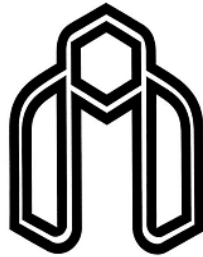


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شاهرود

دانشکده : مهندسی معدن، نفت و ژئوفیزیک

پایان نامه کارشناسی ارشد ژئوفیزیک – گرایش ژئوالکتریک

عنوان:

بهینه سازی پاسخ شبکه‌ی عصبی کمیته‌ای در تخمین اشباع از آب

مخزن توسط الگوریتم ژنتیک

دانشجو :

مکی لویمی

اساتید راهنما :

ابوالقاسم کامکار روحانی

ناصر کشاورز فرج‌خواه

تیر ۱۳۸۹

## دانشگاه صنعتی شاهرود

دانشکده : مهندسی معدن، نفت و ژئوفیزیک

گروه : اکتشاف

پایان نامه کارشناسی ارشد آقای مکی لویمی

تحت عنوان: بهینه سازی پاسخ شبکه عصبی کمیته ای در تخمین اشباع از آب مخزن با استفاده از الگوریتم ژنتیک

در تاریخ ..... توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجه ..... مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور	امضاء	اساتید راهنما
	نام و نام خانوادگی :		نام و نام خانوادگی :
	نام و نام خانوادگی :		نام و نام خانوادگی :

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور
	نام و نام خانوادگی :		نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :

## تشکر و قدردانی:

از اساتید محترم، دکتر کامکار روحانی و دکتر کشاورز فرج‌خواه که زحمت راهنمایی اینجانب را در انجام این تحقیق بر عهده گرفتند تشکر ویژه دارم.

از کلیه اساتید ارجمندم که در این مدت از محضرشان بهره گرفته‌ام تشکر و قدردانی می‌نمایم.

از داوران محترمی که مطالعه و داوری این پایان‌نامه را بر عهده گرفتند تشکر می‌کنم.

دانشجو تأیید می نماید که مطالب مندرج در این پایان نامه ( رساله ) نتیجه تحقیقات خودش می باشد و در صورت استفاده از نتایج دیگران مرجع آن را ذکر نموده است.

کلیه حقوق مادی مترتب از نتایج مطالعات ، آزمایشات و نو آوری ناشی از تحقیق موضوع این پایان نامه ( رساله ) متعلق به دانشگاه صنعتی شاهرود می باشد .

ماه و سال

## چکیده

یکی از پارامترهای بسیار مهم و کاربردی در مخازن هیدروکربنی، پارامتر اشباع از آب می‌باشد که به صورت درصد فضاهای خالی که توسط آب پر شده است تعریف می‌شود. در این تحقیق از شبکه‌های عصبی کمیته‌ای<sup>۱</sup> برای تخمین این پارامتر استفاده شده است. برای این کار از داده‌های ۵ چاه واقع در یکی از میادین جنوب غربی ایران استفاده شده است. مقادیر عددی نگارهای صوتی، چگالی، پرتو گاما، مقاومت ویژه، مقاومت ویژه نرمال بلند و همچنین مقدار تخلخل مؤثر (که از روی نگارهای مختلف و داده‌های مغزه‌ها به دست آمده است) به عنوان ورودی و اشباع از آب به عنوان خروجی در نظر گرفته شد. این روش به مراتب از زمانی که یک تک شبکه داریم قوی‌تر و دقیق‌تر می‌باشد. اجزاء این ترکیبات کمیته‌ای شبکه‌های پرسپترون چند لایه می‌باشند که بهترین ساختار آنها با فرایند سعی و خطا از لحاظ داشتن کمترین خطای آزمون انتخاب شدند. تعداد ۵۹ شبکه با ساختارهای مختلف مورد آموزش و آزمون قرار گرفتند. آموزش این ۵۹ شبکه با ۳ روش فراآموزش، مرتب سازی و توقف سریع انجام شد. از بین این ۵۹ شبکه، ۷ شبکه با کمترین خطا برای ساختن ترکیبات کمیته‌ای انتخاب و ۱۲۰ ترکیب ممکن، یعنی ۲۱ ترکیب دو شبکه‌ای، ۳۵ ترکیب سه شبکه‌ای، ۳۵ ترکیب چهار شبکه‌ای، ۲۱ ترکیب پنج شبکه‌ای، ۷ ترکیب شش شبکه‌ای و ۱ ترکیب هفت شبکه‌ای ایجاد شد. ضرایب وزنی این ترکیبات خطی با استفاده از دو روش الگوریتم ژنتیک و میانگین گیری ساده به دست آورده شدند و سپس نتایج حاصل از این ترکیبات با هم مقایسه شد.

برای روش فرا آموزش، شبکه تک لایه با ساختار (۱-۱۰-۶) به عنوان بهترین شبکه شناخته شد. در حالتی که ضرایب وزنی ترکیب خطی شبکه‌های (۱-۱۰-۶)، (۱-۵-۶-۶)، (۱-۱۳-۶)، (۱-۱۱-۶) و (۱-۹-۴-۶)، با استفاده از روش الگوریتم ژنتیک به دست آورده شد، بیشترین کاهش خطای آزمون در ترکیب این شبکه‌ها مشاهده گردید.

---

<sup>1</sup> Committee neural networks

در روش مرتب سازی، شبکه با ساختار (۶-۱۰-۲-۱) کمترین خطا را در مرحله‌ی تعمیم به همراه داشت. در حالتی که ضرایب وزنی ترکیب خطی شبکه‌های (۶-۱۰-۲-۱)، (۶-۶-۵-۱)، (۶-۵-۶-۱)، (۶-۴-۱۰-۱) و (۶-۶-۶-۱)، با استفاده از روش الگوریتم ژنتیک به دست آورده شد، بیشترین کاهش خطای آزمون در ترکیب این شبکه‌ها مشاهده گردید.

در روش توقف سریع، شبکه با ساختار (۶-۸-۸-۱)، در مرحله‌ی تعمیم دارای کمترین خطا بود. ضرایب وزنی ترکیب خطی شبکه‌های (۶-۸-۸-۱)، (۶-۸-۹-۱)، (۶-۱۵-۱)، (۶-۶-۹-۱) و (۶-۷-۱۰-۱)، با استفاده از روش الگوریتم ژنتیک به دست آورده شد و بیشترین کاهش خطای آزمون در ترکیب این شبکه‌ها مشاهده گردید.

کلمات کلیدی: اشباع از آب، شبکه‌های عصبی مصنوعی، بهینه سازی، ترکیبات آنسامبلی، الگوریتم ژنتیک.

## مقالات مستخرج

۱. لویمی، م. کامکار روحانی، ا. کشاورز، ن. ۱۳۸۸، تخمین نگار نوترون توسط شبکه-

های مصنوعی و الگوریتم ژنتیک، بیست و هفتمین گردهمائی علوم زمین و

سیزدهمین همایش انجمن زمین‌شناسی ایران، تهران.

۲. لویمی، م. کامکار روحانی، ا. کشاورز، ن. ۱۳۸۹، تخمین اشباع از آب توسط

شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک، چهاردهمین کنفرانس ژئوفیزیک

ایران، تهران.



## فهرست مطالب

۱	فصل اول: کلیات
۱-۱	مقدمه
۲-۱	اشباع از آب سازند
۱-۲-۱	اندازه گیری از روی مغزه ها
۱-۱-۲-۱	روش تقطیر
۲-۱-۲-۱	روش دین - استارک
۲-۲-۱	محاسبه از روی نگارها
۱-۲-۲-۱	نگار صوتی
۲-۲-۲-۱	نگار چگالی
۳-۲-۲-۱	نگار پرتو گامای طبیعی
۴-۲-۲-۱	نگارهای مقاومت ویژه الکتریکی
۳-۲-۱	رابطه آرچی
۴-۲-۱	نمودار هینگل
۵-۲-۱	نمودار پیکت
۳-۱	کاربرد روش شبکه های عصبی مصنوعی در تخمین مقدار اشباع از آب سنگ مخزن
۴-۱	مروری بر مطالعات انجام شده
۵-۱	ضرورت و هدف از مطالعه حاضر
۶-۱	ساختار پایان نامه
۲۵	فصل دوم: مبانی شبکه های عصبی مصنوعی
۱-۲	مقدمه
۲-۲	ساختار شبکه های عصبی مصنوعی
۳-۲	آموزش شبکه های عصبی مصنوعی
۴-۲	الگوریتم پس انتشار خطا
۵-۲	بایاس و واریانس
۶-۲	بهبود توان تعمیم شبکه
۱-۶-۲	آموزش به روش مرتب سازی
۲-۶-۲	آموزش به روش توقف سریع
۴۱	فصل سوم: مبانی شبکه های عصبی مصنوعی کمیته ای
۱-۳	مقدمه
۲-۳	سیستم شبکه های چند گانه
۱-۲-۳	ترکیبات آنسامبلی
۲-۲-۳	ترکیب مدوله ای
۳-۳	روش های ترکیب
۴-۳	الگوریتم ژنتیک

۵۰	۱-۴-۳- مقدمه
۵۱	۲-۴-۳- تاریخچه الگوریتم های ژنتیک
۵۲	۳-۴-۳- پیش زمینه بیولوژیکی ژن ها و کروموزوم ها
۵۳	۴-۴-۳- بهینه سازی بالگوریتم های ژنتیک
۵۵	۵-۴-۳- تولید مثل
۵۵	۶-۴-۳- انواع روش های انتخاب
۵۶	۱-۶-۴-۳- روش چرخ گردان
۵۷	۷-۴-۳- عملگر های GA
۵۸	۱-۷-۴-۳- عملگر ادغام
۵۹	۱-۱-۷-۴-۳- روش ادغام تک نقطه ای یا مکانی
۶۰	۲-۱-۷-۴-۳- روش ادغام دو نقطه ای
۶۰	۳-۱-۷-۴-۳- روش ادغام چند نقطه ای
۶۱	۲-۷-۴-۳- جهش
۶۲	۱-۲-۷-۴-۳- عمل جهش یا موتاسیون
۶۴	۸-۴-۳- شرط توقف الگوریتم

#### فصل چهارم: تخمین اشباع از آب سازند سروک واقع در یکی از میادین جنوب غربی ایران، توسط شبکه

۶۵	های عصبی کمیته ای و الگوریتم ژنتیک
۶۶	۱-۴- مقدمه
۶۶	۲-۴- زمین شناسی میدان مورد مطالعه
۶۹	۳-۴- چینه شناسی میدان مورد نظر
۷۰	۱-۳-۴- سازند سروک
۷۳	۴-۴- چاه های مورد مطالعه
۷۴	۵-۴- انتخاب داده های آموزش و آزمون
۷۵	۶-۴- برنامه نویسی رایانه ای
۷۸	۷-۴- ایجاد شبکه های پس انتشار خطا به روش فرا آموزش و ترکیب نتایج
۸۴	۱-۷-۴- ترکیبات آنسامبلی خطی
۹۰	۲-۷-۴- واریانس و بایاس خطا
۹۲	۸-۴- ایجاد شبکه های پس انتشار خطا به روش مرتب سازی و ترکیب نتایج
۱۰۰	۱-۸-۴- ترکیبات آنسامبلی خطی
۱۰۶	۲-۸-۴- واریانس و بایاس خطا
۱۰۸	۹-۴- ایجاد شبکه های پس انتشار خطا به روش توقف سریع و ترکیب نتایج
۱۱۶	۱-۹-۴- ترکیبات آنسامبلی خطی
۱۲۲	۲-۹-۴- واریانس و بایاس خطا

#### فصل پنجم: نتایج و پیشنهادات

۱۲۵	۱-۵- جمع بندی و نتیجه گیری
۱۲۷	۲-۵- پیشنهادات

فهرست منابع	۱۲۹
پیوست	۱۳۵
پیوست الف:	۱۳۶
انواع الگوریتمهای آموزشی پس انتشار خطا و توابع مربوط به آنها در نرمافزار MATLAB	۱۳۶
الف-۱- الگوریتم پسانتشار خطای تفریق گرادیان	۱۳۸
الف-۲- الگوریتم پسانتشار خطای تفریق گرادیان با تکانه	۱۳۹
الف-۳- الگوریتمهای پسانتشار خطا با سرعت یادگیری متغیر	۱۴۰
الف-۴- الگوریتم پسانتشار خطای بهبودپذیر	۱۴۱
الف-۵- الگوریتمهای گرادیان مزدوج	۱۴۲
الف-۵-۱- الگوریتم فلچر - ریویس	۱۴۳
الف-۵-۲- الگوریتم پولاک- ریبر	۱۴۵
الف-۵-۳- الگوریتم شروع پاول- بیل	۱۴۵
الف-۵-۴- الگوریتم گرادیان مزدوج مقیاسی	۱۴۶
الف-۶- الگوریتمهای شبه نیوتنی	۱۴۷
الف-۶-۱- الگوریتم بیافجیاس	۱۴۷
الف-۶-۲- الگوریتم متقاطع یک مرحلهای	۱۴۸
الف-۷- الگوریتم لونبرگ-مارکوت	۱۴۸
الف-۸- تابع مرتبسازی بیزین (TRAINBR)	۱۵۰
پیوست ب: روش های مختلف بهینه سازی	۱۵۲
ب-۱- روش شبیه سازی تبرید تدریجی	۱۵۳
ب-۲- الگوریتم کلونی مورچه ها	۱۵۵
ب-۳- الگوریتم جستجوی تابو	۱۵۷
ب-۴- الگوریتم جستجوی هارومونی	۱۵۹
ب-۵- الگوریتم اتوماتیک سلولی	۱۶۰
ب-۶- الگوریتم پرندگان	۱۶۱

