نشان می‌دهد. برای یک تصویر نگاری کامل از مدل هدف ما بایستی انتظار داشته باشیم که این سطح خطا هر کجا برابر با صفر شود.

آزمایش مختصر ما نشان داد که یک شبکه عصبی می‌تواند برای وارونه‌سازی داده‌های شیب گرانی برای تصویر ساختار چگال که علامت شیب قابل مشاهده ایجاد می‌کند، آموزش یابد. ما ذکر کردیم که اجراء روی یک پردازشگر موازی بهبود خواهد یافت. شایستگی شبکه‌های عصبی این است که آن‌ها به راحتی بر روی یک محاسبه گر موازی به کار برده می‌شوند، آن‌ها تنها محاسبات تابعی هدف - نه مشتقاتش، را لازم دارند، و اینکه نویز داده‌ها به راحتی در مجموعه آموزشی شامل می‌شود. نقص‌های آن‌ها شامل، تولید مینیمم موضعی به سبب غیر خطی بودن ذاتی می‌شود، و اینکه آن‌ها به دقیقی تکنیک‌هایی که از شیب تابعی هدف استفاده می‌کنند، نیستند.

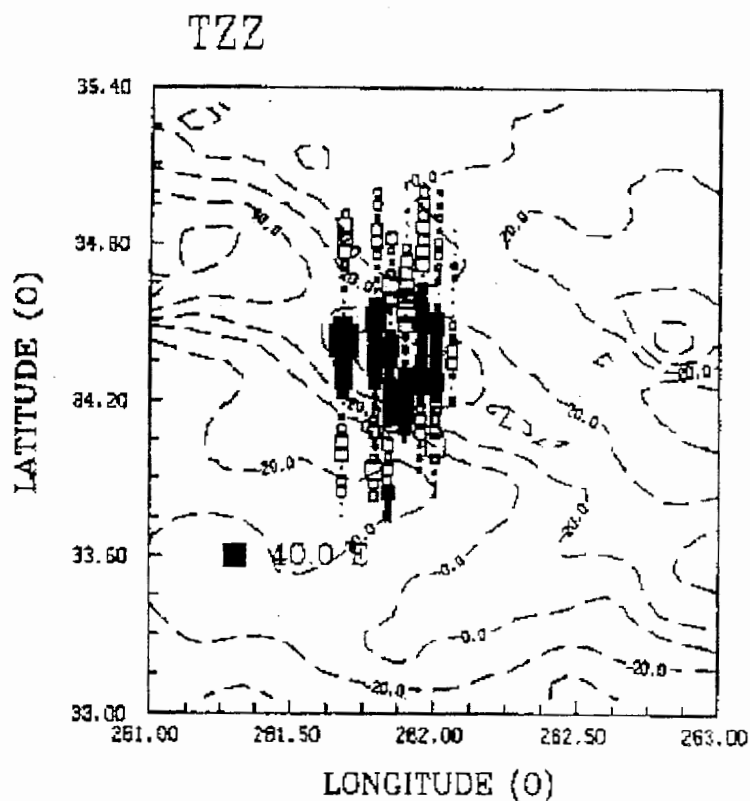
تعدادی از جنبه‌های این تکنیک که فعلا فهمیده نشدند اما ما به شرح آنها پرداختیم اینجا وجود دارد.

آن‌ها شامل:

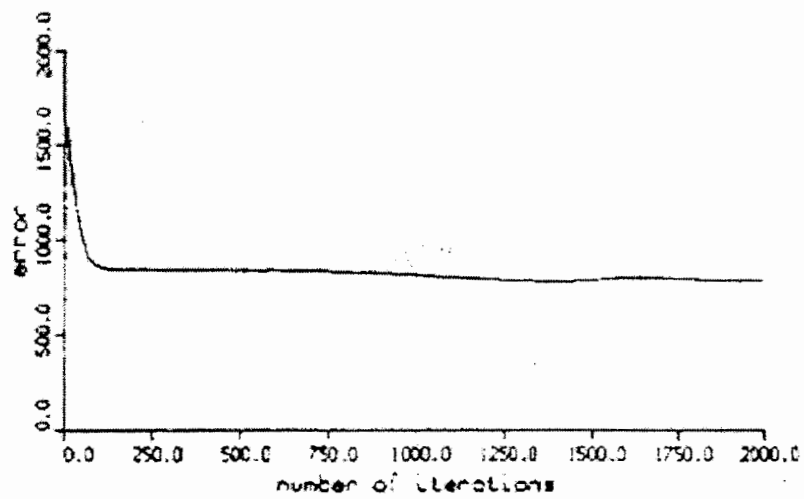
الف) اثر نویز در مشاهدات

ب) روابط بین تعداد مدلها و اندازه مجموعه آموزشی نسبت به توانایی

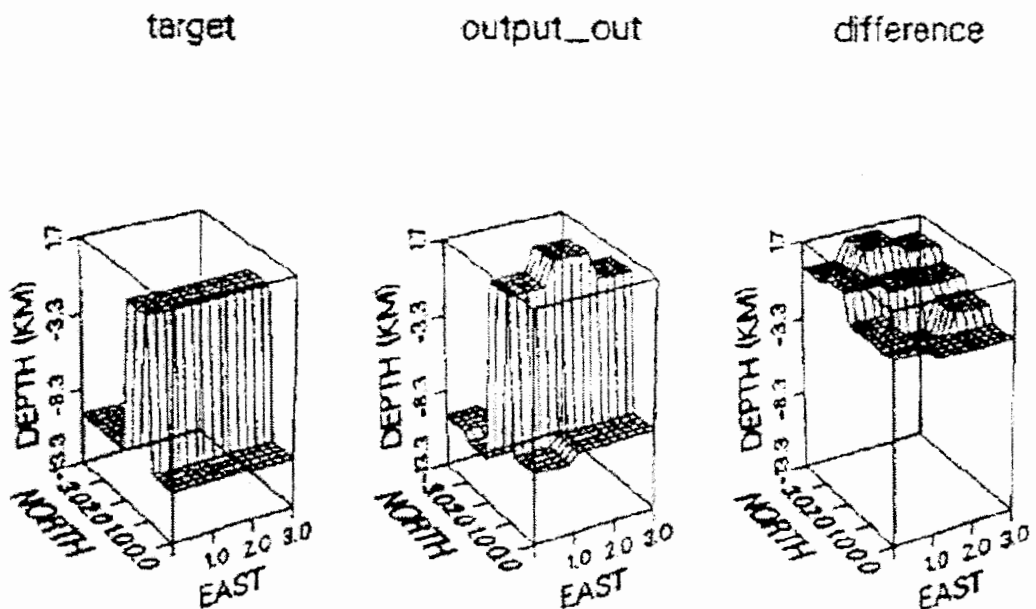
ج) اثرات ساختارهای شبکه‌ای گوناگون، مثلا، تعداد لایه‌های میانی و تعداد نرونها در لایه‌های میانی.



شکل (۶-۹). مولفه عمودی شیب گرانی، اندازه هر مربع، متناسب با مقدار شیب است؛ مقادیر مثبت توپر نشان داده شده‌اند و مقادیر منفی توخالی. گرانی سطح میانگین منطقه‌ای به شکل خطوط میان خط تیره در یک فاصله ۲۰ میلی‌گالی رسم شده است.



شکل ۶-۱۰: خطای باقیمانده نسبت به تعداد تکرارهای آموزش



شکل (۶-۱۱): نقشه‌های سطحی برای نشان دادن مدل هدف (چپ)، خروجی مدل توسط شبکه (مرکز)، و تفاوت بین مدل‌های هدف و خروجی (راست).

فصل هفتم

کاربرد شبکه‌های عصبی در چاه‌پیمایی

۱-۷- محاسبات عصبی در چاه‌پیمایی

شکل (۱-۷) یک شبکه عصبی ساده سه لایه را نمایش می‌دهد که نرون‌ها را با دایره نشان داده، و ارتباطات آن‌ها با وزنهای مربوطه، W_{kj} ، با خطوط نمایش داده شده است.

۱-۱-۷- نمونه آموزش شبکه

بهترین روش برای شرح این روند آموزش بوسیله یک مثالی است که از دو نمودار ترکیبی برای ارزیابی ساختار زیرزمین استفاده می‌کند. فرض کنید ما پتانسیل خودزا و نمودار مقاومت ویژه را برای یک چاه داریم، و می‌خواهیم یک شبکه عصبی را برای تولید یک نمودار ساختار شناسی با استفاده از مقادیر اندازه‌گیری شده از این دو نمودار الکتریکی آموزش دهیم. (شکل ۲-۷). ساختارهای زیرزمین به طور ناخالص به سه رده دسته‌بندی خواهد شد: ماسه سنگ‌ها، مشیل‌ها، و سنگ آهک‌ها. ساختار شبکه عصبی استفاده شده برای این مساله در شکل (۱-۷) نشان داده شده است آن شامل دو نرون ورودی (یک نرون برای دریافت یک مقدار برای نمودار SP، و دومی برای دریافت یک مقدار برای نمودار مقاومت ویژه در عمق یکسان)، و سه نرون خروجی - هر یک برای هر کدام از سه ساختار ممکن. نرونهای خروجی ۱-۳ به ترتیب ماسه سنگ، شیل و سنگ آهک را نشان

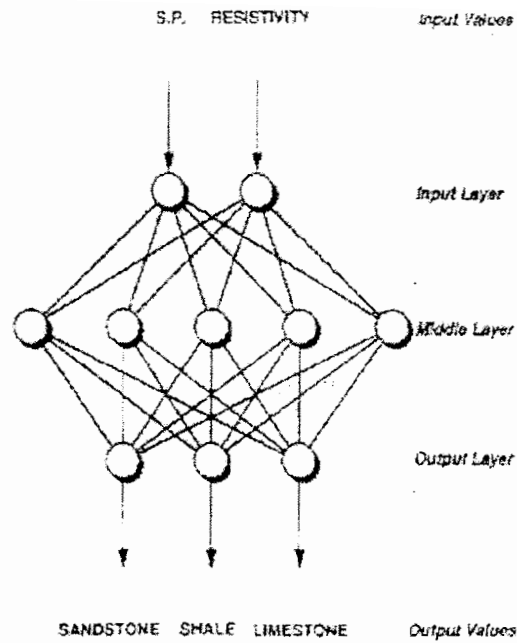
می‌دهند. سیگنال هر نرون خروجی اگر یک عدد صفر باشد، نشان دهنده این است که این ساختار وجود ندارد، و اگر عدد یک باشد نشان دهنده این است که این ساختار وجود دارد. بنابراین یک جفت ورودی از مقادیر نمودار مقاومت ویژه - SP بایستی در نرونهای خروجی (۱۰۰) تولید کند اگر ساختار ماسه سنگ است، (۰۱۰) اگر شیل وجود دارد، و (۰۰۱) برای نشان دادن سنگ آهک. یک لایه میانی شامل ۵ نرون ساختار شبکه را کامل می‌کند. انتخاب تعداد لایه‌های میانی (می‌تواند هر عددی باشد، حتی صفر) و تعداد نرونها در هر لایه میانی یک هنر است. همچنین کم بودن نرونهای لایه میانی آموزش شبکه را مشکل می‌سازد و همین زیاد بودن نرونها باعث می‌شود شبکه نمونه‌های آموزشی را به خاطر بسپارد و هنگام ارزیابی ورودی‌های جدید به طور ضعیفی اجرا شود.

جدول (۱-۷) شامل هفت جفت از مقادیر نمودار SP- مقاومت ویژه است که از عمقهای مشخص شده در شکل (۲-۷) برای سه ساختار انتخاب شده بودند.