## فهرست اشکال

- شکل ۱-۱: تصویری از عملکرد یک ابزار مقاومت نرمال ..... ۱۰
- شکل ۲-۱: نمایی از عملکرد یک ابزار مقاومت جانبی ..... ۱۱
- شکل ۳-۱: نمودار هینگل ..... ۱۵
- شکل ۴-۱: نمودار پیکت ..... ۱۶
- شکل ۱-۲: نمایی از یک نرون بیولوژیکی ..... ۲۷
- شکل ۲-۲: مدل یک نرون ..... ۲۸
- شکل ۳-۲: تابع زیگموئید ..... ۳۱
- شکل ۴-۲: نمایی از ساختار لایه‌های شبکه‌های عصبی ..... ۳۱
- شکل ۱-۳: ترکیب خطی شبکه‌های آموزش دیده ..... ۴۸
- شکل ۲-۳: یک رشته DNA ..... ۵۲
- شکل ۳-۳: مدلی از چرخ گردان ..... ۵۷
- شکل ۴-۳: روش ادغام تک نقطه‌ای ..... ۵۹
- شکل ۵-۳: روش ادغام دو نقطه‌ای ..... ۶۰
- شکل ۶-۳: روش ادغام چند نقطه‌ای (حالتی که تعداد مکان‌ها زوج باشد) ..... ۶۱
- شکل ۷-۳: روش ادغام چند نقطه‌ای (حالتی که تعداد مکان‌ها فرد باشد) ..... ۶۱
- شکل ۸-۳: نیاز یک جمعیت به بیت ۱ ..... ۶۳
- شکل ۹-۳: تاثیر عملگر جهش ..... ۶۳
- شکل ۱-۴: تقسیمات تکتونیکي حوزه زاگرس در ایران، کادر مستطیلی موقعیت میدان مورد نظر را نشان میدهد ..... ۶۷
- شکل ۲-۴: واحدهای پایینی و میانی چینه‌شناسی دوره‌ی کرتاسه در جنوب غربی ایران ..... ۷۲
- شکل ۳-۴: همبستگی سازندهای دوره‌ی کرتاسه شمال شرقی صفحه‌ی عربی از عراق تا جنوب غربی ایران ..... ۷۳
- شکل ۴-۴: موقعیت چاه‌های مورد استفاده ..... ۷۴
- شکل ۵-۴: نمای شماتیک مراحل طی شده در این تحقیق ..... ۷۷
- شکل ۴-۶-الف: روند کاهش خطا برای مرحله‌ی آموزش شبکه‌ی شماره ۱ جدول ۳-۴ ..... ۸۲
- شکل ۴-۶-ب: نمودار همبستگی بین اشباع از آب و تخمین شبکه‌ی شماره ۱ جدول ۳-۴ برای داده‌های آموزش ..... ۸۲
- شکل ۴-۶-ج: مقادیر اشباع از آب و تخمین شبکه شماره ۱ جدول ۳-۴ برای داده‌های آموزش ..... ۸۳
- شکل ۴-۷: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک ..... ۸۹

- شکل ۴-۸- الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۹۱
- شکل ۴-۸- ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آزمون (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۹۱
- شکل ۴-۹- الف: نمایش روند آموزش و همگرایی SSE و SSW و تعداد پارامترهای مؤثر شبکه‌ی شماره‌ی ۱ جدول ۴-۸ ..... ۹۶
- شکل ۴-۹- ب: نمودار همبستگی مقادیر اشباع از آب و تخمین شبکه‌ی ۱ جدول ۴-۸ برای داده‌های آموزش ..... ۹۷
- شکل ۴-۹- د: نمودار همبستگی بین مقادیر اشباع از آب و تخمین شبکه‌ی ۱ جدول ۴-۸ برای داده‌های آزمون ..... ۹۹
- شکل ۴-۱۰: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک ..... ۱۰۵
- شکل ۴-۱۱- الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۱۰۷
- شکل ۴-۱۱- ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آزمون (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۱۰۷
- شکل ۴-۱۲- الف: نمایش روند آموزش و کاهش میانگین مربعات خطا برای داده‌های آموزش و اعتبار سنجی در شبکه شماره ۱ جدول ۴-۱۳ ..... ۱۱۲
- شکل ۴-۱۲- ب: نمودار همبستگی مقادیر اشباع از آب و تخمین شبکه ۱ جدول ۴-۱۳ برای داده‌های آموزش ..... ۱۱۳
- شکل ۴-۱۲- ج: مقادیر اشباع از آب و تخمین شبکه شماره ۱ جدول ۴-۱۳ برای داده‌های آموزش ..... ۱۱۴
- شکل ۴-۱۲- د: نمودار همبستگی بین مقادیر اشباع از آب و تخمین شبکه ۱ جدول ۴-۱۳ برای داده‌های آزمون ..... ۱۱۵
- شکل ۴-۱۲- ه: مقادیر اشباع از آب و تخمین شبکه‌ی شماره‌ی ۱ جدول ۴-۱۳ برای داده‌های آزمون ..... ۱۱۶
- شکل ۴-۱۳: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک ..... ۱۲۱
- شکل ۴-۱۴- الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۱۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۱۲۳
- شکل ۴-۱۴- ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آزمون (۱: شبکه ۱ جدول ۴-۱۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک) ..... ۱۲۳

## فهرست جداول

جدول ۴-۱: چینه شناسی میدان مورد نظر .....	۶۸
جدول ۴-۲: نتایج مربوط به شبکه‌های ساخته شده با یک و دو لایه‌ی میانی به روش فرا آموزش .....	۷۸
جدول ۴-۳: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش فرا آموزش بر اساس داشتن کمترین خطای آزمون .....	۸۱
جدول ۴-۴: نتایج ترکیبات آنسامبلی خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده .....	۸۵
جدول ۴-۵: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک .....	۸۸
جدول ۴-۶: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه ۱ جدول ۴-۳ و بهترین ترکیب آن جدول ..... ۹۰	۹۰
جدول ۴-۷: نتایج مربوط به شبکه‌های پس انتشار خطای آموزش دیده با ۱ و ۲ لایه‌ی میانی به روش مرتب‌سازی .. ۹۲	۹۲
جدول ۴-۸: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش مرتب‌سازی بر اساس داشتن کمترین خطای آزمون .....	۹۵
جدول ۴-۹: نتایج ترکیبات آنسامبلی خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده .....	۱۰۱
جدول ۴-۱۰: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک .....	۱۰۴
جدول ۴-۱۱: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه‌ی ۱ جدول ۴-۸ و بهترین ترکیب آن جدول ۱۰۶	۱۰۶
جدول ۴-۱۲: نتایج مربوط به شبکه‌های ساخته شده با یک و دو لایه‌ی میانی به روش توقف سریع .....	۱۰۸
جدول ۴-۱۳: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش توقف سریع بر اساس داشتن کمترین خطای آزمون .....	۱۱۱
جدول ۴-۱۴: نتایج ترکیبات کمیته‌ای خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده .....	۱۱۷
جدول ۴-۱۵: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک .....	۱۲۰
جدول ۴-۱۶: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه‌ی ۱ جدول ۴-۱۳ و بهترین ترکیب آن جدول .....	۱۲۲

# فصل اول

## کلیات

## ۱-۱- مقدمه

انباشته شدن مواد هیدروکربنی در زیر سطح زمین در سنگ‌هایی صورت می‌گیرد که توانایی نگه‌داری و انتقال سیالات را داشته باشند. این سنگ‌ها، به نام سنگ‌های مخزن<sup>۱</sup> خوانده می‌شوند. تجمع مواد هیدروکربنی به صورت اقتصادی در سنگ مخزن منوط به وجود عوامل متعددی است. به طور کلی وجود پنج عامل زیر برای تجمع اقتصادی نفت و گاز لازم و ضروری است [۹]:

۱. سنگ منشأ بالغ<sup>۲</sup> که تولید هیدروکربن کرده باشد.

۲. سنگ مخزن<sup>۳</sup> که بتواند هیدروکربن را در داخل خود جا دهد.

۳. مهاجرت هیدروکربن بین سنگ منشأ و سنگ مخزن<sup>۴</sup> عملی باشد.

۴. پوش سنگ<sup>۵</sup> ناتراوا که از خروج نفت از داخل سنگ مخزن جلوگیری کند.

۵. تله نفتی<sup>۶</sup> که در آن نفت به صورت اقتصادی متمرکز گردد.

سنگ منشأ از لحاظ تعریف، سنگ دانه ریز غنی از ماده آلی است که قادر است در اثر تکامل حرارتی تولید هیدروکربن نماید. پدیده‌ی مهاجرت به دو بخش مهاجرت اولیه<sup>۷</sup> و مهاجرت ثانویه<sup>۸</sup> قابل تقسیم است. مهاجرت اولیه به حرکت نفت و گاز از سنگ منشأ به لایه‌ی تراواتر و یا مخزن گفته می‌شود. در طی مهاجرت ثانویه نفت یا گاز در لایه‌ها حرکت می‌کنند تا اینکه در یک تله‌ی نفتی متمرکز شوند [۹].

پوش سنگ‌ها سنگ‌هایی هستند که می‌توانند از حرکت رو به بالای هیدروکربن جلوگیری نمایند. ضرورتاً یک پوش سنگ باید در برابر هیدروکربن نفوذ ناپذیر باشد. اساساً سنگی که تراوایی صفر داشته باشد وجود ندارد اما بعضی از سنگ‌ها تراوایی خیلی کمی دارند به گونه‌ای که می‌توان از آن صرف نظر کرد. سنگ‌های دانه ریز مانند شیل می‌توانند یک پوش سنگ خوب محسوب شوند. بیش از ۶۰ درصد میدان‌های بزرگ شناخته شده دارای پوش سنگ شیلی هستند.

1. Reservoir  
5. Cap rock

2. Mature source rock  
6. Oil trap

3. Reservoir rock  
7. Primary migration

4. Migration pathway  
8. Secondary migration



نفتگیر یا تله‌ی نفتی، یکی از عوامل اصلی و ضروری برای تجمع اقتصادی هیدروکربن می‌باشد. ساده-ترین و فراوان‌ترین نفتگیرها نفتگیرهای طاق‌دیسی<sup>۱</sup> هستند.

سنگ مخزن سنگی است که هم دارای ظرفیت ذخیره‌سازی بوده (متخلخل باشد) و هم توانایی عبور سیال از میان خود را داشته باشد (تراوا باشد). نحوه‌ی رفتار هیدروکربن درون مخزن عمدتاً به وسیله برخی خواص ذاتی سیالات درون مخزن و محیط متخلخل کنترل می‌شود. تحلیل خصوصیات زمین-شناسی مخزن که جریان سیال و مکانیسم‌های بازیافت را کنترل می‌کند، باعث ایجاد مدل‌های کاربردی برای شبیه‌سازی مخزن<sup>۲</sup> و توسعه‌ی میدان می‌شود.

در ارزیابی یک مخزن از کلیه‌ی تکنیک‌ها و روش‌ها استفاده می‌شود تا بتوان پارامترهای زمین‌شناسی و پتروفیزیکی کنترل‌کننده‌ی حرکت سیال را در آن مشخص نمود. درک متغیرها و پارامترهای پیچیده‌ی مخزنی نظیر تخلخل، تراوایی، اشباع‌شدگی و ... نیازمند یک کوشش همه‌جانبه است. در این مطالعه از داده‌های منابع مختلف استفاده می‌شود. این منابع برای مثال شامل مغزه‌ها، نمودارها و ... می‌باشند [۱ و ۹].

## ۱-۲- اشباع از آب سازند

یکی از مهمترین پارامترها در مخازن هیدروکربنی پارامتر اشباع از آب می‌باشد. این کمیت به صورت درصدی از محیط متخلخل که توسط آب پر شده است تعریف می‌شود. دو روش عمده برای اندازه‌گیری این پارامتر وجود دارند: ۱- اندازه‌گیری از طریق مغزه‌ها ۲- محاسبه با استفاده از نگارها و روابط تجربی. نمودار مقاومت ویژه ابزار اصلی برای محاسبه این پارامتر است. از طریق نمودارهای دیگر نیز می‌توان این محاسبه را انجام داد. برای مثال از ترکیب نمودار دانسیته و شاخص هیدروژنی به دست

---

1. Anticlinal trap

2. Reservoir simulation

آمده از نگارهای تشدید مغناطیسی هسته‌ای<sup>۱</sup> (NMR) یا نوترون می‌توان اشباع از آب را محاسبه کرد، ولی این روش فقط به گاز حساس بوده و گستره‌ی اندازه‌گیری آن محدود به اطراف چاه است [۳] و [۲۲]. در ادامه به چند روش که برای تعیین اشباع از آب به کار می‌روند اشاره می‌کنیم.

### ۱-۲-۱- اندازه‌گیری از روی مغزه‌ها

علاوه بر محاسبه‌ی اشباع از آب از روی نمودارها و روابط تجربی که رویکرد اصلی صنعت می‌باشد گاهی این امر با اندازه‌گیری از روی مغزه‌ها صورت می‌گیرد. در اینجا به دو روش برای تخمین این کمیت از روی مغزه‌ها اشاره می‌کنیم [۱۸ و ۳۹]:

#### ۱-۱-۲-۱- روش تقطیر<sup>۲</sup>

در این روش، نمونه تحت تأثیر دمای بالا قرار می‌گیرد و میزان آب و نفت مستخرج از آن محاسبه می‌شود. قبل از اینکه نمونه درون دستگاه تقطیر قرار بگیرد، خرد شده و وزن آن را اندازه می‌گیرند. سپس تا دمای  $65.0^{\circ}\text{C}$  گرم می‌شود و سیالات خارج شده به تدریج چگالیده و از یکدیگر جدا می‌شوند. حجم آب و نفت خارج شده به صورت مستقیم محاسبه می‌شود ولی حجم گاز نیز برای اندازه‌گیری‌های دقیق باید تعیین شود. این کار با تزریق جیوه با فشار بالا در نمونه‌ی مشابه و مشاهده میزان جذب شده صورت می‌گیرد. قبل از تزریق، وزن نمونه اندازه‌گیری می‌شود و حجم فضای خالی آن از طریق میزان جیوه‌ی جابجا شده تعیین می‌شود [۱۸ و ۴۲].