برای انجام فرایند آموزشی، وزنهای برای هر اتصال داخلی شبکه، با مقادیر تصادفی کوچک آغاز می‌شوند. همانطور که ممکن است حدس بزنید، وقتی که جفت‌های مقاومت ویژه- SP وارد این شبکه، که مقادیر وزنی تصادفی دارد، می‌شوند سیگنالهای خروجی (و از آنرو تخمین ساختار) کاملاً اشتباه هستند. بنابراین، وقتی که ما ورودی‌ها را وارد می‌کنیم (برای مثال ۱)، مقادیر واقعی برای سه نرون خروجی ممکن است (۰/۴ ۰/۸ ۰/۱)، بیشتر از خروجی مطلوب (۰۰۰)، باشد تفاوت بین مقادیر واقعی و مطلوب در هر نرون خروجی خطای خروجی است. این اختلاف‌های خطا برای اصلاح وزنهای شبکه مورد استفاده قرار می‌گیرند، به طوری که اگر ما ورودی‌ها برای مثال ۱ را دوباره وارد کنیم، مقادیر خروجی واقعی اکنون به خروجی‌های مطلوب نزدیکتر خواهد بود. خطاها در هر نرون در شبکه از لایه خروجی به لایه ورودی به عقب انتشار می‌یابند با تغییرات اعمال شده به هر نرون به در مقایسه با (۱) هم بخشی آن اتصال اعمال شده به خطای سرتاسر و (۲) بزرگی سیر در طول آن انتقال. نام شبکه عصبی «پس انتشار» از روش اصلاح وزنهای اتصالات داخلی توسط انتشار خطا به عقب گرفته شده است. نمونه‌های ۲-۷ از جدول (۱-۷) به ترتیب ورودی برای شبکه هستند. تغییرات برای وزن اتصالات داخلی بعد از هر نمونه به شبکه ارائه می‌شود اگر خروجی‌های واقعی و مطلوب عمدتاً تفاوت داشته باشند. به طور نرمال چند صد یا چندین هزار تکرار قبل از اینکه وزنهای شبکه تنظیم شوند، اتفاق می‌افتد به طوری که خطاهای خروجی برای

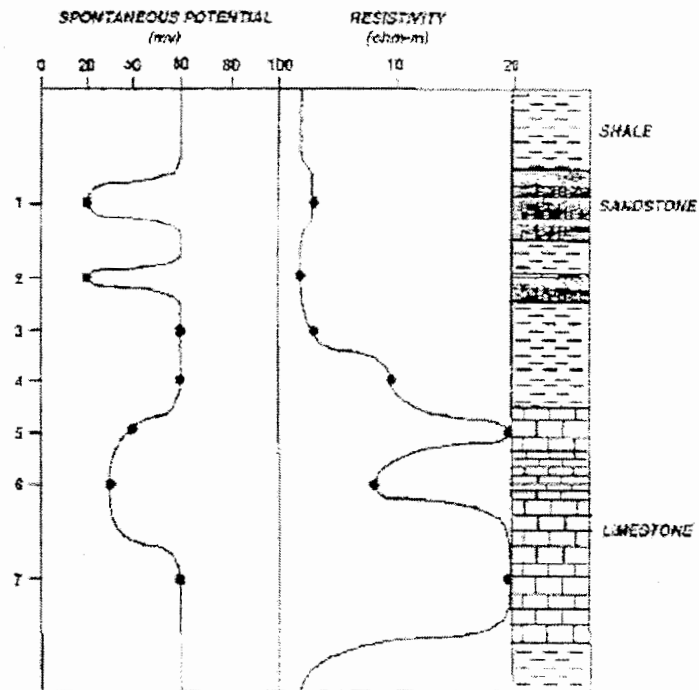
همه هفت مثال کمتر از یک حد مشخص هستند. برای مثال نشان داده شده در جدول ۱، ۸۴۵ تکرار برای تنظیم شدن وزنه‌های اتصالات داخلی لازم بود به طوری که مقادیر نزونهای خروجی واقعی و مطلوب بیشتر از ۱۰ درصد اختلاف نداشتند. در این مرحله، گفته می‌شود شبکه عصبی آموزش یافته و وزنه‌های اتصالات داخلی برای استفاده بعدی ذخیره می‌شوند.

بعد از آموزش کامل، مقادیر جدید مقاومت ویژه - SP برای دیگر عمقها در چاه می‌توانند به شبکه عصبی برای تولید یک نمودار ساختارشناسی برای داخل چاه وارد شوند. نتیجه نمودار ساختار شناسی در کناره راست شکل (۳-۷) نشان داده شده است. شبکه عصبی در این روش از عملکرد خیلی سریع اجرا می‌شود چون وزنه‌های متصل منطبق شده هستند و داده‌ها به جلو از ورودی به خروجی حرکت می‌کنند (در اینجا زمان مصرفی پس انتشار خطا وجود ندارد). شکل (شکل ۳-۷) یک نقشه است که نشان می‌دهد چطور شبکه عصبی آموزش یافته می‌تواند هر جفت ورودی SP - مقاومت ویژه داده شده را به یکی از سه ناحیه ساختاری نگاشت کند. مرزهای بین نواحی ساختاری مختلف به طور خودکار توسط نمونه‌های انتخاب شده در طول دوره آموزشی تعیین می‌شوند و در وزنه‌های اتصالات داخلی کدبندی می‌شوند. اگر تحلیل گر نمودار با این نتیجه راضی نباشد، او می‌تواند دوباره شبکه را بوسیله فراهم آوردن نمونه‌های اضافی حرکت دادن مرزها به موقعیت مناسب‌اشن آموزش دهد.



شکل ۷-۱: یک شبکه عصبی سه لایه کاملاً متصل

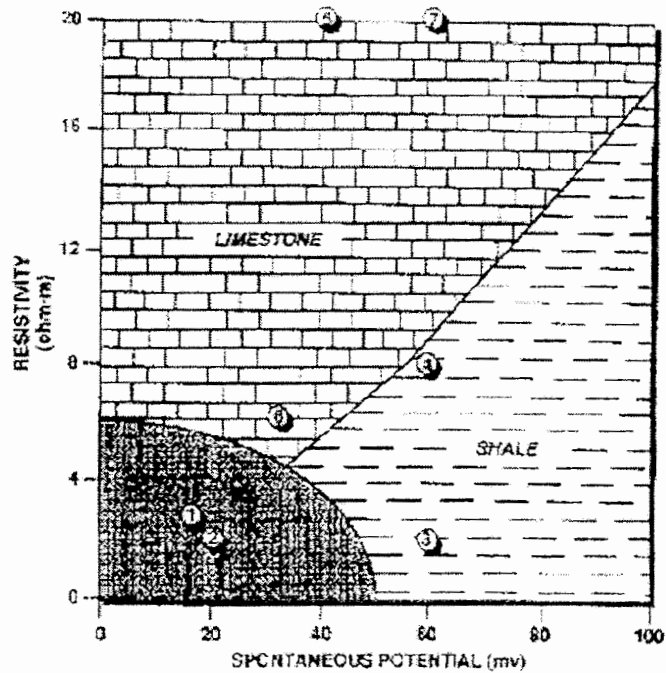
شکل ۷-۲: پتانسیل خودزا مصنوعی (SP) و نمودار مقاومت ویژه در حال نشان دادن موقعیتهای عمقی هفت جفت مقادیر نمودار مقاومت ویژه-SP برای آموزش یک شبکه عصبی پس انتشار. نمودار ساختارشناسی محاسبه شده توسط شبکه عصبی در سمت راست نشان



داده شده است.

جدول ۱-۷: مقادیر نمودار مقاومت ویژه-SP

No.	Lithology	SP (mv)	Resistivity (ohm-m)	Desired output
1	Sandstone	17	3	(1 0 0)
2	Sandstone	20	2	(1 0 0)
3	Shale	60	2	(0 1 0)
4	Shale	60	8	(0 1 0)
5	Limestone	40	20	(0 0 1)
6	Limestone	25	6	(0 0 1)
7	Limestone	60	20	(0 0 1)



شکل ۳-۷: مقطع عرضی مقاومت ویژه-SP در حال نشان دادن سطح تصمیم‌گیری برای سه ساختار توسعه یافته در طول آموزش با هفت نمونه در جدول (۱-۷). موقعیت هفت نمونه آموزشی همچنین در مقطع عرضی نشان داده شده است.

۷-۲- وارونه سازی شبه دوبعدی بلادرنگ داده های چاه پیمایی

آرایش مقاومت ویژه^۱

۷-۲-۱- خلاصه

ما برای یک ابزار آرایه ای برقی جدید ، بوسیله کاهش بعدی و شبیه سازی شبکه عصبی یک الگوریتم وارونه سازی آنی شبه دوبعدی ارائه می دهیم. با استفاده از عمل متقابل و انطباق ، ما یک تکنیک متمرکز کردن عددی برای داده های نامتمرکز بکار می بریم. داده های به طور عددی متمرکز شده خیلی کمتر تحت نتایج دو بعدی و لایه بندی قرار می گیرند و می تواند به همان اندازه از یک زمین یک بعدی استوانه ای تخمین زده شود. ما سپس روی داده های متمرکز شده برای تهیه اطلاعات تقریبی درباره ساختار مقاومت مخصوص دو بعدی وارونه سازی یک بعدی انجام می دهیم. یک شبکه عصبی برای انجام مدلسازی مستقیم در وارونه سازی یک بعدی استفاده شده است ، که چند صد برابر سریعتر است از راه حل های مستقیم عددی مرسوم. تست الگوریتم وارونه سازیمان در هر دو داده های ترکیبی و زمینی نشان می دهد که این الگوریتم وارونه سازی سریع ، برای تهیه اطلاعات مقاومت مخصوص آرایه ای در محل چاه مفید است. شبکه های عصبی به طور فزاینده ای در ژئوفیزیک عمومی شده اند. زیرا آنها تقریب زنده هایی عمومی هستند، این ابزار می تواند هر تابع پیوسته ای را با دقت دلخواه تقریب زنند. از این رو، آنها ممکن است سهم مهمی برای پیدا کردن راه حل هایی برای یک قسم از کاربردهای ژئوفیزیکی ارائه دهند. به هر حال ، علم تعدادی از روشها و تکنیکها که اخیراً برای افزایش اجرا و برای آسان کردن استفاده از شبکه های عصبی گسترش یافته اند به نظر نمی رسد که در جامعه ژئوفیزیکی متداول شده باشد. بنابراین، قدرت این ابزار هنوز به اندازه کامل خودش کشف نگردیده است.

۷-۲-۲- مقدمه

فرآیند حفاری و خواص مایع های حفاری مربوطه عمدتاً می تواند ساختمان و خواص مایع ساختمان نزدیک دیواره چاه را تغییر دهد. منطقی آغشته ممکن است شعاعی بیشتر از چندین متر

۱- مطالب این بخش از منبع [۴۰] آورده شده است.

داشته باشد؛ بنابراین، داده های چاه پیمایی مقاومت مخصوص توسط هر دو منطقه آغشته و دست نخورده تحت تاثیر قرار می گیرند. الگوریتم های وارونه سازی دو بعدی برای ترسیم تغییرات عمودی و افقی ساختارهای مقاومت مخصوص لازم هستند. به علت اینکه میزان نمونه گیری چگال و شمار زیادی از فواصل منابع دریافت کننده در چاه پیمایی مقاومت مخصوص آرایه ای، وارونه سازی های دو بعدی مرسوم داده های چاه پیمایی مقاومت مخصوص آرایه ای خیلی اوقات مصرف می شوند و نمی توانند نتایج وارونه سازی را در محل یک چاه برای تسهیل تصمیمات نهایی خوب ارائه دهند.

ابزار چاه پیمایی برقی آرایه ای جدید اطلاعات مفیدی درباره حجم هیدروکربن در ساختار تهیه می کند. ابزار نمودار گیری جانبی با دقت بالا (HDLL) که توسط Baker Atlas و Hakvoort Shell (et al., 1998) توسعه یافت یکی از چنین ابزار آرایه ای جدید تولید نیرو است. تفسیر سریع و دقیق داده های آرایه ای برای تصمیمات بموقع در تکمیلات خوب لازم است. ابزار HDLL، به هر حال، مجموعه داده های زیادی را تولید می کند که برای پارامترهای ساختار نامتمرکز و غیر خطی هستند. تفسیر این داده ها با استفاده از روشهای وارونه سازی دو بعدی مرسوم می تواند وقت گیر باشد. یک نمودارگیر HDLL نمونه ای ممکن است چندین صد متر را اندازه گیری کند، با استفاده از یک میزان نمونه ۴ سانتیمتری و جمع آوری ۳۸ نقاط داده ای در هر سطح. هر پنجره عمق تفسیر در یک معکوس سازی دو بعدی میتواند به طول ۳۰ متر باشد. از اینرو، وارونه سازی در هر پنجره عمق ۱۰^۴ داده و ۱۰^۲ پارامتر نمونه را درگیر می کند. بنابراین تفسیر داده های HDLL تشکیل یک برنامه معکوس غیرخطی بزرگ مقدار میدهد. بنابراین، معکوس سازی دو بعدی مرسوم ۳۰۰ متری داده های HDLL ممکن است احتیاج به چندین روز در یک ایستگاه کار جدید داشته باشد. توسعه تقریبی لازم است اما الگوریتمهای معکوس سازی دو بعدی سریع برای تسهیل تصمیمات نهایی خوب در محل چاه.

شبکه های عصبی یک راه حل ممکن برای مسائل معکوس بزرگ میزان در ژئوفیزیک و چاه پیمایی ارائه می دهد. (Poulton et al. (1992) شبکه های عصبی را برای تعیین محل اهداف زیر سطح زمین از داده - های ژئوفیزیکی استفاده کرد. (Swiniarski et al. (1993) شبکه های عصبی را برای مسائل معکوس ژئوفیزیکی بکار برد. (Zhang et al. (1999b) از شبکه عصبی پیمانه ای برای شبیه سازی نمودارگیر لاترولوگ دوتایی DLL و نمودارگیر لاترولوگ کوچک MLL

استفاده کرد اما کنترل مشکلی برای خطای rms زیر ۰.۵٪ داشت ، سطح غیر متجانس لازم برای بیشتر وارونه سازی های داده های زمینی .(این به این خاطر است که (۱) پیچیدگی تابعی غیرخطی که مدل را به داده ها ترسیم می کند و (۲) شمار زیادی از پارامترهای متوالی در توالی درگیر شده اند.)

ما یک الگوریتم وارونه سازی شبه دو بعدی آنی را با استفاده از شبکه عصبی گسترش می دهیم. بجای پرورش مستقیماً شبکه عصبی روی مدل‌های دوبعدی ، ما اول پیچیدگی مسئله مدل‌سازی مستقیم را توسط کاهش بعد پایین می آوریم و سپس شبکه عصبی را روی یک مدل زمینی خیلی ساده شده با ایجاد یک الگوریتم وارونه سازی آنی ، پرورش می دهیم .کاهش بعدی مدل دوبعدی مرسوم را در شکل (۷-۴) a) به یک سری از مدل‌های یک بعدی (شکل ۷-۴) b) تبدیل می کند. کاهش بعدی ،توسط وارونه کردن داده های دوبعدی خام به داده های (SFR) مقاومت مخصوص به طور ترکیبی متمرکز شده حاصل می شود. این فرآیند متمرکز کردن آثار s holder را روی داده ها کاهش می دهد؛ از اینرو داده های SFR میتوانند به عنوان یک مدل زمینی استوانه ای تلقی شوند. برای کاهش بیشتر شماره پارامترهای آزاد در پرورش شبکه عصبی ، ما تصحیحات گمانه داده های SFR را انجام می دهیم. مدل زمین بنابراین به یک مدل نشان داده شده در شکل (۷-۴) c) ساده شده ، که می تواند تنها توسط سه پارامتر مدل توصیف شود: R_t ، مقاومت ویژه سازند ؛ R_{∞} ، مقاومت ویژه ناحیه آغشته شده ؛ و L_{∞} ، طول ناحیه آغشته شده است. الگوریتم های مدل سازی مستقیم یک بعدی مرسوم وجود دارند، اما آنها به اندازه کافی برای وارونه سازی های محل چاه سریع نیستند. در این مقاله ، ما از شبکه های عصبی استفاده می کنیم برای انجام مدل سازی مستقیم یک بعدی با سرعت محاسبه ای دو تا سه برابر سریعتر از مدل‌سازی مستقیم یک بعدی مرسوم. شبکه های عصبی برای آموزش روابط بین پارامترهای مدل زمین یک بعدی و داده های SFR پرورش داده شده اند . شبکه های عصبی آموزش داده شده هستند معتبر شده و کاربردی شده در یک الگوریتم وارونه سازی به عنوان اسباب مدل سازی مستقیم برای شبه سازی داده های SFR . ما الگوریتم وارونه سازی شبه دوبعدی آنی خودمان را روی داده های زمینی و ترکیبی تست می کنیم.

۷-۲-۳- روش

ابزار چاه پیمایی مقاومت ویژه آرایه ای HDLL شامل می شود از یک الکتروود جریان و الکتروودهای اندازه گیری هجده (Itskovich et al., 1998). داده های اندازه گیری شده شامل هشت ولتاژ هستند بین دریافت کننده ها و الکتروود منبع در بینهایت و شانزده اختلاف اول مرتبه بین دریافت کننده های مجاور. چهارده تفاضل دوم مرتبه بین سه الکتروود همسایه با استفاده از تفاضلهای اول اندازه گیری شده محاسبه می شوند. مدل دوبعدی مان یک گمانه را شامل می شود، در ناحیه آغشته شده با مقاومت ویژه R_{x0} و به درازای L_{x0} ، و ناحیه بکر که مقاومت ویژه آن با R_t تعریف شده (شکل ۷-۴ a).

۷-۲-۳-۱- کاهش بعدی

کاستن بعد وارونه سازی از دوبعد به یک بعد توسط متمرکز کردن عددی حاصل می شود. اولین کار روی متمرکز کردن عددی می تواند توسط Roy (1981) پیگیری شده باشد، کسی که استفاده کردن از پاسخ های چندین الکتروود را که با فاصله معینی جایگزاری شده اند را برای محاسبه نمودارهای جانبی ترکیبی پیشنهاد کرد. با استفاده از این قاعده، انواع مختلفی از نمودارهای جانبی ترکیبی توسعه یافتند (Schoen, 1988; Jericevic et al., 2000). ایده پایه ای، استفاده از داده های اندازه گیری شده است در موقعیتهای ابزاری مختلف توسط یک ابزار نمودارگیری مقاومت ویژه آرایه ای ساده، تحت قواعد عمل به مثل و انطباق، برای ساخت یک سری از ابزار ترکیبی با عمق های مختلف بررسی. الکتروودهای این ابزار ترکیبی در یک چنین روشی طوری چیده شده اند که جریانهای ترکیبی آنها برای جاری شدن به طور عمودی در دیواره چاه و سازند با فشار جریان تزریق می شوند، بدلیل بدست آوردن نتیجه یکسان به مانند ابزار سخت افزاری متمرکز شده در کمینه کردن اثرهای لایه های مجاور. فایده متمرکز کردن نرم افزاری این است که آن اجازه می دهد مجمع ابزار ترکیبی با عمق های مختلف بررسی از یک مجموعه ساده نمودارهای جانبی آرایه ای.

ابزار SFR منتشر شده رایج SFR ۱۰، ۲۰، ۳۰، ۴۰، ۵۰ هستند. متناظر با عمق های بررسی ۱۰، ۱۲، ۱۴، ۱۶، ۲۰، ۳۰، ۴۰، ۵۰ اینچ (یا ۲۵/۵، ۳۰/۴۸، ۳۵/۵۶، ۴۰/۶۴، ۵۰/۸، ۷۶/۲).

۱۰/۱۶، و ۱۲۷ سانتیمتر) (ضمیمه A). همه این ابزار متمرکز شده ترکیبی متناسب هستند با عمق اندازه گیری و در یک طریق مشابه بوجود آمده اند. ما از S FR20 برای توضیح اینکه چطور ابزار ترکیبی تولید شده اند استفاده می کنیم. هدف ما ساختن یک ابزار ترکیبی است که در شکل 2a رسم شده ، که شامل سه الکتروود جریان و چهار الکتروود پتانسیل است. ابزار متناسب است با الکتروود جریان مرکزی ، که یک جریان اصلی C0 به سازند می فرستد. دو الکتروودی که در دو انتهای ابزار ترکیبی هستند جریانهای متمرکز C1 و C2 را به سازند تزریق می کنند. تمام جریانها توسط C0 هنجار شده اند. چهار الکتروودهای پتانسیل ترکیبی ، ولتاژهای ترکیبی VM1 ، VM2 ، VM3 و VN4 را اندازه می گیرند. ما جریانهای متمرکز را همچو وضعیت تامل متمرکز تنظیم می کنیم:

(۱-۷)

$$VM1 = VM2$$

$$VM3 = VN4$$

بنابراین ، C0 به طور عمودی در دیواره گمانه و سازند جریان پیدا می کند . برای ساخت این ابزار SFR20 ترکیبی ، دوازده ولتاژ ترکیبی در نه عمق نمودارگیری مختلف در شکل (۷-۵-b-d) لازم است. اول ، ولتاژهای ترکیبی در شکل (۷-۵-b) توسط چهار اندازه گیری ولتاژ HDLL در سه عمق نمودارگیری مختلف نشان داده شده در شکل (۷-۶-a) ساخته می شوند. ولتاژهای V1 و V2 قاعده عمل متقابل منتج شده است ،

(۲-۷)

$$V1(z) = V1(z-x)$$

$$V2(z) = V2(z-y)$$

که Z عمق اندازه گیری است و x و y فواصل از عمق اندازه گیری به الکتروودهای ولتاژهای داخلی و خارجی هستند. دوم ، ولتاژهای ترکیبی در شکل 2c مستقیماً از چهار اندازه گیری ولتاژ HDLL در عمق نمودار گیری (z-L/2) بدست آمده اند. که L دازای ابزار ترکیبی است. سوم ، ولتاژهای ترکیبی در شکل (۷-۵-b) توسط چهار اندازه گیری ولتاژ HDLL در چهار عمق نمودارگیری مختلف (شکل ۷-۶-b) با استفاده از قاعده عمل متقابل ساخته شده اند :

(۳-۷)

$$V3(z)=V3(z+y)$$

$$V4(z)=V4(z+x)$$

$$V5(z)=V5(z-x)$$

$$V6(z)=V6(z-y)$$

سرانجام ، ابزار ترکیبی SFR20 در شکل (۴-۷) بوسیله استفاده از قاعده انطباق از تئوری پتانسیل ساخته می شود :

(۴-۷)

$$V2+C1.V3+C2.V6=V M1$$

$$V1+C1.V4+C2.V5=VM2$$

$$V1+C1.V5+C2.V4=VM3$$

$$V2+C1.V6+C2.V3=VM4$$

که " نشان دهنده ضرب اسکالر است.