---

1. Nuclear Magnetic Resonance

2. Retort method

## ۱-۲-۱-۲- روش دین - استارک<sup>۱</sup>

در این روش با جوشاندن حلال آب موجود خارج و در ظرف دیگری چگالیده می‌شود. حلال مزبور نیز چگالیده و بار دیگر نمونه را در بر می‌گیرد و با تکرار مرحله‌ی قبل نفت را استخراج می‌کند. این فرایند حداقل برای ۲ روز ادامه پیدا می‌کند تا جایی که حلال استخراج شده حاوی هیچ‌گونه ناخالصی نباشد. وزن نمونه قبل و بعد از استخراج اندازه‌گیری می‌شود و از این طریق میزان تخلخل، اشباع از آب و نفت محاسبه می‌شود [۱۸ و ۴۲].

## ۱-۲-۲- محاسبه از روی نگارها

در طی حفاری، گرفتن مغزه از سنگ‌های مورد نظر معمولاً وقت‌گیر و مستلزم مخارج زیادی است. خرده سنگ‌های به دست آمده از حفاری نیز به دلیل کوچک بودن، مخلوط بودن باهم و شکستگی توسط سیال حفاری، اغلب در تفسیر مورد اطمینان نیستند. به این دلیل اطلاعات کمی از خواص فیزیکی اولیه‌ی سازندهایی که از آن گرفته شده‌اند، در اختیار قرار می‌دهند [۲۸]. نمودارهای چاه-پیمایی کمک بزرگی برای حل این مشکل و کسب اطلاعات بیشتر از فواصل حفاری شده می‌باشد. نگار چاه، نموداری است که تغییرات یک متغیر فیزیکی یا مشتقات آن را نسبت به عمق نشان می‌دهد. از اهداف مهم نمودارگیری تهیه‌ی یک سری اطلاعات درجا از مخزن است. پارامترهای مهم قابل دستیابی از نمودار عبارتند از: تخلخل، نسبت اشباع آب و نفت، ضخامت زون تولید<sup>۲</sup>، لیتولوژی<sup>۳</sup>، تراوایی و ... [۹ و ۱۳].

در اینجا به توضیح چند نمونه از نگارها می‌پردازیم.

---

1. Dean – Stark method  
3. Lithology

2. Net pay thickness

## ۱-۲-۲-۱- نگار صوتی

در این نوع نگار ثبت پارامترهایی مورد نظر است که در ارتباط با فرستادن موج به داخل سازند به دست می‌آید. یکی از این پارامترها، مدت زمان عبور موج از دیواره چاه و رسیدن آن از فرستنده به گیرنده است. زمانی که به وسیله‌ی سوند ثبت می‌شود ( $T_t$ ) تابع سه عامل اصلی زیر است:

۱. مدت زمان سیر موج در خمیره یا ماتریس سنگ ( $T_m$ ).

۲. مدت زمان سیر موج در سیال پرکننده‌ی فضاهای خالی ( $T_f$ ).

۳. حجم خلل و فرج و یا همان تخلخل ( $\varphi$ ).

بنابراین می‌توان نوشت [۱۳]:

$$T_t = (1 - \varphi)T_m + \varphi T_f \quad (1-1)$$

با دانستن نوع خمیره و سیالی که فضاهای خالی را پر کرده، می‌توان تخلخل را محاسبه کرد. رابطه‌ی فوق برای سازندهای فشرده و فاقد شیل صادق است، در غیر این صورت تصحیحات لازم باید صورت بگیرد.

چون موج کوتاهترین مسیر را انتخاب می‌کند، در صورتی که در سازند حفره و شکستگی (تخلخل ثانویه) وجود داشته باشد، موج صوتی تمایل به نادیده گرفتن آن‌ها و عبور از نقاط دیگر را داشته و بنابراین نگار صوتی فقط تخلخل اولیه را ثبت می‌کند.

در صورتی که سازند فشرده و متراکم نباشد از ضریب فشردگی ( $C_p$ ) که معمولاً بین ۰/۸ تا ۱/۲ تغییر

می‌کند، استفاده می‌کنیم و  $\varphi$  را در  $\frac{1}{C_p}$  ضرب می‌کنیم [۱۳].

## ۱-۲-۲-۲- نگار چگالی

وقتی که پرتوهای گامای مصنوعی به سازند تابیده می‌شوند، بر حسب انرژی‌شان می‌توانند سه نوع برهمکنش با سازند داشته باشند: اثر فوتوالکتریک، پراکندگی کامپتون و تولید جفت.

در نگار چگالی از اثر کامپتون استفاده می‌شود. پراکندگی کامپتون از برخوردهای با انرژی متوسط با سازند، ایجاد می‌شود. گامای پراکنده شده توسط گیرنده‌ها دریافت می‌شود و شمارش می‌گردد. دستگاه، چگالی الکترونی<sup>۱</sup> ( $\rho_e$ ) سازند را بررسی خواهد کرد زیرا هرچه تعداد برخوردها بیشتر و یا به عبارت دیگر تعداد الکترون‌ها در واحد حجم سازند زیادتر باشد، شدت ثبت شده کمتر خواهد بود. از طرف دیگر، چگالی الکترونی نسبت مستقیم با چگالی سازند<sup>۲</sup> ( $\rho_b$ ) دارد [۱۳]:

$$\rho_e = \left( \frac{2Z}{A} \right) \rho_b \quad (۲-۱)$$

در این رابطه  $Z$  عدد اتمی و  $A$  جرم اتمی عناصر و ترکیبات تشکیل دهنده‌ی سنگ است. در اغلب عناصر و ترکیبات، چگالی الکترونی به چگالی ظاهری حاصل از دستگاه خیلی نزدیک است بنابراین مقدار  $\rho_b$  به دست می‌آید.

چگالی کل به سه عامل چگالی خمیره سنگ ( $\rho_{ma}$ )، چگالی سیالات پرکننده‌ی فضاهای خالی ( $\rho_f$ ) و تخلخل سازند وابسته است [۱۳]:

$$\rho_b = (1 - \varphi) \rho_{ma} + \varphi \rho_f \quad (۳-۱)$$

بنابراین با دانستن چگالی خمیره و چگالی سیال سازند با استفاده از  $\rho_b$  حاصل از نگار، می‌توانیم تخلخل را محاسبه کنیم. مانند نگار صوتی، در نگار چگالی نیز اثرات مربوط به ناخالصی‌های شیلی، اثر گاز و اثرات محیطی باید اصلاح شود [۱۳].

### ۱-۲-۳- نگار پرتو گامای طبیعی

موادی که از خود پرتوهای  $\alpha$ ،  $\beta$  و  $\gamma$  گسیل می‌دارند را مواد رادیواکتیو گویند. ماده ممکن است به طور طبیعی از خود پرتو گسیل کند، که در این صورت به آن رادیواکتیو طبیعی و یا با پرتو دهی توسط عناصر دیگر به ماده رادیواکتیو تبدیل شود که به آن رادیواکتیو مصنوعی گویند. از جمله رادیواکتیوهای طبیعی با ارزش در مطالعات چاه‌پیمایی می‌توان به عناصر پتاسیم، توریوم و اورانیوم اشاره کرد. پرتوی گاما به دلیل انرژی زیادی که دارد در چاه‌نگاری نقشی اساسی ایفا می‌کند [۱۳].

با توجه به اینکه اورانیوم، توریوم و پتاسیم به مقدار زیادی در کانی‌های رسی تمرکز می‌یابند جهت برآورد حجم شیل سازندهای غیر تمیز رسوبی از این نمودار می‌توان کمک گرفت.

واحد قدیمی ثبت پرتو گاما بر حسب گرم رادیوم بر تن (Gr Ra Eq/Ton) یا رونتگن بر ساعت (r/h) است که معادل گرم رادیوم بر گرم سنگ نیز است. امروزه به افتخار انستیتوی نفت آمریکا<sup>۱</sup> واحد

A.P.I را به کار می‌برند که به علت بزرگ بودن از  $\frac{1}{200} API$  استفاده می‌شود که به آن GAPI می‌گویند [۱۳].

---

1. American Petroleum Institute

## ۱-۲-۲-۴- نگارهای مقاومت ویژه الکتریکی

یکی از متغیرهای فیزیکی که اغلب در نمودارگیری مد نظر قرار می‌گیرد و در ارزیابی مخازن هیدروکربنی و بررسی سازندهای آبدار مورد توجه می‌باشد، مقاومت ویژه الکتریکی است. به طور کلی به منظور بررسی مقاومت الکتریکی سازند، یکی از دو دستگاه زیر را می‌توان به کار برد:

۱. دستگاه‌های با سیستم الکترودی

۲. دستگاه‌های با سیستم القایی (دارای سیم پیچ)

در سیستم الکترودی، از طریق الکترودها (توسط چشمه‌ی جریان) جریان به درون سازند ارسال و اختلاف پتانسیل توسط یک سری الکتروود دیگر اندازه‌گیری می‌شود و نهایتاً نمودار مقاومت ویژه الکتریکی ثبت می‌شود. برخی از نمودارهای الکترودی مقاومت ویژه عبارتند از [۹ و ۱۳]:

- نمودارهای مقاومت ویژه‌ی غیر متمرکز<sup>۱</sup> که به نمودارهای مقاومت ویژه‌ی الکتریکی نرمال یا معمولی<sup>۲</sup> و جانبی<sup>۳</sup> تقسیم‌بندی می‌شوند.
- نمودارهای مقاومت ویژه‌ی متمرکز که شامل نمودارهای لاترولوگ<sup>۴</sup>، لاترولوگ متمرکز کروی<sup>۵</sup> و لاترولوگ دوگانه<sup>۶</sup> می‌باشد.

در اینجا به توضیح دو آرایش الکترودی نرمال و جانبی می‌پردازیم:

۱. روش مقاومت ویژه‌ی نرمال: در این روش مطابق شکل ۱-۱ یکی از الکترودهای پتانسیل (M) و یکی از الکترودهای جریان (A) در داخل چاه قرار می‌گیرند و دو الکتروود دیگر (B,N) در سر چاه (خارج از چاه) نصب می‌گردند. فاصله‌ی بین دو الکتروود A و M را بازه‌ی الکتروودها-

1. Unfocused

2. Normal resistivity log (NL)

3. Lateral resistivity log (LL)

4. Laterolog

5. Spherical focusing laterolog

6. Dual laterolog

می‌نامند. بر حسب اندازه‌ی این بازه، سوند نرمال به کوتاه و بلند تفکیک می‌شود که در سوند

با بازه‌ی کوتاه، فاصله‌ی الکترودی ۱۶ اینچ و در بازه‌ی بلند ۶۴ اینچ است [۱۳].

۲. روش مقاومت ویژه‌ی جانبی: در این روش هر دو الکتروود پتانسیل (M,N) در داخل چاه و

پایین‌تر از الکتروود جریان A قرار دارند. الکتروود جریان دیگر (B) در سر چاه (خارج از چاه)

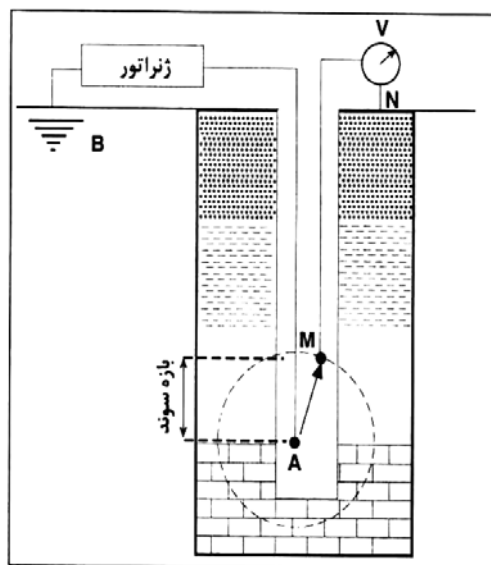
نصب می‌گردد (شکل ۱-۲). فاصله الکتروود جریان تا وسط الکتروودهای پتانسیل (O) به نام

فاصله‌ی الکتروودی تعریف می‌شود و بر حسب بزرگی این بازه نمودارهای جانبی به کوتاه و

بلند نام‌گذاری می‌شوند. بسته به اینکه بازه چقدر باشد عمق نفوذ جریان به داخل سازند

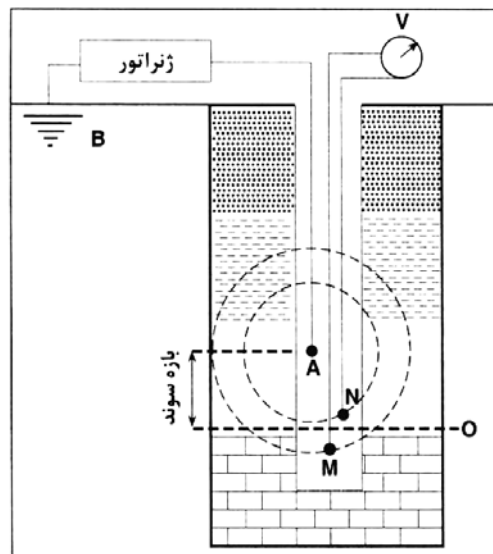
تغییر می‌کند. هرچه عمق نفوذ جریان بیشتر باشد، اطلاعات ثبت شده به اندازه‌های واقعی

نزدیک‌تر می‌شود. بنابراین در روش نرمال عمق نفوذ جریان کمتر از جانبی است [۱۳].



شکل ۱-۱: تصویری از عملکرد یک ابزار مقاومت نرمال [۹].





شکل ۱-۲: نمایی از عملکرد یک ابزار مقاومت جانبی [۹].

در نمودارهای الکتریکی هر قدر بازه‌ی سوند بیشتر باشد، عمق بررسی سوند افزایش می‌یابد. برای مثال یک بازه‌ی سوند بسیار کوچک ممکن است تنها گل موجود در چاه را بررسی کند. هر چه عمق نفوذ جریان بیشتر باشد، اطلاعات ثبت شده به اندازه‌های واقعی نزدیکتر می‌شود [۱۳].

برای آن که هرچه بهتر لایه‌های نازک را بررسی کنیم و ناحیه‌ی آغشته شده به آب گل حفاری را از ناحیه‌ی بکر و دست نخورده تشخیص دهیم و همچنین برای به دست آوردن مقاومت ویژه‌ی واقعی و بالاخره تعیین حد و مرز طبقات نازک از نمودارگیرهای تمرکز دهنده‌ی جریان استفاده می‌کنیم. در نمودارهای لاترولوگ جریان به صورت یک صفحه با ضخامت کم ایجاد می‌گردد و به این دلیل سازندهای نازک که در برابر دستگاه قرار می‌گیرند شناسایی می‌شوند و از به هدر رفتن جریان در داخل گل حفاری ممانعت به وجود می‌آید [۹ و ۱۳].

در دستگاه‌های با سیستم القایی، میدان الکترومغناطیسی توسط سیم پیچ اولیه به درون گمانه ارسال می‌گردد که این میدان جریانی را در سازند القا می‌کند. جریان القایی سبب تشکیل میدان الکترومغناطیسی ثانویه می‌شود که به وسیله‌ی سیم پیچ گیرنده ثبت و تبدیل به هدایت ویژه‌ی الکتریکی می‌گردد. برخی از نمودارهای القایی عبارتند از:

- نمودار ۶ سیم پیچ با بازه‌ی الکترودی ۲۸ اینچ

- نمودار ۶ سیم پیچ با بازه‌ی الکترودی ۴۰ اینچ

- نمودار جریان القایی ۲ گانه<sup>۱</sup>

روش القایی در چاه‌های حفاری شده با گل‌های پایه روغنی، چاه‌های فاقد گل و حتی چاه‌های جداره-گذاری شده قابل استفاده است [۲]، در صورتی که استفاده از نمودارگیرهای با سیستم الکترودی به دلیل نیاز آن‌ها به یک میانجی رسانا امکان پذیر نیست. این ابزارها نیز قادرند مقاومت در عمق کم و زیاد سازند را اندازه‌گیری کنند. در ادامه به تعدادی از روابط و نمودارهایی که برای محاسبه اشباع از آب کاربرد دارند اشاره می‌شود.

### ۱-۲-۳- رابطه‌ی آرچی<sup>۲</sup>

مقدار اشباع از آب با کمک مقادیر مقاومت ویژه و تخلخل در رابطه آرچی محاسبه می‌شود. آرچی (۱۹۴۲) نشان داد در سنگ‌های با درجه‌ی اشباع ۱۰۰٪ آب، مقاومت ویژه سنگ به تخلخل ( $\phi$ )، مقاومت ویژه‌ی آب ( $R_w$ ) و ساختار منافذ بستگی دارد. ارتباط این پارامترها در رابطه‌ی آرچی به خوبی نشان داده شده است.

ارتباط ضریب مقاومت ویژه سازند با تخلخل توسط رابطه‌ی تجربی زیر که برای اولین بار توسط آرچی (۱۹۴۲) بیان گردید، بیان می‌شود [۳۶]:

$$F = \frac{R_o}{R_w} = \frac{a}{\phi^m} \quad (۴-۱)$$

---

1. Dual induction log (DIL)

2. Archie equation

که در این رابطه:

$R_o$  مقاومت ویژه کل سازند با اشباع ۱۰۰٪ آب،  $R_w$  مقاومت ویژه‌ی آب سازند،  $a$  ضریب ثابتی که تابعی از جنس سنگ و پیچاپیچی مسیر حرکت سیال است،  $\phi$  تخلخل و  $m$  ضریب سیمان شدگی است. همچنین آرچی نشان داد هرگاه هیدروکربن وجود داشته باشد حجم آب موجود در منافذ کاهش و مقاومت ویژه با توجه به مقدار هیدروکربن افزایش می‌یابد. مقاومت ویژه‌ی سازند در حضور هیدروکربن توسط رابطه‌ی زیر محاسبه می‌شود [۳۶]:

$$I = \frac{R_t}{R_o} = S_w^{-n} \quad (5-1)$$

در این رابطه:

$I$  شاخص مقاومت ویژه،  $R_t$  مقاومت ویژه‌ی واقعی سازند،  $R_o$  مقاومت ویژه‌ی سنگ با اشباع ۱۰۰٪ آب،  $S_w$  درصد اشباع از آب و  $n$  توان اشباعی سازند است.

ترکیب دو رابطه (۴-۱) و (۵-۱)، رابطه‌ی معروف آرچی را به ما می‌دهد [۳۶]:

$$R_t = a \times R_w \times \phi^{-m} \times S_w^{-n} \quad (6-1)$$

یا

$$S_w = \left( \frac{a \times R_w}{R_t \times \phi^m} \right)^{\frac{1}{n}} = \left( \frac{F \times R_w}{R_t} \right)^{\frac{1}{n}} \quad (7-1)$$

آنچنان که واضح است، دقت این رابطه به دقت پارامترهای ورودی  $F$ ،  $R_w$  و  $R_t$  بستگی دارد. ضریب  $a$  معمولاً از ۰/۶۲ تا ۱/۶۵ متغیر است. برای ماسه سنگ سخت نشده (تراکم کم) ۰/۶۲ و برای ماسه

سنگ شیل دار این عدد،  $1/65$  در نظر گرفته می‌شود. این ضریب را معمولاً  $1$  در نظر می‌گیرند.  $R_w$  به میزان شوری و دما بستگی دارد. نگرانی اصلی، تعیین این پارامتر در صورت نبود نتایج آزمایش‌های بهره‌برداری و نگارهای SP می‌باشد که معمولاً با چنین شرایطی در میادین نفتی مواجه هستیم. به علاوه اگر جنس سازند مشخص نباشد در مقدار تخلخل نیز خطا ایجاد می‌شود. در نهایت اطمینانی در مورد توان اشباعی سازند و ضریب سیمان شدگی مورد استفاده نمی‌توانیم داشته باشیم.  $R_t$  از نگارهای مقاومت ویژه و تخلخل با استفاده از نگار نوترون و یا نگار صوتی محاسبه می‌شود. مقدار  $n$  بین  $1/7$  و  $2/2$  در نوسان است ولی مناسب‌ترین مقدار آن  $2$  می‌باشد [۹، ۱۳، ۲۲ و ۳۶].

دو روش گرافیکی برای تعیین اشباع از آب ( $S_w$ ) وقتی مقدار  $R_w$  ثابت و نامعلوم است وجود دارد:

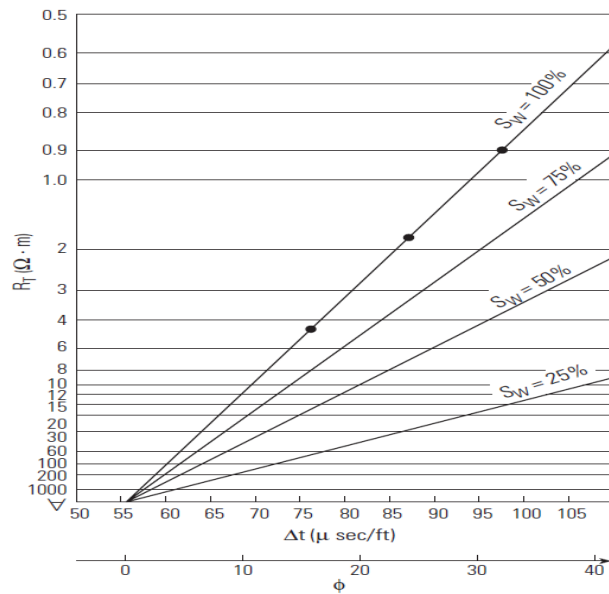
### ۱-۲-۴- نمودار هینگل<sup>۱</sup>

اولین روش برای جبران نقایص رابطه‌ی آرچی در سال ۱۹۵۹ توسط هینگل ارائه شد. در این روش فرض بر این است که اندازه‌گیری‌های تخلخل معلوم است (در اینجا  $\Delta t$ ) هرچند ممکن است جنس سازند نامعلوم باشد. نموداری را که می‌توان از این طریق ترسیم نمود مقادیر تخلخل و اشباع از آب را مستقیماً تعیین می‌کند. اگر به رابطه‌ی (۷-۱) نگاه کنیم متوجه می‌شویم به ازای  $m = n = 2$  و  $a = 1$  در صورت ثابت بودن مقدار اشباع و مقاومت ویژه آب، تخلخل متناسب با  $\frac{I}{\sqrt{R_t}}$  تغییر می‌کند.

این موضوع منجر به ترسیم شکل ۱-۳ می‌شود [۲۲].

---

1. Hingle plot



شکل ۳-۱: نمودار هینگل [۲۲].

رابطه‌ی (۷-۱) را می‌توان به شکل زیر نوشت [۲۲]:

$$\frac{I}{\sqrt{R_t}} = S_w \frac{I}{\sqrt{R_w}} \phi \quad (۸-۱)$$

همانطور که از شکل ۳-۱ دیده می‌شود به ازای اشباع از آب ۱۰۰٪، نقاط روی خط با بیشترین شیب قرار می‌گیرند. نقاط با درصد اشباع کمتر و با مقدار ثابت تخلخل دارای مقاومت ویژه‌ی بالاتر هستند و در نتیجه زیر این خط قرار می‌گیرند.

### ۱-۲-۵- نمودار پیکت<sup>۱</sup>

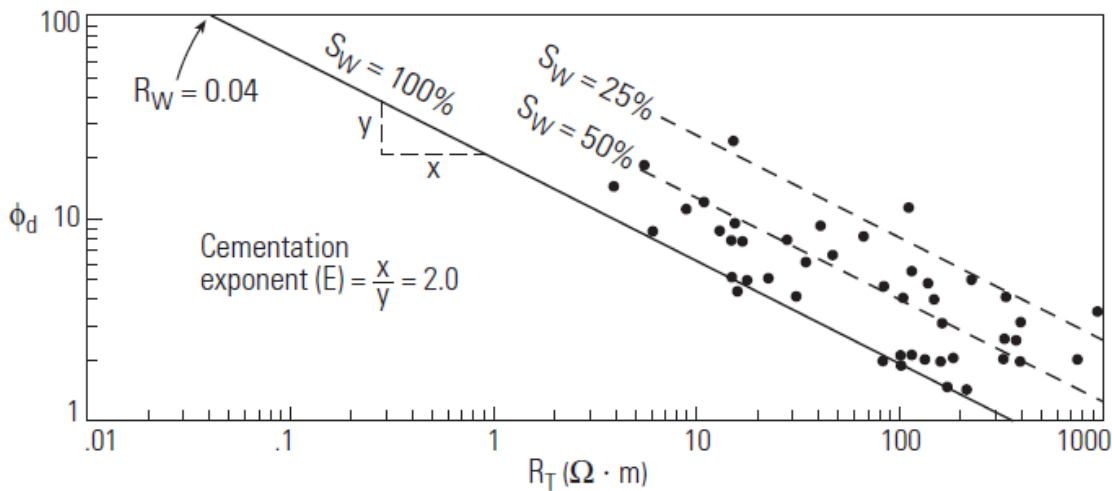
روش گرافیکی کارآمد دوم، روش ارائه شده توسط پیکت (۱۹۶۳) می‌باشد. در این روش اطلاع داشتن از تخلخل ضروری است و در مقابل مقادیر  $m$ ،  $R_w$  و  $S_w$  را می‌توان به دست آورد. رابطه‌ی (۷-۱) در

1. Pickett plot

مقیاس لگاریتمی ترسیم می‌شود و مورد بهره‌برداری قرار می‌گیرد. با لگاریتم گرفتن از دو طرف رابطه (۷-۱) و مرتب کردن آن به رابطه‌ی زیر می‌رسیم [۲۲]:

$$\log(\phi) = -\frac{I}{m} \log(R_t) + \frac{I}{m} (\log(a) + \log(R_w) - n \log(S_w)) \quad (۹-۱)$$

بنابراین وقتی اشباع از آب مقدار ثابتی باشد، نمودار لگاریتمی تخلخل بر حسب  $R_t$  یک خط صاف با شیب منفی خواهد بود که این شیب مقدار ضریب سیمان شدگی را تعیین می‌کند و مقدار آن نزدیک ۲ می‌باشد (شکل ۴-۱).



شکل ۴-۱: نمودار پیکت [۲۲].

اگر مقدار  $a$  را ۱ در نظر بگیریم رابطه (۹-۱) به صورت زیر تبدیل می‌شود [۲۲]:

$$\log(\phi) = -\frac{I}{m} \log(R_o) + \frac{I}{m} \log(R_w) \quad (۱۰-۱)$$

این رابطه خط مربوط به اشباع شدگی ۱۰۰٪ را نشان می‌دهد. محل تلاقی این خط با نقطه‌ی با تخلخل ۱۰۰٪،  $R_w$  را مستقیماً به ما می‌دهد. برای مقادیر  $S_w$  کمتر از ۱۰۰٪ ارتباط بین  $R_t$  و  $\phi$  به

صورت خطوط موازی با خط مربوط به اشباع ۱۰۰٪ است با این تفاوت که به سمت راست جابجا شده‌اند [۲۲].

آنچه تاکنون بیان شد مربوط به تعیین اشباع‌شدگی در سازندهای تمیز (فاقد شیل) می‌باشد که رابطه‌ی آرچی در آن‌ها صادق است. این رابطه به دلیل سادگی نقایصی دارد که از جمله‌ی آن‌ها می‌توان به عدم کاربردی بودن آن در سازندهای شیلی و یا غیر همگن اشاره کرد. همچنین ثابت‌های مورد نیاز باید برای هر مسأله‌ی خاص تعیین شوند. هیچ نظریه‌ی مشخصی برای تعیین اشباع از آب در سازندهای شیلی وجود ندارد. روابط متعددی پیشنهاد شده‌اند که قسمتی از آن‌ها منطبق بر مشاهدات تجربی محلی هستند و از اعتبار محدودی برخوردارند.

در سازندهای کربناته و سنگ‌های حاوی حفرات غیر همگن، این مشکل به گونه‌ی دیگری وجود دارد. در این نوع سنگ‌ها روش‌های تعیین اشباع روشن و صریحی ارائه نشده است. اگر هم روش مشخصی وجود داشته باشد، اندازه‌گیری نسبت انواع حفرات در این سازند به مراتب دشوارتر از اندازه‌گیری حجم شیل در ماسه سنگ‌های شیلی می‌باشد [۲۲].

در کل برای هر دو نوع سازند روابط متعددی پیشنهاد شده است که از ذکر جزئیات در مورد آن‌ها امتناع می‌کنیم.