از معادلات (۱-۷) و (۴-۷) ، ما می توانیم برای C1 و C2 حل کنیم به خوبی چهار اندازه گیری پتانسیل ترکیبی VM1 تا VM4 . این سیستم از معادلات خوب مطرح شده است و می تواند بوسیله استفاده از حل کننده - های سیستم خطی استاندارد دوباره حل شود. ما سپس پتانسیل ها را به مقاومت ویژه ظاهری تبدیل می کنیم ، یا داده های SFR :

(۵-۷)

$$R_a=(K_n/4C0)(VM1+VM2+VM3+VM4)$$

که K_n عامل هندسی برای $SFR_n(n=10,12,14,16,20,30,40,50)$ است ، که از میان مدل سازی در گمانه ۸-اینچ (۰/۲ متر) بدست آمده است ، ۱۰۰ ohm-m مقاومت ویژه سازند ، و ۱ ohm-m مقاومت ویژه گل (ضمیمه A).

بعد از فرایند متمرکز سازی ، منحنی SFR نتیجه خیلی کم به اثرات بستر لایه های مجاور بستگی دارد و می تواند تقریب زده شود به همان اندازه از یک زمین یک بعدی استوانه ای که شامل یک گمانه ، ناحیه آغشته و بکر می باشد. برای کاهش بیشتر شمار پارامترهای آزاد در پرورش شبکه

عصبی ، ما یک تصحیح گمانه روی داده های SFR انجام می دهیم(ضمیمه B). مدل زمین ساده شده بنابراین توسط تنها سه پارامتر مدل در هر لایه شرح داده می شود: R_{X0} ، R_{X0} و L_{X0} . شبکه های عصبی به طور مؤثر می توانند این مدل های ساده شده را بکارگیرند.

۷-۲-۳-۲- شبیه سازی شبکه عصبی

هدف ما پرورش یک شبکه عصبی است برای آمیختن داده های SFR گمانه تصحیح شده ، داده شده به عنوان ورودی یک مدل یک بعدی استوانه ای . ما از یک شبکه عصبی گسترش (توزیع) برگشتی (BP) در کاربردهایمان استفاده می کنیم. شبکه ما شامل سه لایه (لایه های ورودی ، پنهانی (میانی) ، و خروجی) است با سه گره ورودی ، سیزده گره میانی ، و یک گره خروجی . میزان یادگیری برای لایه های میانی و خروجی بترتیب ۰/۰۰۶ و ۰/۰۰۳ هستند. اندازه حرکتها برای لایه های میانی و خروجی همه ۰/۲ هستند. اساتید (۱۹۹۳) توضیح می دهند که چطور این پارامترها را انتخاب کنیم.

برای آموزش شبکه ، ما اول طیف درونی از مدل زمین را برای ایجاد نمونه های پرورش از میان مدل سازی عددی نمونه گیری می کنیم(de Witte and Gould, 1959). مقاومت ویژه سازند و آغشته در تولید نمونه های پرورش استفاده شده از حدود ۰/۰۱ تا ۱۰۰۰ h m-m با چهار نمونه در دهه لگاریتمی ، و تغییرات طول آغستگی از ۰/۰۳ تا ۱/۲ متر. این پارامترهای مدل در لایه ورودی شبکه عصبی به عنوان محرک به کار برده شده اند. پاسخ های HDLL یک بعدی شبیه سازی شده سپس برای تهیه پاسخ های SFR به طور عددی متمرکز شده اند. این تمرکز عددی ، داده های اثر آزاد گمانه و لایه مجاور ، برای لایه خروجی شبکه عصبی بکار برده شده اند برای نظارت پرورش شبکه عصبی. پرورش زمانی متوقف می شود که خطای میانگین در لایه خروجی کمتر از ۱٪ است. یک مجموعی از ۱۳۶۳۰ تکرار برای هر دوره پرورش انجام شده اند. بهترین خطای r ms برای SFR ۱۰، ۲۰، ۳۰، ۴۰، ۵۰، ۱۲، ۱۴، ۱۶، ۲۰، ۳۰، ۴۰، ۵۰ به ترتیب ۰/۰۱۹٪، ۰/۰۲٪، ۰/۰۲٪، ۰/۰۲٪، ۰/۰۲٪، ۰/۰۲٪، ۰/۰۳٪ هستند.

فرایند پرورش ، یک مجموعه از ضرایب وزن برای استفاده در شبیه سازی پاسخ های ابزار HDLL تولید می کند. قبل از اینکه ما بتوانیم از شبکه عصبی پرورش داده شده استفاده کنیم ، معتبر سازی بر روی نمونه آزمایشی که با نمونه پرورشی مشترک نیستند لازم است. اگر خطای

ناجور بین داده های خروجی از شبکه عصبی و داده های آزمایش ، محاسبه شده از الگوریتم های مدل سازی مستقیم عددی ، زیر سطح آستانه دلخواه باشد ، آنگاه شبکه عصبی معتبر خواهد بود و می تواند برای شبیه سازی داده های HDLL استفاده شود.

شبکه های عصبی نمی توانند با شبیه سازی عددی عرفی در دوره های درست مطابق شوند. به همان اندازه افزایش های میزان دینامیک ، یک دقت شبکه عموماً کاهش داده می شود. به هر حال هدف ما روی هم رفته جایگزینی شبیه سازی عددی عرفی نیست. ترجیحاً ، ما در جستجوی یک مدل سازی مستقیم تقریبی سریع با خطای بقدر کافی کوچک هستیم برای استفاده در معکوس سازی مان. برای کاربردمان ، این خطای ناجور داده های دلخواه دارای ۵٪ خطای rms است ، که داده های پایانی ناجور را منعکس می سازد که وارونه سازی های دو بعدی با استفاده از شبیه سازی عددی مرسوم می توانند برای بیشتر موارد داده های زمینی بدست آورند. خطاهای rms در آزمایشات تاییدی برای SFR ۱۰، ۳۰، ۲۰، ۱۶، ۱۴، ۱۲، ۴۰، و ۵۰ بترتیب ۱/۴٪، ۴/۷۵٪، ۴/۲٪، ۳/۹٪، ۳/۲٪، ۴/۱۰٪، و ۴/۷٪ هستند. این خطاها کوچکتر از ۵٪ خطای rms سطح ناجور دلخواه هستند ؛ بنابراین ، شبکه عصبی معتبر می شود. ما سرعت های محاسبه ای مدل سازی مستقیم شبکه عصبی مان و مدل سازی مستقیم مرسوم شرح شده در (Fabris(1991) با استفاده از یک مدل یک بعدی مقایسه می کنیم. شبکه عصبی ۳/۵۸ ثانیه بر ای محاسبه ۳۰ متر از داده با ۸۰۱ نمونه وقت لازم دارد ، در حالیکه الگوریتم مرسوم ۲۸ دقیقه و ۵ ثانیه برای بدست آوردن نتیجه مشابه زمان می برد. بنابراین ، الگوریتم شبکه عصبی مان حدوداً ۴۷۰ مرتبه سریعتر است.

ما سپس شبکه عصبی پرورش داده شده را به عنوان اسباب مدل سازی مستقیم در یک الگوریتم شرح داده شده توسط Zhang et al.(1999a) بکار می بریم. چون مدل سازی مستقیم تنها یک قسمتی از ثانیه وقت می گیرد ، معکوس سازی ما می تواند در یک زمان واقعی انجام شود. کیفیت نتایج معکوس سازی به این بستگی دارد که چقدر خوب کاهش بعدی انجام شده باشد. وقتی که کنتراست مقاومت ویژه بالا است و ضخامت بستر نازک است ، متمرکز سازی عددی ممکن نیست به طور کامل اثرات دوبعدی را برطرف کند. این بقایای اثرات دوبعدی در حال حاضر در نتایج وارونه سازیمان در شکلی از نوسان فرکانس بالا دامنه پایین هستند. برای چیره شدن این مسئله ، ما این خصیصه های کوچک خیالی در نتایج وارونه سازی را با یک فیلتر میانگین ۱۱ نقطه ای خارج از صافی می کنیم.

۷-۲-۴- مثالهای عددی

ما مدل سازی مستقیم شبکه عصبی بنیان شده مان را داخل می کنیم در یک الگوریتم وارونه سازی شرح داده شده توسط Zhang et al.(1999a). ما حساسیت ها را توسط روش تفاضل محدود محاسبه می کنیم. ما سپس این الگوریتم وارونه سازی سریع را برای داده های ترکیبی تولید شده از مدل شکل (۷-۷) بکار می بریم. این مدل شامل یک رشته از لایه های نفوذپذیر و نفوذناپذیر می باشد. دو لایه نفوذپذیر با مایع های هادی آغشته شده اند. مقاومت ویژه گل ۰/۰۵ ohm-m است، و قطر گمانه ۸ اینچ (۰/۲ متر) می باشد. داده های دوبعدی خام وجود دو ناحیه آغشته مختلف را نشان می دهند اما در نشان دادن بطور دقیق مقاومت ویژه سازند درست اشتباه کردند. این وقتی نتیجه می دهد که ابزار HDLL نامتمرکز شده باشند و، از اینرو، داده های HDLL تحت تاثیر بستر لایه های مجاور و اثرات گمانه هستند.

بعد از متمرکز سازی عددی و تصحیح گمانه، اثرات لایه های مجاور و گمانه به طور زیادی کاهش می یابند (شکل ۷-۸). منحنی های SFR به طور واضحی دو ناحیه آغشته شده هادی را نشان می دهد. دامنه منحنیهای SFR، به هر حال، هنوز دور از مدل صحیح هستند. همچنین تعدادی محصول های مصنوعی در نتیجه گیری داده های SFR از رفع اثر بستر لایه مجاور ناقص وجود دارد. تنها بالای ضخامت، لایه های آغشته نشده، SFR مدل زمین صحیح را نشان خواهد داد.

برای معتبر ساختن نتیجه متمرکز سازی عددی نرم افزار، ما همچنین پاسخ ابزار D LL سخت افزاری متمرکز شده را بیشتر از مدل مشابه محاسبه می کنیم که از مدل سازی مستقیم عددی دوبعدی استفاده می کند و پاسخ DLL دوبعدی را با منحنیهای SFR متمرکز شده عددی در شکل (۷-۸) رسم می کند. ابزار مجازی ما یک شکل متفاوت از ابزار D LL دارد. اینرو، ما نمی توانیم پاسخهای DLL را با منحنیهای SFR در یک روش محض مقایسه کنیم. به هر حال مقایسه شکلهای پاسخ DLL سخت افزار متمرکز شده با داده های SFR نرم افزار متمرکز مفید است. برای مدل تحت مطالعه، داده های DLL یک شکل مشابه دارند، همچنین داده های SFR. منحنی عمیق DLL به طور نزدیکی از منحنی SFR30 پیروی می کند، و منحنی کم عمق DLL حدوداً همانند SFR12 می خواند. مصنوعات اطراف مرزهای بستر نشان داده شده در منحنی های SFR

همچنین در پاسخ DLL نیز وجود دارند. این مطالعه مقایسه ای نشان می دهد که متمرکز سازی عددی می تواند نتایج مشابه به متمرکز کردن سخت افزاری بدست دهد.

سپس، ما منحنیهای SFR در شکل (۷-۸) را معکوس می کنیم. در وارونه سازی، ما از SFR50 (منحنی عمیقتر) استفاده می کنیم به عنوان اولین مقدار برای R_t و SFR10 (منحنی کم عمقتر) به عنوان اولین مدل برای R_{x0} . مقدار اولیه برای L_{x0} یک مقدار ثابت ۱۲ اینچی (۰/۳ متر) است. شکل (۷-۹) مدل‌های مقاومت ویژه صحیح و بهبودیافته از وارونه سازی نشان می دهد. R_{x0} بهبود یافته نمایش خوب از مدل صحیح است، و همچنین دامنه R_t بر روی SFR50 بهبود یافته است. علاوه، وارونه سازی یک تخمین معقول از طول آغشته شده تهیه می کند. ناجوری داده ها حدود یک ۲٪ خطای rms است برای همه عمقها. این مثال داده های ترکیبی وقتی آنرا نشان می دهد که طول آغستگی مطابق قدرت تفکیک پذیری ابزار HDLL (بین ۰/۱ و ۱ متر) باشد، الگوریتم وارونه سازی سریع ما می تواند اطلاعات مفیدی درباره ساختار مقاومت ویژه تهیه کند.

ما الگوریتم وارونه سازیمان را روی یک مجموعه داده های زمین از هلند آزمایش می کنیم. فاصله تحت مطالعه یک زاویه انحراف 78° دارد. مقاومت ویژه گل ۰/۰۲۷ ohm-m است و نسبتاً هادی است (بین ۰/۵ و ۵۰ ohm-m). شکل (۷-۱۰) داده های SFR را همراه با داده های پرتو گاما و قطرسنجی نشان می دهد. برای نشان دادن منظور، تنها منحنی های SFR10، 16 و 50 اینجا ارائه شده اند. داده های SFR چندین ناحیه آغشته شده را نشان می دهد. منحنی قطرسنجی وضعیتهای گمانه خوب را نشان می دهد. مقادیر اولیه برای R_t و R_{x0} منحنی های عمیق (SFR5) و کم عمق (SFR10) هستند. مقدار اولیه برای L_{x0} ، ۰/۴ متر است. منحنی پرتو گاما در وارونه سازی برای نشان دادن آغستگی و تحمیل کردن L_{x0} در هر تکرار بکار برده می شود. در ناحیه شیل، L_{x0} به صفر رسانده شده است. سر حد در ۰/۴۰ مقدار پیک منحنی پرتو گاما قرار گرفته بود، برطبق سازند فیزیک نفت (پتروفیزیک) موجود. همه هشت منحنی SFR در وارونه سازی استفاده شده اند. وارونه سازی می تواند ۶۰ متر از داده ها را در هر دقیقه روی ایستگاه کارمان پردازش کند، که خیلی سریعتر از سرعت نمودارگیری است (حدود ۱۰ متر در دقیقه).

شکل (۷-۱۱) مدل مقاومت ویژه بهبودیافته از وارونه سازی را نشان می دهد. جایی که سازند متراکم است، جدایی بین منحنی های SFR کم عمق و عمیق کوچک است و نتیجه وارونه سازی به طور نزدیکی از منحنی های SFR پیروی می کند. معکوس سازی چندین ناحیه آغشته

شده را جایی که مقادیر مقاومت ویژه بهبود یافته ۱- ۲۰ ohm-m بودند، نمایان می کند. معکوس سازی همچنین اطلاعاتی درباره عمق آغستگی ارائه می کند، که منحنی های S FR نمی توانند تهیه کنند. داده ناچور برای هر سطح عمق متفاوت بود اما معمولاً خطای rms بین ۲-۴٪ بود.

برای معتبر ساختن نتیجه وارونه سازی مان، ما یک وارونه سازی سه بعدی روی مجموعه داده های یکسان با استفاده از الگوریتم شرح داده شده در (Lugao et al. (1999) انجام دادیم. مقادیر اولیه برای R_i ، R_{x0} و L_{x0} به ترتیب SFR50، SFR10، و 12 اینچ (۰/۳ متر) بودند. وارونه سازی سه بعدی ۳۶ ساعت در همان ایستگاه کار برای همگرا کردن به مدل نهایی در شکل (۷-۱۲) زمان می برد. نتایج وارونه سازی سه بعدی یک R_i مقاومتی زیاد نشان داد نسبت به وارونه سازی بلادرنگ روی لایه های نازک، e.g.، اطراف لایه ۹۸۰۰ فوت (۲۹۹۰ متر)، و دوباره بدست آوردن یک ناحیه آغشته کمی عیقتر. خصوصیات کلی دو ساختار مقاومت ویژه سازند بازیافت شده، به هر حال، خیلی شبیه بودند. مقاومت مخصوص های ناحیه آغستگی دوباره بدست آمده از دو وارونه سازی به طور مساوی برابر بودند. این مثال نشان می دهد که الگوریتم معکوس سازی شبه دوبعدی ما می تواند به طور مؤثر داده های زمین را به کار برد و اطلاعات مفیدی به طور آبی و در محل چاه تهیه کند.

۷-۲-۵- نتیجه

ما یک شبکه عصبی مبنا، الگوریتم وارونه سازی شبه دوبعدی سند می کنیم که میتواند مجموعه های داده ای نمودارگیری مقاومت ویژه آرایه ای بزرگ را به طور آبی بکار گیرد. این الگوریتم اطلاعات مفیدی درباره ساختار مقاومت مخصوص سازند در محل چاه تهیه می کند. ما مسأله وارون دو بعدی را به یک مسأله یک بعدی توسط متمرکز سازی عددی کاهش می دهیم. داده های S FR سپس تصحیح گمانه و وارونه می شوند با استفاده از الگوریتم وارونه سازی شبکه عصبی مبنای مان. مدل سازی مستقیم شبکه عصبی ما چند صد برابر سریعتر از الگوریتمهای مدل سازی مستقیم یک بعدی مرسوم است. آزمایشهای داده های ترکیبی به خوبی آزمایش ما نشان می دهد که این الگوریتم وارونه سازی می تواند اطلاعات قابل قبول و مفیدی وقتی که عمق آغستگی بین ۰/۱ و ۱ متر باشد تولید کند. کیفیت نتایج وارونه سازی بستگی به متمرکز سازی عددی دارد. وقتی که لایه های نازک و تباین بالا وجود دارد و آثار دو بعدی برای به طور مؤثر

موقوف شدن خیلی قوی باشند، موقع تعبیر و تفسیر نتایج وارونه سازی احتیاط لازم است زیرا داده های SFR تحت این موقعیت ها کمتر متمرکز شده هستند. وارونه سازی شبه دوبعدی بلادرنگ ما یک تفسیر بهتر از ساختار مقاومت مخصوص نسبت به منحنی های SFR تهیه می کند و می تواند مستقیماً در تفسیر کامل یا به عنوان مدل های اولیه در وارونه سازی های دو بعدی مابعد استفاده شود.

ضمیمه A

عمق بررسی

عمق بررسی هر SFR بر پایه مفهوم یک فاکتور هندسی شعاعی مشتق شده بنا شده است. این فاکتورها برای ابزار مقاومت ویژه برقی فاکتورهای هندسی کاذب شعاعی مشتق شده نامیده شده اند برای تاکید بر این که آنها ثابت نیستند اما به طور زیاد بستگی به تباین R_t/R_m و ضخامت بستر دارد. عامل هندسی کاذب HDLL برای هر SFR می تواند به عنوان کمک تعریف شود برای پاسخ SFR کلی دریافت شده از یک سازند یکنواخت (i.e. یک بستر ضخیم) احاطه شده بین یک استوانه به عنوان اندازه افزایشهای گمانه از صفر تا بینهایت.

یک روش استاندارد (Flaum and Theys, 1991) برای تعیین عمق بررسی یک ابزار مقاومت مخصوص، تعریف یک فاکتور هندسی کاذب $J(r_i)$ به عنوان یک تابع از شعاع آغستگی r_i است که مقاومت ویژه ظاهری R_a را برحسب پارامترهای سازند R_{x0} و R_t بیان می کند:

$$Ra = J(r_i)R_{x0} + [1 - J(r_i)]R_t \quad (A-1)$$

که

$$J(r_i) = \frac{R_t - R_a}{R_t - R_{x0}} \quad (A-2)$$

اینجا، r_i مجموع شعاع گمانه و L_{x0} است. این فرمول یک خانواده از منحنی ها را تعریف می کند، که به R_t/R_{x0} بستگی دارد. این منحنی ها می توانند با استفاده از مدل سازی عددی محاسبه شده باشند. برای یک R_t/R_{x0} ، عمق بررسی به عنوان مقداری از L_{x0} تعریف می شود که $J=0.5$ مربوط به آن است. برای ابزار هندسی داده شده، عمق بررسی، یک تابع از تباین مقاومت ویژه است

بین ناحیه آغشته و تمیز. معمولاً عمق بررسی برای یک تباین مقاومت ویژه مشخص بیان شده است. برای پاسخهای SFR عمق بررسی از یک میانگین مقادیر داده شده توسط R_t/R_{x0} برابر با ۱۰ و ۱۰۰ تعیین شده است. عمق بررسی برای هر SFR اکثراً به نسبت R_t/R_{x0} بستگی دارد. وقتی که نسبت افزایش می یابد، عمق بررسی برای SFR به طور مهبجی کاهش می یابد. بنابراین عمق بررسی برای هر SFR بایستی استفاده شود تنها به عنوان یک راهنما برای مقایسه پاسخهای SFR، که به عنوان عدد مطلق نیست.

نمودار فاکتور شبه هندسی برای SFR20 به عنوان یک مثال در شکل A-1 محاسبه شده است. مدل یک بستر ضخیم محدود ساده با دو ناحیه شعاعی دارد. مقاومت ویژه ناحیه داخلی و خارجی بترتیب توسط R_{x0} و R_t مشخص شده اند. مدلسازی عددی برای HDLL در این سازند با تباین مقاومت ویژه $R_t/R_{x0}=0.1, 0.5, 2.0, 5.0, 10, 50, 100, 500, 1000, 5000, 10000$ انجام شده است. تنوع عمق آغشتگی از $L_{x0}=0$ تا ۱۰۰ اینچ پاسخهای SFR گردآوری شده بودند از اندازه گیری های یک الکتروود و سپس هنجار شده با رابطه R_t در یک سازند همگن.

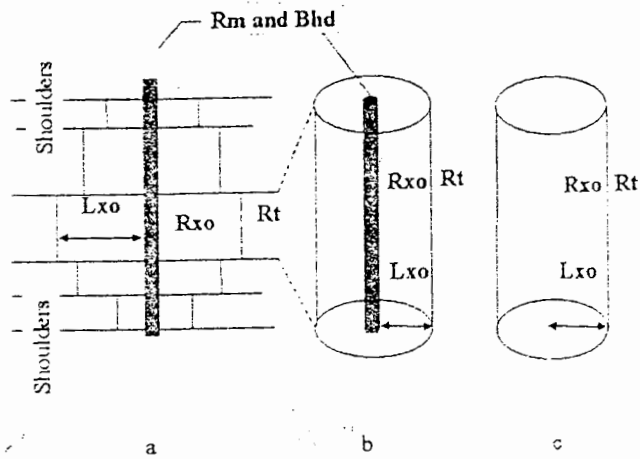
ضمیمه B

تصحیح گمانه برای داده های SFR

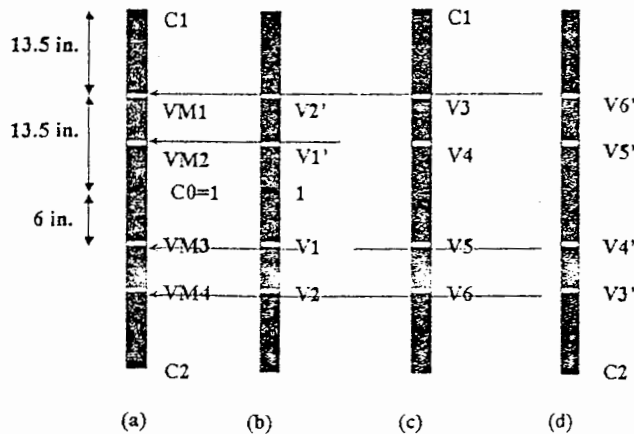
تصحیحات گمانه برای ابزار HDLL با استفاده از شبیه سازی عددی یک بعدی محاسبه شده اند. مدل شامل یک گمانه و یک سازند ضخیم محدود است (Western Atlas, 1985, 7-15). شبیه سازی عددی ولتاژهای الکتروود HDLL را تهیه می کند، که با تولید پاسخهای SFR ترکیب شده اند. ما سپس از معادله (۷-۳) استفاده می کنیم برای تعریف فاکتور سودمند ابزاری مجازی K روی یک مدل با یک گمانه ۸ اینچ (۰/۲ متر) و $R_t/R_m=100$ ، که R_m مقاومت ویژه مایع گمانه است. برای مدل های متفاوت از این مدل منبع، مقاومت ویژه ظاهری برابر با R_t نیست و تصحیح گمانه لازم است.