### ۱-۳- کاربرد روش شبکه‌های عصبی مصنوعی در تخمین مقدار اشباع از آب

#### سنگ مخزن

روش‌هایی که به طور معمول از آن‌ها برای تعیین اشباع از آب و پارامترهای دیگر مخزنی، از نگارهای چاه استفاده می‌شود، شامل فرض‌ها و شرایطی هستند که در صورت اقناع آن‌ها، نتایج حاصل صحیح است و در غیر این صورت می‌بایست تصحیحات لازم به روشی مناسب اعمال شود. همچنین روش‌ها

و معادلاتی که از آن‌ها در این روابط استفاده می‌شود، شامل ضرایب و عباراتی هستند که به نقطه قرائت و سنگ‌شناسی آن بستگی دارد و به‌طور کلی ناشناخته‌اند و بایستی با تحلیل نمونه‌های سنگی تعیین شوند [۸ و ۳۰].

چنانچه از شبکه عصبی برای تخمین اشباع از آب استفاده شود، حداقل می‌توان گفت بعد از آن که شبکه‌ی مناسبی طراحی شد، در مرحله‌ی اجرا و کاربرد، ما را از فرمول‌ها و روابطی که نیاز به تعیین ضرایب و عبارات لازم دارند و همچنین تصحیحات ناشی از عدم تحقق فرض‌ها و شرایط حاکم بر روش مورد استفاده بی‌نیاز می‌کند، زیرا شبکه این موارد را در مرحله‌ی آموزش آموخته و در مرحله‌ی تعمیم، توانایی شبکه در اعمال آن‌ها مورد آزمون قرار گرفته است [۸].

در مواردی که راه‌حل‌های متعارف می‌توانند مسأله را به‌طور مؤثر حل کنند، لزومی ندارد که از روش شبکه‌ی عصبی یا هر روش هوش مصنوعی دیگر استفاده کرد، اما مسائل متعددی وجود دارند که برای آن‌ها، این روش، راه‌حلی منحصر به فرد است. روش شبکه‌ی عصبی دارای پتانسیل بالایی برای تولید نتایج تقریباً دقیق از پایگاه داده‌های بزرگ می‌باشد و در شرایط زیر از قابلیت و توانایی مناسبی نسبت به سایر روش‌ها برخوردار است و می‌تواند به‌طور مؤثر به کار رود:

- ۱- هنگامی که داده‌ها خطای زیادی دارند و قطعیت آن‌ها کم است.
- ۲- الگوی تصمیم‌گیری، بسیار پیچیدی و مبهم است، به‌طوری که مغز انسان و روش‌های آماری استاندارد قادر به استنباط آن‌ها نیست.

۳- داده‌ها کاملاً غیرخطی، آشفته و نابسامان هستند به گونه‌ای که همه‌ی پارامترهای دخیل در یک فرآیند بخصوص شناخته شده نیستند و یا اینکه رابطه‌ی پارامترها برای مدل کردن سیستم بسیار پیچیده است و در نتیجه امکان استفاده از مدل‌سازی ریاضی وجود ندارد. داده‌های مربوط به علوم



زمین، در اغلب موارد، چنین شرایط و پیچیدگی‌هایی را دارند و به همین دلیل در این علوم از روش شبکه عصبی مصنوعی به طور روزافزون استفاده می‌شود [۱۷ و ۱۸].

ویژگی‌ها و مزایایی از روش شبکه‌ی عصبی که آن را در مواجهه با داده‌هایی با شرایط مذکور توانمند و ممتاز می‌سازد عبارتند از:

۱- قابلیت یادگیری: قابلیت یادگیری یعنی توانایی تنظیم پارامترهای شبکه وقتی که محیط شبکه تغییر می‌کند و شبکه شرایط جدید را تجربه می‌کند، با این هدف که اگر شبکه برای یک وضعیت خاص آموزش دید و تغییر کوچکی در شرایط محیطی آن (وضعیت خاص) رخ داد، شبکه بتواند با آموزش مختصر برای شرایط جدید نیز کارآمد باشد. دیگر اینکه اطلاعات در شبکه‌های عصبی در پارامترها ذخیره می‌گردد و هر نرون در شبکه، به صورت بالقوه از کل فعالیت سایر نرون‌ها متأثر می‌شود. در نتیجه اطلاعات از نوع مجزا از هم نبوده بلکه متأثر از کل شبکه می‌باشد [۱۴].

۲- پراکندگی اطلاعات در تمام شبکه: اطلاعات و دانشی که شبکه فرا می‌گیرد، در پارامترهای شبکه مستتر می‌باشد. رابطه‌ی یک به یک بین ورودی‌ها و پارامترها وجود ندارد. می‌توان گفت که هر پارامتر مربوط به همه‌ی ورودی‌هاست ولی به هیچ یک از آن‌ها به طور منفرد و مجزا مربوط نیست. به عبارت دیگر هر نرون در شبکه، از کل فعالیت سایر نرون‌ها متأثر می‌باشد. بر این اساس چنانچه بخشی از نرون‌های شبکه حذف شوند و یا عملکرد غلط داشته باشند باز هم احتمال رسیدن به پاسخ صحیح وجود دارد. اگرچه این احتمال برای تمام ورودی‌ها کاهش یافته ولی برای هیچ یک از بین نرفته است [۱۴].

۳- قابلیت تعمیم<sup>۱</sup>: پس از آنکه مثال‌های اولیه به شبکه داده شد، شبکه می‌تواند در مقابل یک ورودی آموزش داده نشده قرار گیرد و یک خروجی مناسب ارائه نماید. این خروجی بر اساس مکانیسم تعمیم

که همانا چیزی جز فرآیند درون‌یابی<sup>۱</sup> نیست، به دست می‌آید. به عبارت روشن‌تر، شبکه تابع را یاد می‌گیرد، الگوریتم را می‌آموزد و یا رابطه‌ی تحلیلی مناسبی را برای تعدادی نقاط در فضا به دست می‌آورد [۱۴].

۴- پردازش موازی: هنگامی که شبکه عصبی در قالب سخت افزار پیاده می‌شود، نرون‌هایی که در یک تراز قرار می‌گیرند، می‌توانند به طور همزمان به ورودی‌های آن تراز پاسخ دهند. این ویژگی باعث افزایش سرعت پردازش می‌شود. در واقع در چنین سیستمی، وظیفه‌ی کلی بین پردازنده‌های کوچک-تر توزیع می‌گردد [۱۴].

۵- مقاوم بودن: در یک شبکه عصبی، رفتار کلی شبکه، برآیند رفتارهای محلی نرون‌های متعدد است. این ویژگی باعث می‌شود تا خطاهای محلی در خروجی نهایی تأثیر چشمگیر نداشته باشد. به عبارت دیگر، نرون‌ها در یک روند همکاری، خطاهای محلی یکدیگر را تصحیح می‌کنند. این خصوصیت باعث افزایش قابلیت مقاوم بودن<sup>۲</sup> و تحمل‌پذیری خطا در سیستم می‌گردد [۱۴].

## ۱-۴- مروری بر مطالعات انجام شده

مطالعات زیادی در زمینه تخمین پارامترهای پتروفیزیکی توسط شبکه‌های عصبی مصنوعی صورت گرفته است که در اینجا به تعدادی از آن‌ها اشاره می‌کنیم. اولسون (۱۹۹۸) تخلخل و تراوایی را در میدان بزرگ گازی هوگوتن<sup>۳</sup> در کانزاس<sup>۴</sup> توسط شبکه‌های عصبی مصنوعی تخمین زده است [۳۸]. هلی<sup>۵</sup> و همکارانش تخلخل و تراوایی را در یکی از مخازن دریای شمال با استفاده از شبکه‌ی عصبی مصنوعی تخمین زده‌اند [۲۹]. هوانگ<sup>۶</sup> و همکارانش از روش شبکه‌ی عصبی مصنوعی به منظور

---

1. Interpolation  
5. Helle

2. Robustness  
6. Huang

3. Hugoton

4. Kansas

تخمین تراوایی و تخلخل در میدان گازی ونچر در شرق کانادا بهره برده‌اند [۳۲ و ۳۳]. البولوچی<sup>۱</sup> (۲۰۰۷) و همکارانش از داده‌های یک سازند ماسه سنگی در عمان برای تخمین اشباع از آب استفاده کرده‌اند [۱۶]. در ایران نیز افراد زیادی در این زمینه در قالب پایان‌نامه و مقاله به تحقیق پرداخته‌اند. از جمله آن‌ها می‌توان به تخمین خواص پتروفیزیکی با استفاده از شبکه‌ی عصبی توسط تدینی و همکاران (۱۳۸۳) [۴]، تخمین نفوذپذیری توسط حبیبیان و همکاران (۱۳۸۴) [۵] و تخمین تخلخل، تراوایی و اشباع از آب توسط نصرآبادی و همکاران (۱۳۸۲) [۱۵]، اشاره کرد.

با توجه به اینکه امکان تخمین ۱۰۰٪ در شبکه‌های عصبی مصنوعی وجود ندارد، این امر انگیزه کافی را به ما می‌دهد تا در جستجوی راه‌هایی برای افزایش قابلیت این گونه شبکه‌ها باشیم. سیستم‌های چند شبکه‌ای که از آن‌ها به عنوان شبکه‌های کمیته‌ای یا ماشین‌های کمیته‌ای یاد می‌شود گوشه‌ای از این تلاش‌ها است و می‌توان گفت ترکیبی از چند شبکه هستند که به صورت مجزا آموزش دیده‌اند.

پارامنتو و همکارانش (۱۹۹۶) روش میانگین‌گیری ساده شبکه‌های مختلف با مجموعه‌ی آموزشی یکسان و پارامترهای اولیه‌ی گوناگون را نسبت به انتخاب بهترین شبکه برای کاهش واریانس پیشنهاد دادند. نفتالی<sup>۲</sup> و همکاران (۱۹۹۷) در تحقیق خود متوجه شدند شبکه‌های کمیته‌ای واریانس خطا را کاهش می‌دهند ولی بایاس خطا بدون تغییر باقی می‌ماند. این در حالیست که هاشم<sup>۳</sup> (۱۹۹۷) ترکیبات بهینه‌ی نامقید را برای حذف بایاس پیشنهاد کرد. شارکی و همکارانش روش آزمون و انتخاب را برای انتخاب ترکیبات کارآمد پیشنهاد کرده‌اند [۱۷، ۱۸، ۲۶ و ۴۰].

میزان استفاده از شبکه‌های عصبی کمیته‌ای بسیار کمتر از شبکه‌های عصبی منفرد می‌باشد. از جمله مواردی که از این روش استفاده شده است می‌توان به رساله‌ی دکتری بات<sup>۴</sup> (۲۰۰۲) اشاره کرد که در آن به تخمین پارامترهای پتروفیزیکی از جمله تخمین مقدار اشباع از آب مخزن توسط شبکه‌های

---

1. AL-bulushi  
3. Hashem

2. Naftaly  
4. Alpana Bhatt

عصبی کمیته‌ای پرداخته است [۱۸]. چن<sup>۱</sup> و لین<sup>۲</sup> نیز از ماشین کمیته ای، متشکل از فرمول‌های تجربی برای تخمین تراوایی استفاده کرده است و الگوریتم ژنتیک را برای تعیین ضرایب شبکه‌های جزء به کار برده است [۱۹]. گوپینات<sup>۳</sup> و ردی<sup>۴</sup> (۲۰۰۰) مزیت شبکه‌های کمیته ای را نسبت به تک شبکه‌ها برای یک مورد مطالعاتی بررسی کرده‌اند [۲۵]. فانگ<sup>۵</sup> و همکارانش (۱۹۹۶) از ترکیب مدوله-ای برای تشخیص خصوصیات پتروفیزیکی از روی نگارهای چاه استفاده کرده‌اند [۲۳]. چو<sup>۶</sup> (۱۹۹۹) در یک مسأله‌ی تشخیص الگو از الگوریتم ژنتیک برای ترکیب شبکه‌های عصبی مصنوعی بهره برد [۲۰]. ذاکری، م. (۱۳۸۷) نیز در ایران برای اولین بار با استفاده از شبکه‌های عصبی کمیته ای تداخل را تخمین زده است و توانسته نتیجه خوبی در خصوص کاهش خطای تخمین به دست آورد [۸]. همچنین پایان نامه ذاکری و مقالات مستخرج از این پایان نامه، تنها منابع و مطالعاتی است که تا کنون در زمینه‌ی استفاده از شبکه‌های عصبی کمیته‌ای در تخمین یکی از پارامترهای مخزن (تداخل) در ایران ارائه شده است.

## ۱-۵- ضرورت و هدف از مطالعه‌ی حاضر

با توجه به مطالعات انجام شده، شبکه‌های عصبی مصنوعی از قابلیت بالایی برای تخمین پارامترهای پتروفیزیکی برخوردارند. استفاده از این شبکه‌ها، تا حد زیادی ما را نسبت به اطلاعات زمین شناسی و تصحیحات ضروری بی‌نیاز می‌کند.

با توجه به مواردی که قبلاً اشاره شد تخمین ۱۰۰٪ برای داده‌های جدید ممکن نیست و باید به دنبال راه‌هایی برای افزایش قابلیت این شبکه‌ها باشیم. یکی از این روش‌ها، استفاده از ترکیبات کمیته ای این شبکه‌ها می‌باشد. این رویکرد در مقابل انتخاب بهترین شبکه و در نظر نگرفتن شبکه‌های دیگر

---

1. Chen  
4. Reddy

2. Lin  
5. Fung

3. Gopinath  
6. Cho

قرار می‌گیرد و این در صورتی است که این شبکه‌های حذف شده می‌توانند حاوی اطلاعات مفیدی باشند.

از آنجا که مطالعات کمتری در زمینه‌ی قابلیت این روش نسبت به شبکه‌های عصبی مصنوعی منفرد صورت گرفته است، نقاط قوت و ضعف روش مزبور به خوبی روشن نیست. در ایران ذاکری (۱۳۸۷) برای اولین بار تأثیر این روش را در تخمین تخلخل مورد بررسی قرار داد. وی از دو روش تحلیلی هاشم و میانگین‌گیری ساده این کار را انجام داد. در اینجا ما قصد داریم قابلیت الگوریتم ژنتیک در تعیین ضرایب بهینه‌ی شبکه‌ها را در ترکیب کمیته‌ای مورد بررسی قرار دهیم و نتایج را با روش میانگین‌گیری ساده مقایسه کنیم.