ما برای شرح روش از SFR20 استفاده می کنیم (شکل B-1). در نمودار تصحیح، مقاومت ویژه ظاهری به عنوان R_{HDLL} مشخص می شود. چون ما معمولاً پیشاپیش R_t را نمی دانیم، ما R_{HDLL}/R_m را به عنوان یک تابع از R_t/R_m برای ساختن نمودار تصحیح گمانه محاسبه می کنیم.

مقاومت ویژه ظاهری تصحیح شده R_{corr} است ، که بایستی با R_t روی مدل‌های یک بعدی آغشته نشده برابر باشد . نمودار ، نسبت R_{corr}/R_{HDLL} را به عنوان یک تابع از R_{HDLL}/R_m نشان می دهد. در یک سازند واقعی تصحیحات دیگر (برای مثال تصحیح لایه های مجاور ، تصحیح آغستگی ، و تصحیح انیزوتروپی) ممکن است برای برابری R_t با R_{corr} لازم باشد. [۴۰]

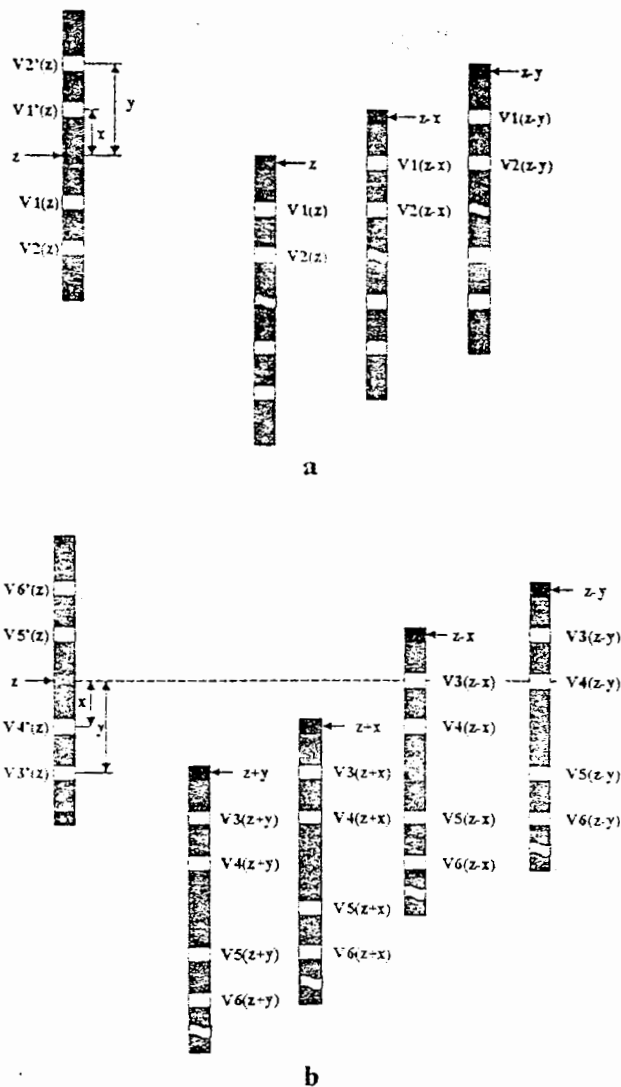


شکل ۴-۷: (a) مدل دوبعدی زمین قبل از متمرکز کردن عددی عبارت می باشد از R_m ، مقاومت ویژه گل؛ B_{hd} ، قطر گمانه؛ R_t ، مقاومت ویژه سازند؛ R_{xo} ، مقاومت ویژه ناحیه آغشته؛ L_{xo} ، طول آغستگی؛ و لایه های مجاور. (b) مدل زمین بعد از متمرکز کردن عددی که عبارت می باشد از R_m ، R_{xo} ، R_t ، B_{hd} و L_{xo} . (c) مدل ساده شده یک بعدی زمین که در پرورش شبکه عصبی استفاده می شود، تنها شامل سه پارامتر مدل است: R_t ، R_{xo} و L_{xo} .

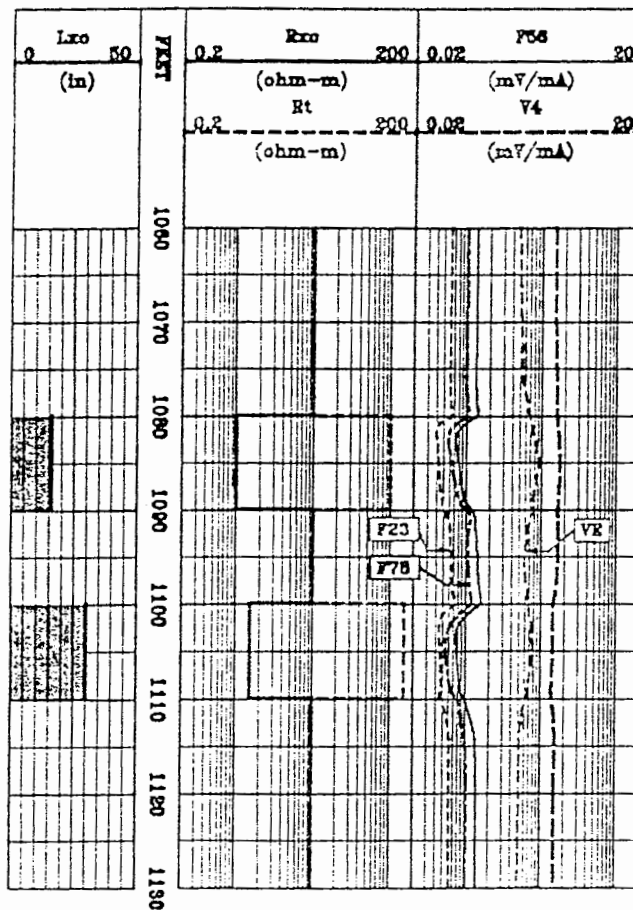


شکل ۵-۷: ساختار SFR20. (a) ابزار SFR20 ساخته شده (متناسب با C0). (b) وضعیت ابزار HDLL برای V_1, V_2, V_1', V_2' ؛ (c) V_3, V_4, V_5, V_6 ؛ (d) V_3, V_4, V_5, V_6 و V_3', V_4', V_5', V_6' ؛ و C1 و C2 جریانهای هنجار شده ابزار مجازی هستند.

شکل ۶-۷: وضعیت ابزار HDLL استفاده شده در شکل 2b و 2d. (a) اندازه گیری های ابزار HDLL برای ساختن اندازه گیری های ترکیبی در شکل 2b لازم شده. (b) اندازه گیری های ابزار HDLL استفاده شده در شکل (۷-۵ d): عمق ضبط توسط z نشان داده شده است. فواصل از z تا دو الکتروود اندازه گیری به ترتیب x و y هستند.

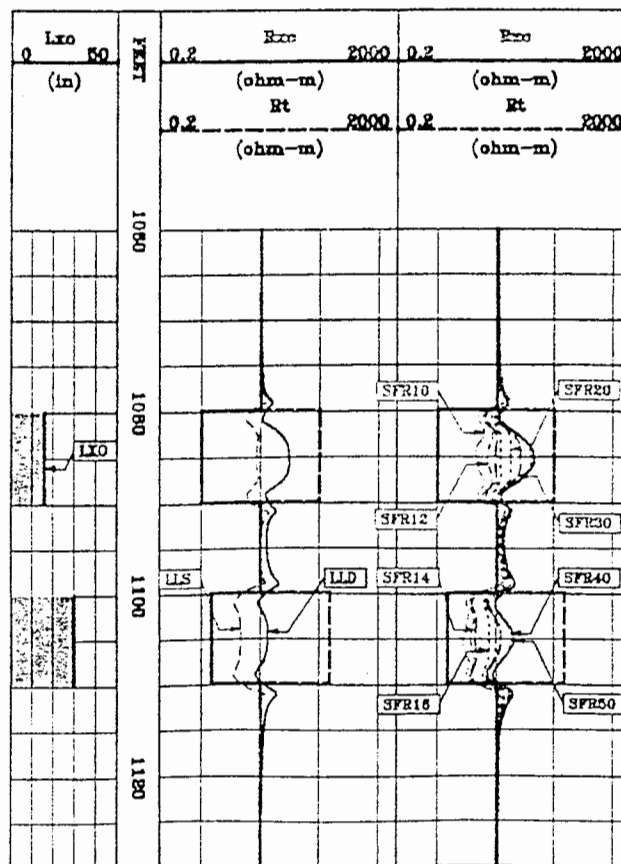


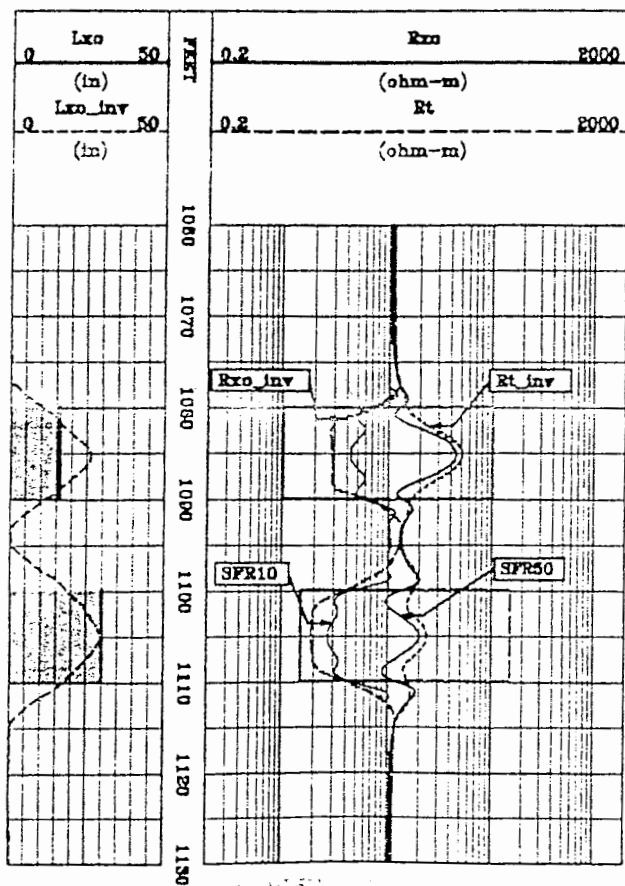
شکل ۷-۷: مدل مصنوعی و داده های خام برای خواندن آسان همه داده ها نشان داده نشده اند. مقاومت ویژه گل ۰/۰۵ ohm-m، و قطر گمانه ۰/۲ متر است. F23 اولین اندازه گیری تفاوت بین الکترودهای دوم و سوم است و مشابه F56 و F78 است. V4 و VE پتانسیلهای اندازه گیری شده در الکترودهای 4 هستند (نگاه کنید به شکل ۷-۴).



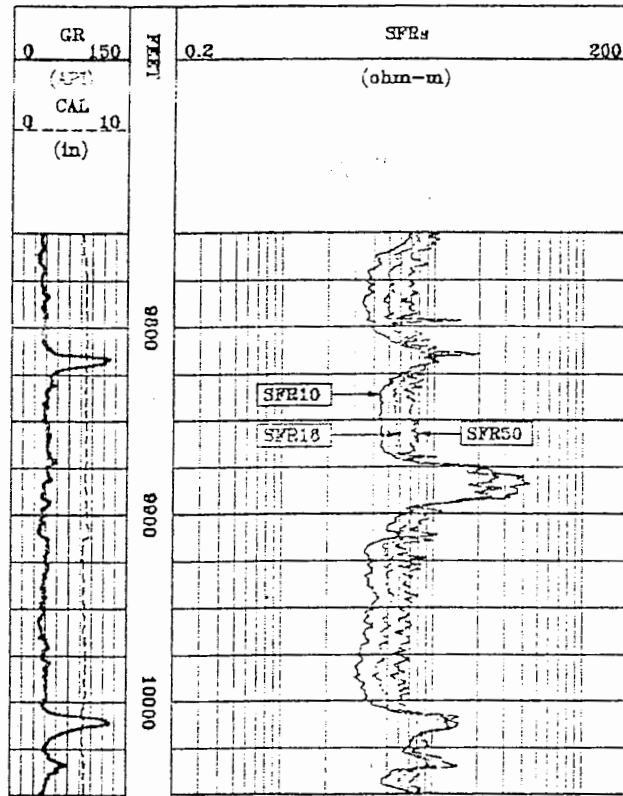
شکل ۷-۸: مقایسه منحنی های SFR و داده های LLS LLD.DLL منحنیهای عمیق و کم عمق ابزار

DLL هستند. R_{xo} ، R_t و پارامترهای مدل صحیح هستند

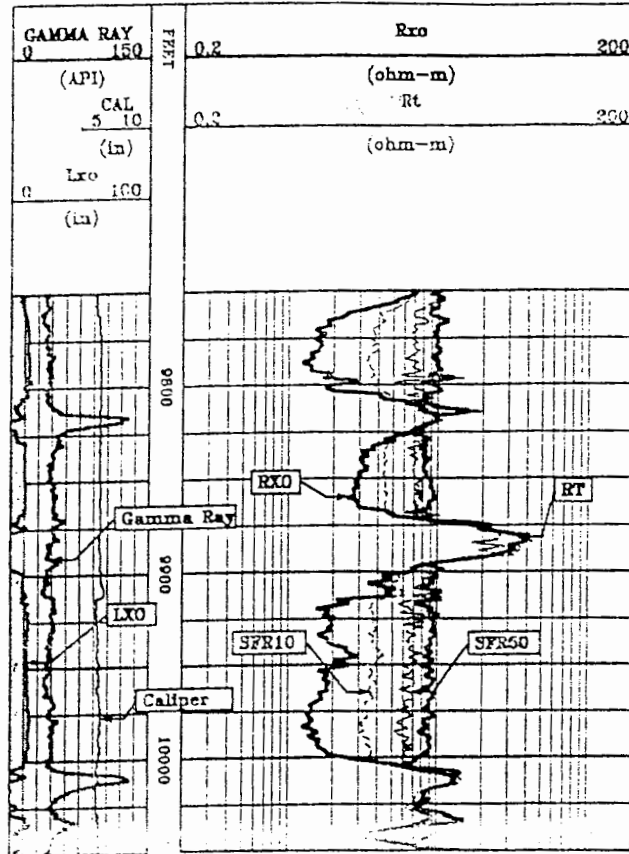




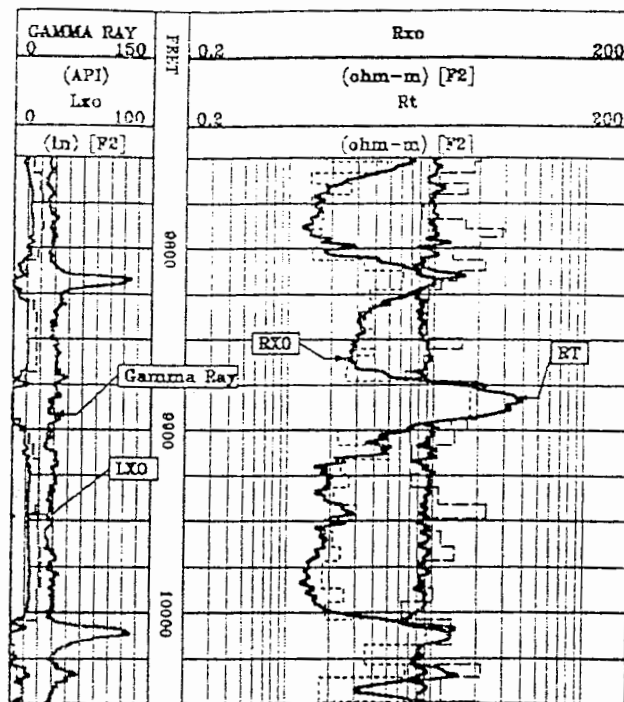
شکل ۷-۹: نتیجه وارونه سازی برای داده SFR در شکل (۷-۸). Lxo_inv ، Rxo_inv ، Rt_inv نتایج وارونه سازی هستند. SFR50 و SFR10 برای مقایسه ارائه شده اند.



شکل ۷-۱۰: منحنی های SFR تولید شده از مجموعه داده های زمین از نترلند (Netherlands). GR و CAL بترتیب منحنیهای پرتو گاما و قطر سنجی را نشان می دهند.



شکل ۷-۱۱: مثال داده های زمین از Netherlands، R_{xo} ، R_t و نتایج وارونه سازی هستند با استفاده از الگوریتمهای وارونه سازی آنی ما. CAL منحنی قطرسنجی را مشخص میکند. SFR10 و SFR50 برای مقایسه هستند.



شکل ۷-۱۲: مقایسه نتایج وارونه سازی سه بعدی و شبه دوبعدی برای مجموعه داده های زمینی در شکل (۷-۱۰). (چپ) منحنی های پرتو گاما و L_{xo} منحنی سایه دار L_{xo} است از وارونه سازی شبه دو بعدی؛ خط تیره L_{xo} از وارونه سازی سه بعدی را نشان می دهد. (راست) R_{xo} و R_t بهبود یافته از وارونه سازی شبه دوبعدی و سه بعدی. خطهای نقطه چین و تیره R_{xo} و R_t از وارونه سازی دو بعدی را نشان می دهد.

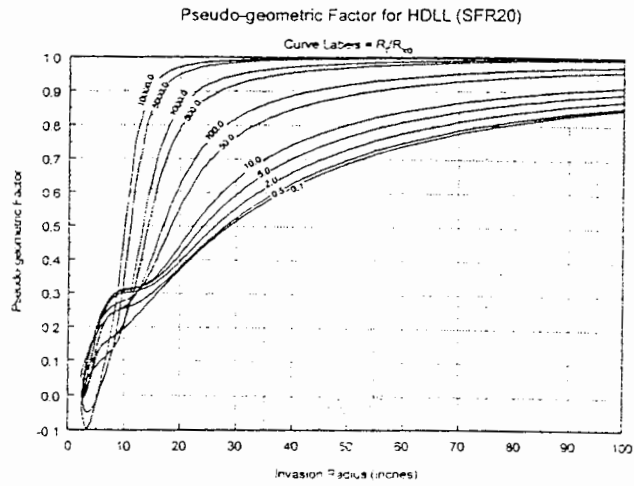
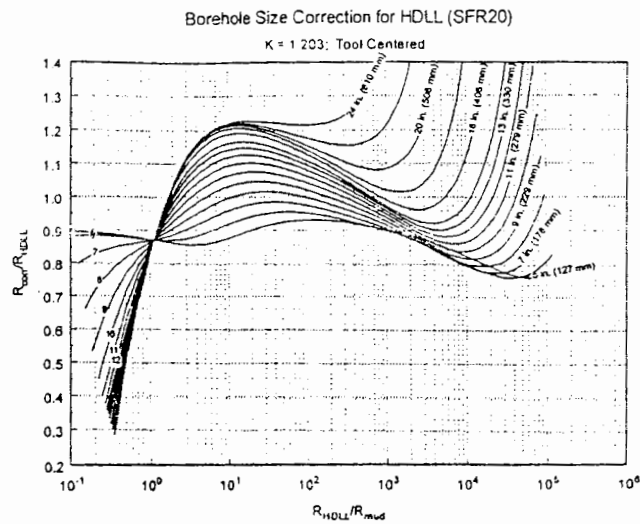


FIG. A-1. Pseudogeometric factor chart for SFR20.

شکل A-1: نمودار فاکتور شبه هندسی برای SFR20



شکل B-1: نمودار تصحیح اثر چاه برای SFR20. R_m مقاومت ویژه گل است، K فاکتور وسیله برای SFR20، R_{HDLL} مقاومت ویژه ظاهری SFR20 قبل از تصحیح چاه است، در حالی که داده‌های R_{corr} بعد از تصحیح چاه است.

فصل هشتم

قوتها و ضعفهای شبکه‌های عصبی

شبکه‌های عصبی راه حل همه مسائل محاسبه‌ای نیستند. آن‌ها برخی محدودیت‌ها دارند که بایستی در تصمیم‌گیری اینکه آیا از این تکنولوژی استفاده شود یا نه، در نظر گرفته شود. محدودیت‌های رایج و مشکلات با شبکه‌ای عصبی شامل نکات زیر است:

- سرعت یادگیری پایین. برای مسائلی که یک ساختار شبکه‌ای پیچیده و برگ نیاز دارند با یک تعداد زیادی از نمونه‌های آموزشی دارند، زمان مورد نیاز برای آموزش شبکه می‌تواند بیش از اندازه طولانی باشد، 10^4 - 10^5 تکرار دوره آموزشی، دقیق یا ساعات ثابت، برای مسائل پیچیده نرمال یک چیز مشترک و عمومی است. این مفید بودن شبکه‌های عصبی را برای پردازش‌هایی که پاسخ زمانی واقعی نیاز دارند، مانند تفسیر فعل و انفعالی، محدود می‌کند.