## ۱-۶- ساختار پایان‌نامه

در فصل دوم مبانی شبکه‌های عصبی مصنوعی مورد بررسی قرار گرفته است به گونه‌ای که خواننده دیدی هرچند مختصر نسبت به این مقوله پیدا کند.

مبانی مربوط به شبکه‌های عصبی کمیته‌ای در فصل سوم آورده شده است. در این فصل اصول ماشین‌های کمیته‌ای و الگوریتم ژنتیک تا حد مناسبی مورد بررسی قرار گرفته است.

در فصل چهارم نتایج کار عملی در مورد تخمین اشباع از آب، در یکی از میادین جنوب غربی ایران ارائه شده است.

در نهایت در فصل پنجم که فصل آخر پایان‌نامه‌ی حاضر می‌باشد، نتایجی که از این تحقیق گرفته شده است ارائه شده و همچنین راه‌کارهایی که می‌تواند به منظور تبیین و وضوح بیشتر در آینده انجام

شود، پیشنهاد شده است. همچنین در انتهای پایان نامه حاضر، منابع مورد استفاده به صورت فهرست وار ذکر شده‌اند.

## فصل دوم

### مبانی شبکه‌های عصبی مصنوعی

## ۲-۱- مقدمه

یک شبکه عصبی را می‌توان به شکل یک پردازنده‌ی قوی در نظر گرفت که از واحدهای ساختاری به نام نرون<sup>۱</sup> تشکیل شده است و این قابلیت را دارد تا دانش تجربی را به خوبی برای کاربردهای بعدی ذخیره کند. این شبکه‌ها از دو جهت عمده به عملکرد مغز شباهت دارند: ۱- اطلاعات از طریق فرایند یادگیری حاصل می‌شوند. ۲- قدرت ارتباطی بین نرون‌ها که تحت عنوان وزن های سیناپسی شناخته شده‌اند برای ذخیره‌ی این اطلاعات استفاده می‌شوند [۱۸].

مغز انسان را می‌توان به صورت یک سیستم کامپیوتری پیچیده، غیر خطی و موازی در نظر گرفت که قابلیت‌های فکر کردن، به خاطر سپردن و حل مسأله را دارد. این عضو مهم قابلیت این را دارد که اجزای سازنده‌ی خود را (نرون‌ها) به نحو مطلوبی نظم ببخشد تا بتواند محاسبات دقیق از جمله تشخیص الگو، ادراک و غیره را به بهترین شکل انجام دهد. تلاش‌های زیادی برای تقلید از شیوه‌ی عملکرد مغز توسط مدل‌های کامپیوتری صورت گرفته است. علی‌رغم نتایج با ارزشی که حاصل شده است این تلاش‌ها در مقابل عملکرد بسیار پیچیده مغز قابل اغماض می‌باشند [۱۸].

واحد سلولی اصلی و اساسی سیستم عصبی مغز، نرون می‌باشد. نرون‌ها واحد پردازشی ساده می‌باشند که پیام‌های دریافتی را از طریق دندریت‌ها<sup>۲</sup> (مسیر ورودی به نرون‌ها) از نرون‌های دیگر دریافت و ترکیب می‌کنند. سپس پیام از طریق اکسون<sup>۳</sup> که به دندریت‌های تعداد زیادی نرون دیگر متصل است به نرون‌های دیگر منتقل می‌شود. شکل ۱-۲، نمایی از یک نرون را با اجزای مختلف آن نشان می‌دهد. هر پیامی که از دندریت‌ها به نرون وارد می‌شود از سیناپس<sup>۴</sup> یا گره‌ی سیناپسی عبور می‌کند. در واقع محل تلاقی یک اکسون از یک سلول به دندریت‌های سلول دیگر را سیناپس می‌گویند. سیناپس‌ها واحدهای ساختاری کوچکی هستند که ارتباط بین نرون‌ها را برقرار می‌کنند. سیناپس‌ها انواع

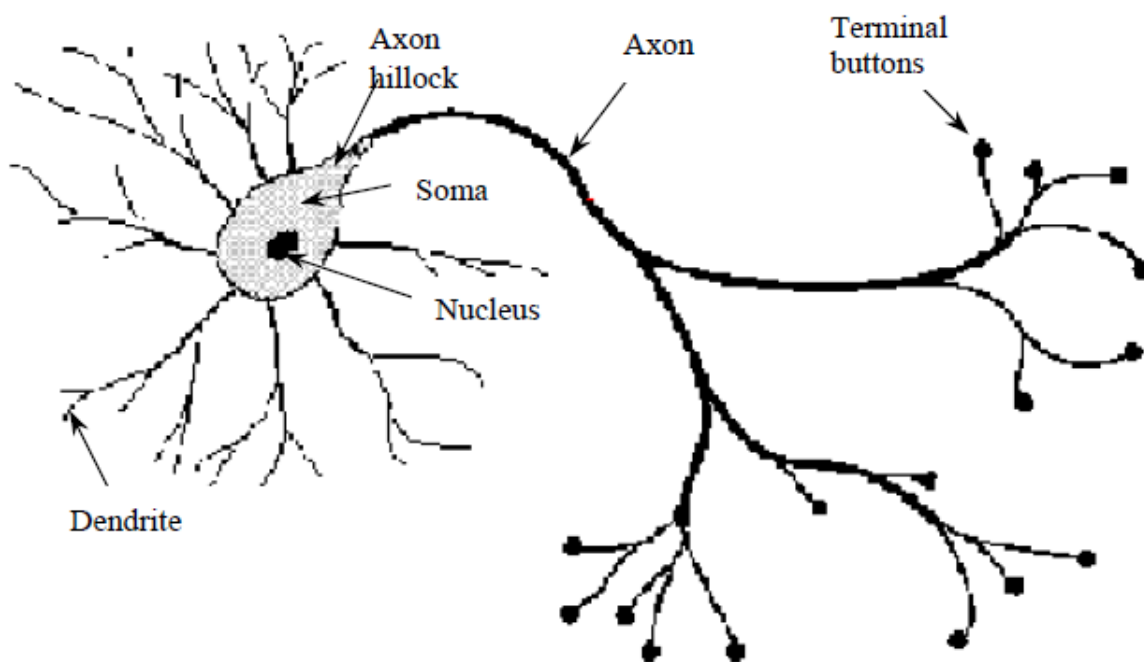
---

1. Neuron  
3. Axon

2. Dendrite  
4. Synapse



مختلفی دارند که یکی از مهمترین آنها سیناپس‌های شیمیایی هستند. به طور خلاصه آنچه که مهم است، ترتیب نرون‌ها و شدت سیناپس‌های بین دو نرون می‌باشد، که توسط یک فرایند شیمیایی خیلی پیچیده به دست می‌آید. بعضی از ساختارهای عصبی هنگام تولد ایجاد می‌شوند و قسمت‌های دیگر، در خلال یادگیری به عنوان ارتباط‌های جدید به وجود می‌آیند و یا از بین می‌روند. به عبارت دیگر هر نرون دارای توانایی تطبیق‌پذیری می‌باشد که این خود اجازه می‌دهد سیستم عصبی، خودش را با محیط اطراف وفق دهد. با دو مقوله می‌توان خاصیت تطبیق‌پذیری نرون‌ها را توجیه نمود: ایجاد ارتباط‌های جدید سیناپسی بین نرون‌ها و تغییرات در شدت و ضعف سیناپس‌های موجود. از این رو ساختارهای عصبی، در خلال تجربه‌های زندگی در حال تغییر هستند [۱۴ و ۱۸].

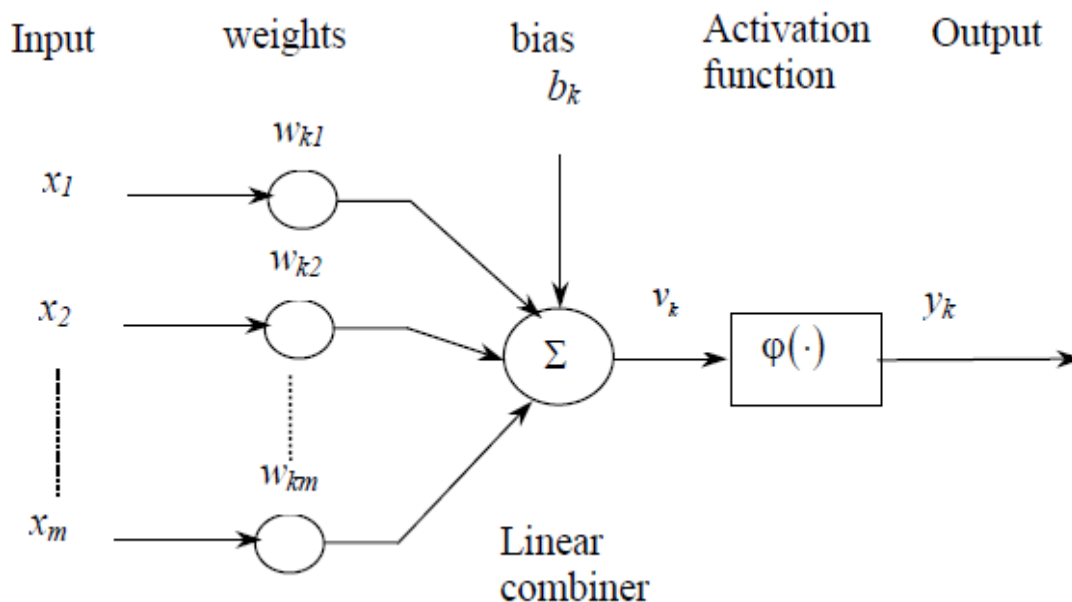


شکل ۱-۲: نمایی از یک نرون بیولوژیکی [۱۸].

## ۲-۲- ساختار شبکه‌های عصبی مصنوعی

همانطور که اشاره شد نرون واحد اساسی سیستم عصبی مغز می‌باشد. به طور مشابه نیز یک نرون اساس عملکرد شبکه‌های عصبی مصنوعی را تشکیل می‌دهد. شکل (۲-۲) مدل یک نرون را نشان می‌دهد که از اجزای زیر تشکیل شده است:

۱. در این مدل سیناپس‌ها به شکل وزن مشخص شده‌اند. سیگنال  $x_j$  در ورودی سیناپس  $j$  ام که به نرون  $k$  ام متصل است در وزن سیناپسی  $w_{kj}$  ضرب می‌شود.
۲. یک جمع کننده که سیگنال‌های وزن دار را با هم جمع می‌کند.
۳. تابع انتقال یا فعال ساز برای محدود ساختن خروجی نرون.
۴. همچنین این مدل نرونی دارای یک مقدار بایاس ( $b_k$ )، می‌باشد که بسته به اینکه منفی یا مثبت باشد ورودی تابع فعال ساز را کاهش یا افزایش می‌دهد.



شکل ۲-۲: مدل یک نرون [۱۸].

اگر بخواهیم عملکرد یک نرون را به شکل ریاضی بیان کنیم می‌توانیم بنویسیم [۱۸]:

$$y_k = \varphi(u_k + b_k) \quad (1-2)$$

که در این عبارت [۱۸]:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2-2)$$

$x_j$  سیگنال ورودی،  $w_{kj}$  وزن‌های سیناپسی نرون  $k$  ام،  $u_k$  خروجی وزن داده شده مربوط به سیگنال‌های ورودی،  $b_k$  بایاس،  $\varphi$  تابع انتقال و  $y_k$  خروجی نهایی نرون می‌باشد. مقدار درون پرانتز در رابطه‌ی (۲-۱) را با  $v$  نشان می‌دهیم.

سه نوع تابع انتقالی که معمولاً مورد استفاده قرار می‌گیرند عبارتند از [۱۸]:

۱. تابع دو مقداره‌ی حدی<sup>۱</sup>:

مقدار خروجی این تابع برابر ۰ یا ۱ است. اگر ورودی کوچکتر از صفر باشد خروجی این تابع صفر و در غیر این صورت ۱ است.

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \quad (3-2)$$

۲. تابع خطی<sup>۲</sup>:

خروجی این تابع برابر ورودی آن است. از این تابع برای مسائل تشخیص الگو و تخمین تابع استفاده می‌شود.

---

1. Hardlimit function

2. Linear function

$$\varphi(v) = v \quad (4-2)$$

۳. تابع زیگموئیدی<sup>۱</sup>:

این تابع با فرمول زیر بیان می‌شود:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (5-2)$$

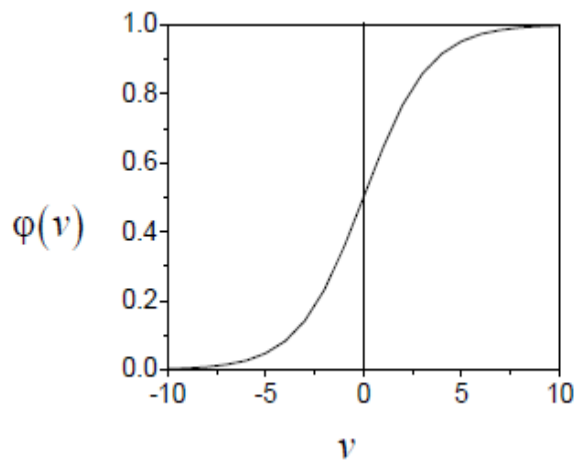
$a$  پارامتر مربوط به میزان شیب منحنی تابع زیگموئیدی است. در شبکه‌های چند لایه‌ای که با الگوریتم پس انتشار خطا آموزش می‌بینند توابع انتقال باید مشتق‌پذیر باشند بنابراین معمولاً از این تابع استفاده می‌شود. تابع زیگموئیدی یک تابع پیوسته می‌باشد که در محدوده‌ی ۰ و ۱ قرار دارد. گاهی مطلوب است که نتیجه تابع انتقال بین -۱ و ۱ باشد (متقارن نسبت به مرکز). همچنین  $\varphi(v)$  می‌تواند به صورت تانژانت هیپربولیک بیان شود [۱۸]:

$$\varphi(v) = \tanh(v) \quad (6-2)$$

لازم به ذکر است که نرون‌ها می‌توانند توابع انتقال متفاوتی داشته باشند. در جعبه ابزار برنامه MATLAB، به صورت پیش فرض برای لایه‌های میانی تانژانت هیپربولیک و برای لایه‌ی خروجی تابع خطی در نظر گرفته شده است. در شکل ۲-۳ تابع زیگموئید نشان داده شده است [۸، ۱۴ و ۱۸].

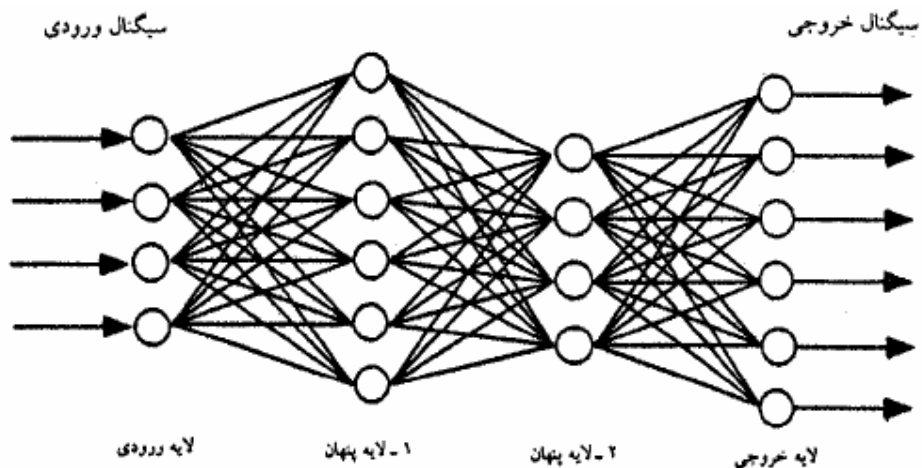
---

1. Sigmoid function



شکل ۲-۳: تابع زیگموئید [۱۸].

شبکه‌های عصبی ساختار لایه‌ای دارند و نرون‌ها در این لایه‌ها، تعبیه شده‌اند. اولین لایه، لایه ورودی است که صرفاً ورودی‌ها را در قسمت خروجی خود کپی می‌کند و کار خاص دیگری انجام نمی‌دهد. لایه آخر لایه خروجی است و لایه‌های بین لایه ورودی و لایه خروجی را لایه‌های میانی و یا پنهان می‌نامند. به جز لایه ورودی و تمام لایه‌های دیگر شبکه در کار پردازش شرکت می‌کنند. در شکل ۲-۴ نمایی از ساختار لایه‌ای شبکه‌های عصبی نشان داده شده است.



شکل ۲-۴: نمایی از ساختار لایه‌ای شبکه‌های عصبی [۸].

بر اساس نحوه‌ی اتصال گره‌ها به یکدیگر، شبکه‌های عصبی به دو گروه بزرگ تقسیم می‌شوند:

۱. شبکه‌های عصبی با تغذیه‌ی پیش‌رو<sup>۱</sup> که در آن‌ها تنها جریان یک طرفه از لایه‌ی ورودی به سمت لایه‌ی خروجی وجود دارد و هیچ مسیر برگشتی وجود ندارد. این نوع شبکه ساده‌ترین نوع شبکه عصبی است و برای مقاصد پیش‌بینی، بسیار مناسب است.
۲. شبکه‌های با تغذیه‌ی برگشتی<sup>۲</sup> که حداقل یک حلقه‌ی برگشتی فعال دارند. شبکه‌ای که خروجی هر سلول آن به تمام سلول‌های دیگر آن لایه وارد می‌شود مثالی از این دسته است. وجود حلقه‌ی برگشتی اثر زیادی بر قابلیت یادگیری شبکه می‌گذارد [۸، ۶ و ۱۸].

## ۲-۳- آموزش شبکه‌های عصبی مصنوعی

قاعده یا الگوریتم آموزش عبارت است از اصلاح وزن‌ها و بایاس‌های شبکه برای آنکه خروجی شبکه در نزدیکترین حالت نسبت به خروجی مطلوب قرار گیرد [۱۸]. الگوریتم‌های متعددی برای این امر وجود دارند که می‌توان آنها را به ۲ دسته اصلی تقسیم کرد:

۱. آموزش با ناظر<sup>۳</sup>: در این نوع آموزش، به شبکه آموخته می‌شود که بین ورودی‌ها و خروجی‌های مجموعه‌ی آموزشی ارتباط برقرار کند. این کار از طریق حداقل کردن اختلاف بین نتایج حاصل از شبکه و خروجی‌های مطلوب انجام می‌شود، بنابراین وجود مجموعه‌ی آموزشی ضروری است. بعد از تنظیم پارامترها، یک آزمون برای تعیین میزان دقت جواب‌های شبکه انجام می‌شود تا تضمین کند که شبکه، الگوی داده‌ها را شناخته است.
۲. آموزش بدون ناظر<sup>۴</sup>: در این نوع آموزش، خروجی‌های داده‌های آموزشی به شبکه عرضه نمی‌شود و پارامترها، فقط بر اساس ورودی‌ها تعیین می‌شوند. در این حالت، هدف یافتن الگویی

---

1. Feed forward networks  
3. Supervised learning

2. Feed backward networks  
4. Unsupervised learning

بین ورودی‌هاست که بتواند بر اساس خصوصیات مشابه، آنها را گروه بندی کند. روش اصلی

انجام آموزش بدون ناظر، خوشه بندی است. این شبکه‌ها را خود سازمان‌ده<sup>۱</sup> نیز می‌گویند.

انتخاب یک مجموعه آموزشی مناسب که دامنه‌ی وسیعی از حالات تحقق یافته را در برگیرد از اهمیت زیادی برخوردار است، زیرا کیفیت عملکرد نهایی شبکه‌های عصبی مستقیماً تابع مثال‌های آموزشی به کار رفته در آموزش آن است. مهم‌ترین شرط لازم برای مجموعه‌ی آموزشی آن است که تمام دامنه‌ی مقادیر ورودی‌ها و خروجی‌ها را پوشش دهد. در بسیاری از مثال‌های آموزشی، پوشش مناسب ورودی‌ها به دامنه‌ی تغییراتشان، موجب پوشش مناسب به خروجی‌ها می‌شود. همچنین بهتر است تعداد مثال‌های آموزشی برای هر محدوده از خروجی‌ها یکسان باشد. اگر شبکه‌ای به قصد پیش‌بینی مواردی با فراوانی کم طراحی می‌شود، باید مطمئن شد که مجموعه‌ی آموزشی به تعداد کافی، مثال‌هایی از موارد نادر دارد. به این دلیل، نمونه‌گیری تصادفی از داده‌های موجود نامناسب است زیرا نمونه‌هایی با فراوانی بیشتر، از حضور نمونه‌هایی با فراوانی کمتر جلوگیری می‌کنند و نقش آن‌ها را کم‌رنگ می‌سازند. برای رفع این مشکل باید از موارد نادر به طور نظام‌مند، نمونه‌های اضافی برداشت کرد. به طور کلی، مجموعه‌ی آموزشی باید به گونه‌ای انتخاب شود که الگوی صحیح را در اختیار شبکه بگذارد.

پیش پردازش ورودی‌ها و خروجی‌ها نیز به آموزش شبکه و شناخت الگوها کمک شایانی می‌کند. روش‌های مختلفی برای این کار وجود دارد. یکی از این روش‌ها، تغییر مقیاس داده‌های آموزشی به گونه‌ای است که تمام ورودی‌ها و خروجی‌ها بین صفر و یک قرار گیرند. روش دیگر تبدیل ورودی‌ها و خروجی‌ها به توزیع گوسی استاندارد است که در آن میانگین توزیع، صفر و انحراف معیار آن ۱ است. هر چه تعداد متغیرهای ورودی شبکه بیشتر شود نرون‌های میانی بیشتری لازم است تا شبکه الگوهای بین متغیرها را در حد کافی بشناسد. از طرفی با افزایش تعداد نرون‌های میانی، به مجموعه داده‌های آموزشی بیشتری نیاز است تا شبکه به جواب‌های نامناسب همگرا نشود. همچنین زمان لازم برای

---

1. Self organization

همگرایی شبکه بیشتر خواهد بود. بنابراین باید همیشه به این نکته توجه داشت که افزایش تعداد نرون‌ها باید با توجه به تعداد داده‌های آموزشی که در اختیار داریم صورت بگیرد [۶، ۱۸ و ۲۱].

## ۲-۴- الگوریتم پس انتشار خطا<sup>۱</sup>

رایج‌ترین روش آموزش با ناظر، الگوریتم پس انتشار خطا است. یادگیری این الگوریتم بر اساس قانون تصحیح خطا بنا شده است که می‌توان آن را تعمیم الگوریتم موسوم به حداقل میانگین مربعات<sup>۲</sup> دانست. قانون پس انتشار خطا از دو مسیر اصلی تشکیل می‌شود. مسیر اول به مسیر رفت موسوم است که در این مسیر، بردار ورودی به شبکه اعمال می‌شود و تأثیراتش از طریق لایه‌های میانی به لایه‌های خروجی انتشار می‌یابد. بردار خروجی تشکیل یافته در لایه‌ی خروجی، پاسخ واقعی شبکه را تشکیل می‌دهد. در این مسیر پارامترهای شبکه ثابت و بدون تغییر در نظر گرفته می‌شوند.

مسیر دوم به مسیر برگشت موسوم می‌باشد. در این مسیر، برعکس مسیر رفت پارامترهای شبکه تغییر و تنظیم می‌گردند. این تنظیم مطابق با قانون اصلاح خطا انجام می‌گیرد. سیگنال خطا، در لایه‌ی خروجی شبکه تشکیل می‌گردد.

بردار خطا برابر با اختلاف بین پاسخ مطلوب و پاسخ واقعی شبکه می‌باشد. مقدار خطا، پس از محاسبه، در مسیر برگشت از لایه‌ی خروجی و از طریق لایه‌های شبکه در کل شبکه توزیع می‌گردد. چون توزیع اخیر، در خلاف مسیر ارتباطات وزنی سیناپس‌ها صورت می‌پذیرد، کلمه پس انتشار خطا جهت توضیح اصلاح رفتاری شبکه انتخاب شده است. پارامترهای شبکه طوری تنظیم می‌شوند که پاسخ واقعی شبکه هر چه بیشتر، به سمت پاسخ مطلوب نزدیکتر شود. این فرایند تا وقتی که معیار آموزش تحقق یابد ادامه می‌یابد. چندین معیار برای پایان آموزش وجود دارد. یکی از این معیارها این است که

---

1. Error back propagation

2. Least mean square



تغییرات پارامترها بسیار کوچک شود و تقریباً، تغییری از خود نشان ندهند. معیار دیگری که می‌تواند مورد استفاده قرار بگیرد آن است که آموزش به تعداد دوره‌های خاصی مثلاً ۱۰۰۰ بار تکرار شود. هر دوره‌ی آموزش<sup>۱</sup> عبارت است از یک بار ارائه‌ی کامل همه‌ی الگوهای مجموعه‌ی آموزشی به شبکه. کاهش خطای شبکه برای داده‌های آموزش تا مقداری مشخص نیز می‌تواند معیاری برای توقف باشد.

هورنیک<sup>۲</sup> و همکارانش (۱۹۸۹) اعلام کردند که اگر یک شبکه‌ی پرسپترون سه لایه با الگوریتم آموزشی پس انتشار خطا که تابع انتقال لایه‌ی میانی آن، تابع زیگموئیدی و تابع انتقال لایه‌ی خروجی آن، تابع خطی است، در لایه‌ی میانی تعداد کافی نرون داشته باشد، می‌تواند هر تابعی با نقاط ناپیوستگی محدود را با هر دقت دلخواه تخمین بزند [۷، ۸ و ۱۸].

در محیط نرم افزار MATLAB برای آموزش شبکه‌ها با استفاده از الگوریتم پس انتشار خطا، ۱۳ تابع مختلف طراحی شده است. تمامی این توابع از شیب تابع عملکرد که معمولاً میانگین مربعات خطا می‌باشد برای تنظیم وزن‌ها و بایاس‌ها استفاده می‌کنند. پارامترهای مزبور در جهت کمینه کردن تابع عملکرد تنظیم می‌شوند. این توابع عبارتند از: تابع پس انتشار خطای کاهش گرادیان<sup>۳</sup>، تابع پس انتشار خطای کاهش گرادیان با تکانه<sup>۴</sup>، تابع پس انتشار خطا با سرعت یادگیری متغیر و با تکانه<sup>۵</sup>، تابع پس انتشار خطا با سرعت یادگیری متغیر و با تکانه<sup>۶</sup>، تابع پس انتشار خطای ارتجاعی<sup>۷</sup>، تابع شیب مزدوج فلچر - ریوز<sup>۸</sup>، تابع شیب مزدوج پولاک - ریبیرر<sup>۹</sup>، تابع شیب مزدوج پاول - بیل<sup>۱۰</sup>، تابع شیب مزدوج مقیاس شده<sup>۱۱</sup>، تابع شبه نیوتونی بی اف جی اس<sup>۱۲</sup>، تابع شبه نیوتونی متقاطع یک مرحله‌ای<sup>۱۳</sup>، تابع لونبرگ - مارکوت<sup>۱۴</sup> و تابع مرتب سازی بیزین<sup>۱۵</sup>. شرح این توابع در پیوست الف ارائه شده است.

دو روش Traingd و Traingdm گاهی بسیار کند عمل می‌کنند. همچنین دو روش Traingda و Traingdx که از سرعت متغیر استفاده می‌کنند نیز می‌باشند [۲۱].

1. Epoch	2. Hornik	3. Traingd	4. Traingdm	5. Traingda
6. Traingdx	7. Trainrp	8. Traincgf	9. Traincgp	10. Traincgb
11. Trainscg	12. Trainbfg	13. Trainoss	14. Trainlm	15. Trainbr

در مسائل تخمین تابع وقتی که تعداد پارامترها در حدود ۲۰۰-۳۰۰ باشد الگوریتم لونبرگ - مارکوت بهترین سرعت و کارایی را از خود نشان می‌دهد. در مقابل در مسائل تشخیص الگو کارایی ضعیف‌تری دارد. این الگوریتم فضای زیادی را برای محاسبه اشغال می‌کند. البته روش‌هایی برای کاهش این فضا وجود دارند که از طرف دیگر باعث افزایش زمان محاسبه می‌شوند [۲۱].