- بدام افتادن در مینیمم موضعی. آموزش یک شبکه عصبی پس انتشار یک جستجو برای مجموعه‌ای از وزنهای اتصال داخلی است که اختلاف خطای بین مقادیر نرون خروجی مطلوب و واقعی را مینیمم کند. استراتژی جستجو از یک تکنیک نزول شیب برای پیدا کردن وزنهای وابسته به این خطای مینیمم استفاده می‌کند. اینجا راهی برای داشتن اینکه آیا مینیمم پیدا شده توسط این روند یک مینیمم مطلق است یا یک مینیمم موضعی (که می‌تواند بزرگتر از مینیمم مطلق باشد) وجود ندارد. یک شبکه که در تله یک مینیمم موضعی می‌افتد نمی‌تواند یک راه حل بهینه تهیه کند. این مساله بدام افتادن متشابه با سعی کردن برای یافتن کمترین ارتفاع در آمریکای شمالی و برای کسی که از هر موقعیت دلخواه شروع می‌کند، است. یک استراتژی جستجوی خوب، حرکت کردن همیشه در جهتی که ماکزیمم شیب روبه پایین را دارد - روش نزول شیب - خواهد بود. در حالی که این روش تضمین می‌کند که یک نفر همیشه در یک ارتفاع مینیمم به انتها برسد، ممکن است آن مینیمم قاره‌ای نباشد (مثلا اگر شما در کانزاس شروع کنید، شما احتمالا در نیوآرلین که بیشتر از valley است به انتها برسید - که ممکن است قابل ترجیح باشد برای یک تعداد از دلایل اما به طور واضح شما با هدف اصلی‌تان مواجه نمی‌شدید).

• ابهام. شبکه‌های عصبی جوابهای عددی دقیقی تولید نمی‌کند، اما بیشتر یک نمونه ورودی را به بیشترین احتمال حالت خروجی مرتبط می‌سازد. تکنیکهای برنامه‌نویسی مرسوم بایستی وقتی که دقت مورد نیاز است مورد استفاده قرار گیرد.

• روش. «جعبه سیاه» برای حل مساله. شبکه‌های عصبی می‌توانند برای تبدیل نمونه ورودی به یک خروجی آموزش یابد، اما هیچ بصیرتی را نسبت به فیزیک تبدیل نمی‌کند. آن برای اختصاص دادن هر مفهوم فیزیکی به اندازه یا علامت وزنه‌های اتصالات برای اکثر نمونه‌های شبکه عصبی، امکان پذیر نیست. به هر حال، تحقیقات اخیر روی شبکه‌های عصبی پس انتشار نتیجه گرفته که وزنه‌های اتصالات، در حالی که باز بدون داشتن مفهوم فیزیکی هر ثانیه، می‌تواند برای تعیین توزیع نسبی هر ورودی به هر خروجی مورد استفاده قرار گیرد.

با وجود کمبودها و محدودیت‌های زیاد که در بالا ذکر شد، شبکه‌های عصبی فواید مهمی روی سیستم‌های ماهر، دیگر تکنولوژی هوش مصنوعی، و روشهای برنامه‌نویسی مرسوم دارد. سیستم ماهر که به طور تخصصی کد بندی می‌کند درباره یک حوزه دانش در یک دنباله از قوانین نمی‌تواند به آسانی مسائل شناسایی نمونه پیچیده را به کار برد مانند آنهایی که در تشخیص صورت یا تصویر مواجه شدند. شبکه‌های عصبی برای این کارهای سطح پایین خوب وفق داده می‌شود. بعلاوه، سیستم‌های ماهر در جایی که آن‌جا دانش ناکامل یا اشتباه درباره فیزیک‌های یک مساله وجود دارد، قوی نیستند، و درباره دسته‌بندی شوند به ذهنیات و قواعد عملی برای تولید راه حل‌های قابل قبول، بعلاوه، شبکه‌های عصبی یک خط سیر عملی و مناسب برای حل تهیه می‌کند. برنامه‌های مرسوم معمولاً از یک الگوریتم ثابت و ایستا بر پایه یک مدل پردازش برای تهیه نتایج دقیق عددی استفاده می‌کند. برای مثال، برنامه‌های spking decon فرض می‌کند که طیف عوامل مشترک انعکاسی لایه‌ها در زمین «سفید» است، شکل موج لرزه‌ای با زمان غیرقابل تغییر است، و بنابراین خطاهای بدی می‌تواند رخ دهد وقتی که این فرضیات مختل شوند، شبکه‌های عصبی، و به عبارت دیگر، به طور جهشی تنزل می‌یابد وقتی که داده‌های ورودی نویزدار و ناکامل می‌شوند، یا حتی وقتی که یک بخش از خود شبکه خراب شود. سرانجام، شبکه‌های عصبی بایستی از یک میزان بالایی از پذیرش توسط اجتماعی از استفاده‌کننده‌ها برخوردار باشد. از آن‌جا که شبکه با استفاده از یک مجموعه از نمونه‌ها که استفاده‌کننده انتخاب کرده، آموزش یافته است، شبکه آموزش یافته از رحجان شخصی استفاده‌کننده و روش منحصر به فرد تحلیل، تقلید خواهد کرد. [۱]، [۴] و [۶]

فهرست منابع

فهرست منابع فارسی

- ۱- آربیل و تی. جکسون، البرزی، م.، ۱۳۸۰، آشنایی با شبکه های عصبی، مؤسسه انتشارات علمی دانشگاه صنعتی شریف
- ۲- عربانی، م.، ۱۳۸۱، تخمین تخلخل از روی نگاره های چاه با استفاده از شبکه های عصبی مصنوعی، پایان نامه کارشناسی ارشد، دانشکده مهندسی معدن و متالورژی و نفت، دانشگاه صنعتی امیرکبیر
- ۳- گرامی زادگان، ا.، ۱۳۸۲، مدل سازی عددی در ژئوفیزیک، سمینار کارشناسی ارشد، دانشکده معدن و ژئوفیزیک، دانشگاه صنعتی شاهرود
- ۴- منهاج، م.ب.، ۱۳۷۹. مبانی شبکه های عصبی مصنوعی، مرکز نشر دانشگاه صنعتی امیرکبیر

فهرست منابع انگلیسی

- 6- Baan, M. and Jutten, C. .2000. Neural networks in geophysics applications: Geophysics, 65, No.4, 1032-1047.
- 7- Bhatt, A. and Helle, H.B. .2002, Committee neural networks for porosity and permeability prediction from well logs: Geophysical Prospecting, 50, 645-660.
- 8- Calderon-Macias, C., Sen, M.K. and Stoffa, P.L. .2000, Artificial neural networks for parameter estimation in geophysics: Geophysical Prospecting, 48, 21-47.
- 9- Constable, S. C., Parker, R.L. and Constable, C.G. .1987. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data: Geophysics, 52, No.3, 289-300.
- 10- El-Qady, G. and Ushijima, K. .2001, Inversion of DC resistivity data using neural networks: Geophysical Prospecting, 49, 417-430.
- 11- Ellis, R. G. and Oldenburg, D. W. .1994b. The pole-pole 3-D DC resistivity inverse problem: a conjugate gradient approach: Geophysical Journal International, 119, 187-194.
- 12- Griffiths, D. H. and Barker, R. D. .1993, Two-dimensional resistivity imaging and modelling in areas of complex geology: Journal of applied geophysics, 29, 211-226.

- 13- Interpex Limited Golden Co.,1993,RESIX,V3.0, User's Manual-Resistivity data interpretation software.
- 14- The Mathworks Inc.,2002, Matlab,V6.5. User Manual for the Language of technical computing, The Mathworks Inc.
- 15- Lippmann, R. P. ,1987, An Introduction to Computing with Neural Nets: IEEE ASSP Mag.,**4-22**,417-435.
- 16- Loke, M. H. and Barker, R.D. ,1996, Rapid least -squares inversion of apparent resistivity pseudosections using quasi - Newton meyhod: Geophysical prospecting,**44**,499-523.
- 17- Loke, M. H. and Dahlin, T. .1997. A combined Gauss - Newton and Quasi - Newton inversion method for the interpretation of apparent resistivity pseudosections,Proceeding: 3rd meeting of the European Association for Environmental and Engineering Geophysics
- 18- McCormack, M. D. ,1991, Neural computing in geophysics: Geophysics,**10**,No.1,11-15.
- 19- Milsom, J. ,1939, Field Geophysics.University College.
- 20- Mooney, M. ,1980, Handbook of Engineering Geophysics,Minnesota,Bison Instruments,Inc.
- 21- Murat, M. E. and Rudman,A.J. ,1992, Automated first arrival picking: A neural network approach: Geophys.Pros.,**40**,587-604.
- 22- Olayinka, A. I. and Yaramanci, U. ,2000, Use of block inversion in 2-D interpretation of apparent resistivity data and its comparison with smooth inversion: Journal of Applied Geophysics,**45**,63-82.
- 23- Pearson, W., Wiener, J. and Moll,R. .1990, Aeromagnetic structural interpretation using neural network: a case study from the northern Denver-Julesburg Basin,61th SEG meeting
- 24- Pelton, W. H., Rijo, L. and Swift, C. M., Jr .1978, Inversion of two-dimensional resistivity and induced polarization data: Geophysics.**43**.788-803.
- 25- Poulton, M. and El-fouly,A. .1991. Preprocessing GPR signatures for cascading neural network classification,SEG meeting
- 26- Poulton, M. M., Sternberg,B.K. and Glass,C.E. ,1992, Location of subsurface targets in geophysics data using neural networks: Geophysics,**57**,No.12,1534-1544.
- 27- Poulton, M. M. .2001, Computational Neural Networks for Geophysical Data Processing,Arizona.Department of Mining & Geological Engineering.
- 28- Spichak, V. and Popova,I. .2000, Artificial Neural Network Inversion of Magnetotelluric Data in Terms of Three Dimentional Earth Macroparameters:

Geophysical Journal International, **142**, 15-26.

29- Telford, W. M., Geldart, L.P. and Sheriff, R.E., 1990, Applied Geophysics, Cambridge University Press.

30- Trip, A. C., Hohmann, G. W. and Swift, C. M. Jr., 1984. Two-dimensional resistivity inversion: Geophysics, **49**, 1708-1717.

31- Zhang, L., Poulton, M. and Wang, T., 2002. Borehole electrical resistivity modeling using neural networks: Geophysics, **67**, No. 6, 1790-1797.

32- Calderon-Macias, C. and Sen, M.K., 1993. Geophysical Interpretation by Artificial Neural Systems: A Feasibility Study: 63rd Ann. Internat. Mtg: Soc. Of Expl. Geophys., 254-257.

33- Cisar, D. and Novotny, T.J., 1992. Electromagnetic Data Evaluation Using a Neural Network: Initial Investigation – Underground Storage Tanks: 62nd Ann. Internat. Mtg: Soc. Of Expl. Geophys., 409-411.

34- Fei, D., Teng, Y.-C. and Kuo, J.T., 1994, Detection of conductive thick plate based on finite-element method and neural networks: 64th Ann. Internat. Mtg: Soc. Of Expl. Geophys., 636-639.

35- Poulton, K.Y., Glass, C.E. and Sternberg, B.K., 1989. Recognizing EM ellipticity patterns with neural network : 59th Ann. Internat. Mtg: Soc. Of Expl. Geophys., 208-212.

36- Zerilla, A., Massimo, F., Ronchini, G. and Apolloni, B., 1992, Lineaments analysis for potential fields data using neural networks: 62nd Ann. Internat. Mtg: Soc. Of Expl. Geophys., 6-9.

37- Guo, Y., Hansen, R.O., and Harthill, N., 1992, Feature recognition from potential fields using neural networks: 62nd Ann. Internat. Mtg: Soc. Of Expl. Geophys., 1-5.

38- Cartabia, G., Zerilla, A. and Apolloni, B., 1994, Lineaments recognition for potential fields images using a learning algorithm for Boltzmann machines: 64th Ann. Internat. Mtg: Soc. Of Expl. Geophys., 432-435.

39- Cai, Y., 1994, The Artificial neural network for research of the recovery ratio of oil fields: 64th Ann. Internat. Mtg: Soc. Of Expl. Geophys., 791-793.

40- Zhang, Z. and Zhou, Z., 2002, Real-Time quasi-2-D inversion of array resistivity logging data using neural network: Geophysics, **67**, 517-524