تابع Trainrp دارای بیشترین سرعت در مسائل تشخیص الگو می‌باشد ولی این الگوریتم در مسائل تخمین تابع خوب عمل نمی‌کند، اما در مقابل حافظه مورد نیاز آن در مقایسه با بقیه الگوریتم‌ها کمتر است [۲۱].

الگوریتم‌های شیب مزدوج مخصوصاً Trainscg، به نظر می‌آید که برای حل گسترده‌ی وسیعی از مسائل، به خصوص مسائلی که تعداد پارامترها در آن‌ها زیاد باشد کارایی خوبی دارند. در مسائل تخمین تابع، الگوریتم Trainscg سرعتی همانند Trainlm دارد (برای شبکه‌های بزرگ سریع‌تر است) و همچنین در مسائل تشخیص الگو به اندازه Trainrp سریع است با این تفاوت که با افزایش دقت مطلوب کاراییش به اندازه Trainrp افت نمی‌کند. الگوریتم‌های شیب مزدوج معمولاً حافظه‌ی متوسطی به کار می‌گیرند [۲۱].

کارایی Trainbfg بسیار شبیه Trainlm می‌باشد. البته نیاز آن به حافظه نسبت به Trainlm کمتر می‌باشد اما حجم محاسبات در آن با افزایش اندازه شبکه به صورت مضاعف افزایش می‌یابد [۲۱].

الگوریتم Traingdx نسبت به سایر الگوریتم‌ها کندتر و حافظه‌ی مورد نیاز آن مشابه Trainrp است اما به هر حال این الگوریتم در برخی از مسائل مفید می‌باشد. در بعضی موقعیت‌ها بهتر است که عمل همگرایی کندتر صورت بگیرد [۲۱].

## ۲-۵- بایاس<sup>۱</sup> و واریانس<sup>۲</sup>

یک شبکه‌ی عصبی می‌تواند برای تخمین تابع مورد استفاده قرار بگیرد. اگر برای ورودی  $\bar{x}$ ، تابع هدف،  $E(\bar{x})$  و تخمین شبکه  $F(\bar{x})$  باشد، میانگین مربعات خطای تخمین  $F$  به صورت زیر محاسبه می‌شود [۱۸]:

$$E \left[ (F(\bar{x}) - E(\bar{x}))^2 \right] \quad (۷-۲)$$

در این جا  $E$  عملگر امید ریاضی است. اگر این رابطه را به صورت بایاس و واریانس تفکیک کنیم به رابطه‌ی زیر می‌رسیم [۱۸]:

$$E \left[ (F(\bar{x}) - E(\bar{x}))^2 \right] = \underbrace{E \left[ F(\bar{x}) - E(\bar{x}) \right]^2}_{MSE} + \underbrace{E \left[ (F(\bar{x}) - E(\bar{x}) - E \left[ F(\bar{x}) - E(\bar{x}) \right])^2 \right]}_{Variance} \quad (۸-۲)$$

بایاس، میانگین خطای تخمین و واریانس، پراکندگی خطا حول بایاس است. بایاس می‌تواند مثبت یا منفی باشد ولی واریانس همواره مثبت است. معادله‌ی فوق را می‌توان به صورت زیر نوشت [۱۸]:

میانگین مربعات خطا = مجذور بایاس خطا + واریانس خطا

مجذور انحراف معیار همان واریانس است. بنابراین:

میانگین مربعات خطا = مجذور بایاس خطا + مجذور انحراف معیار

---

1. Bias

2. Variance

## ۲-۶- بهبود توان تعمیم شبکه

یکی از مشکلات موجود در آموزش شبکه‌های عصبی بیش‌برازش<sup>۱</sup> می‌باشد. منظور از بیش‌برازش این است که پس از آموزش شبکه، خطا بر روی مجموعه‌ی آموزشی به حداقل خود می‌رسد اما با ارائه‌ی داده‌های جدید به شبکه میزان خطا به مقدار قابل توجهی بالا می‌رود. در واقع شبکه بر روی داده‌های آموزشی احاطه‌ی کامل می‌یابد اما نمی‌تواند در قبال داده‌های جدید عکس‌العمل مناسب نشان دهد. در اصطلاح می‌توان گفت شبکه عمومیت کافی ندارد.

یک روش برای بهبود توان تعمیم شبکه استفاده از شبکه‌ای با ابعاد مناسب است که برازش کافی را برای داده‌های آموزشی در بالاترین سطح ایجاد نماید. متأسفانه انتخاب این ابعاد مناسب کار بسیار دشواری می‌باشد. بنابراین باید راه‌های تازه‌ای جستجو کرد. دو راه حل برای افزایش عمومیت شبکه در جعبه ابزار شبکه‌های عصبی در نظر گرفته شده است: ۱- مرتب‌سازی<sup>۲</sup>. ۲- توقف سریع<sup>۳</sup> [۱۸ و ۲۱]. در ادامه به بررسی این دو روش خواهیم پرداخت.

## ۲-۶-۱- آموزش به روش مرتب‌سازی

این روش شامل اصلاح توابع کارایی که معمولاً میانگین مربعات خطا (MSE) در نظر گرفته می‌شود می‌باشد. برای بهبود توان تعمیم می‌توان تابع کارایی را با اضافه کردن یک جمله شامل میانگین مجموع مربعات پارامترهای شبکه اصلاح نمود [۲۱]:

$$F = \gamma MSE + (1 - \gamma) MSW \quad (9-2)$$

که در آن [۲۱]:

---

1. Over fitting  
3. Early stopping

2. Regularization

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (10-2)$$

در این دو رابطه  $MSW$  میانگین مربعات پارامترهای شبکه و  $\gamma$  پارامتر مرتب‌سازی است که مقدار آن بین ۰ و ۱ تغییر می‌کند. انتخاب تابع کارایی به این شکل باعث می‌شود شبکه وزن‌ها و بایاس‌های کوچکتری داشته باشد و در نتیجه پاسخ‌های آن هموارتر و احتمال بیش‌برازش کاهش می‌یابد [۲۱].

مشکل اصلی در این روش یافتن مقدار مناسب و بهینه‌ی  $\gamma$  می‌باشد. اگر این پارامتر خیلی بزرگ در نظر گرفته شود بیش‌برازش رخ می‌دهد و اگر کوچکتر از حد مناسب انتخاب شود، شبکه قابلیت انطباق کافی نخواهد داشت. در نرم‌افزار MATLAB برای رفع این مشکل تابع Trainbr در نظر گرفته شده است. توضیح دقیق شیوه‌ی کار این تابع فراتر از این متن می‌باشد. در استفاده از این تابع هرگاه تعداد پارامترهای مؤثر شبکه و یا مجموع مربعات خطا (SSE) و یا مجموع مربعات پارامترهای شبکه (SSW) در چند تکرار به مقدار نسبتاً ثابتی برسند، شبکه همگرا شده و آموزش پایان می‌یابد [۲۱].

## ۲-۶-۲- آموزش به روش توقف سریع

در این روش داده‌های آموزش به دو دسته تقسیم می‌شوند. مجموعه‌ی اول همان مجموعه‌ی آموزشی است که برای محاسبه‌ی شیب و به روز رسانی وزن‌ها و بایاس‌ها از آن استفاده می‌شود. دسته‌ی دوم به سری اعتبارسنجی موسوم است که در تعیین پارامترها نقشی ندارند ولی در طول آموزش خطای شبکه برای آن‌ها محاسبه می‌شود. خطای مجموعه‌ی اعتبارسنجی مانند خطای مجموعه‌ی آموزشی در مراحل اولیه‌ی آموزش کاهش می‌یابد. زمانی که شبکه شروع به بیش‌برازش کند خطای مجموعه‌ی اعتبارسنجی شروع به افزایش می‌کند. وقتی که خطای اعتبارسنجی برای تعداد خاصی از تکرارها

افزایش یابد آموزش متوقف می‌شود و وزن‌ها و بایاس‌های مربوط به زمانی که این خطا حداقل بود، در نظر گرفته می‌شوند.

هر دو روش مرتب سازی و توقف سریع وقتی که به درستی به کار گرفته شوند می‌توانند قابلیت تعمیم شبکه را تضمین کنند.

نکته مهم در مورد روش توقف سریع این است که نباید آن را همراه با الگوریتمی که سریع همگرا می‌شود به کار برد. اگر این امر رخ دهد پارامترهای الگوریتم مزبور باید به گونه‌ای تنظیم شوند که همگرایی کندتر صورت بگیرد.

در الگوریتم Trainbr شبکه باید تا دستیابی به همگرایی آموزش یابد. مجموع مربعات خطا، مجموع مربعات پارامترها و تعداد پارامترهای مؤثر شبکه هنگامی که شبکه همگرا می‌شود ثابت می‌شوند.

بر اساس تجربه در مسأله‌ی تخمین تابع و زمانی که تعداد داده‌ها کم است روش مرتب سازی کارایی بهتری نسبت به روش توقف سریع دارد. دلیل این امر آن است که این روش نیازی به داده‌های اعتبارسنجی ندارد و از تمام داده‌ها استفاده می‌کند. این مزیت به خصوص وقتی که مجموعه‌ی داده‌ها کوچک است از اهمیت بیشتری برخوردار است [۲۱].

## فصل سوم:

مبانی شبکه‌های عصبی مصنوعی  
کمیته‌ای

### ۳-۱- مقدمه

تا کنون شبکه‌های عصبی برای تخمین پارامترهای گوناگون مخزنی مورد استفاده قرار گرفته‌اند. این روش به عنوان روشی ساده و دقیق برای به دست آوردن پارامترهای گوناگون مانند اشباع از آب، از نگارهای چاه مورد استفاده قرار گرفته است.

در حال حاضر شبکه پرسپترون سه لایه متشکل از یک لایه ورودی، یک لایه میانی و یک لایه خروجی که با یکی از الگوریتم‌های پس انتشار خطا (مانند لونیبرگ مارکوت) آموزش دیده است روش معمول می‌باشد. شبکه‌های عصبی به طور کامل قابلیت تعمیم به داده‌های آزمون را ندارند به این دلیل باید به دنبال روش‌هایی برای افزایش کارایی شبکه باشیم. یک شبکه چند لایه وقتی با الگوهای یکسان مکرراً آموزش می‌بیند برای تابع هدف مینیمم‌های مختلف حاصل می‌شود و شبکه مقادیر وزنی متفاوتی به دست می‌دهد. رویکرد معمول این است که شبکه‌های مختلف آموزش داده می‌شوند و شبکه‌ای که بهترین تعمیم را دارد انتخاب می‌شود. انتخاب بهترین شبکه اگر چه الگوی بین داده‌ها را مشخص می‌کند ولی باعث از دست رفتن اطلاعات و جزئیات موجود در شبکه‌های دیگر می‌شود. بنابراین باید به گونه‌ای عمل کنیم که اطلاعات را از شبکه‌های مختلف استخراج کنیم نه اینکه آنها را از دست بدهیم. این امر انگیزه‌ی اصلی برای استفاده از شبکه‌های عصبی کمیته‌ای می‌باشد به گونه‌ای که تعدادی از شبکه‌هایی که بهترین تعمیم را داشته‌اند با هم ترکیب می‌شوند و در نتیجه قدرت و دقت شبکه افزایش می‌یابد [۱۷ و ۱۸].



### ۳-۲- سیستم شبکه‌های چند گانه<sup>۱</sup>

یک شبکه چند گانه ترکیبی از شبکه‌های مختلف می‌باشد. مسائل مختلفی وجود دارند که با شبکه‌های منفرد قابل حل نیستند و یا اینکه اگر از شبکه‌های چندگانه استفاده کنیم جواب بهتری به دست می‌آوریم. در صورت استفاده از چند شبکه کارایی به مقدار قابل توجهی افزایش پیدا می‌کند. به طور معمول دو روش برای ترکیب شبکه‌ها وجود دارد، (۱) ترکیب آنسامبلی<sup>۲</sup> (۲) ترکیب مدوله‌ای<sup>۳</sup>. این دو روش با هم سازگارند به این معنی که یک سیستم چند شبکه‌ای می‌تواند شامل هر دو نوع ترکیب باشد. های‌کین<sup>۴</sup> (۱۹۹۹) سیستم‌های چند شبکه‌ای را ماشین کمیته‌ای<sup>۵</sup> نامید [۱۸].

### ۳-۲-۱- ترکیبات آنسامبلی

انگیزه‌ی اصلی برای ترکیب شبکه‌ها افزایش قابلیت تعمیم آنها می‌باشد. ترکیب آنسامبلی شبکه‌های آموزش دیده روشی است که می‌تواند به تلفیق بهتر شبکه‌های جزء و در نتیجه عملکرد بهتر شبکه منجر شود. در یک آنسامبل، هر یک از شبکه‌های جزء به تنهایی حلی را برای مسأله‌ی مورد نظر ارائه می‌دهند و نتایج به روشی مناسب با هم ترکیب می‌شوند. از آنجایی که شبکه‌های مختلف به مینیمم‌های مختلف در سطح خطا همگرا می‌شوند انتظار می‌رود ترکیب خروجی‌ها توان تعمیم را بهبود بخشد. واضح است که ترکیب شبکه‌های یکسان هیچ فایده‌ای ندارد زیرا به طور یکسان تعمیم می‌یابند. تأکید می‌شود که شبکه‌ها به صورت متفاوت یا حداقل مشابه و نه یکسان تعمیم یابند. به طور کلی شبکه‌ها با توجه به وزن‌ها، مدت زمان لازم برای همگرایی و حتی ساختار خود با یکدیگر متفاوت باشند ولی در عمل از آنجا که نتیجه برای داده‌های آزمون یکسان می‌باشد این شبکه‌ها تفاوتی با هم ندارند. در نتیجه هدف در اینجا یافتن شبکه‌هایی است که به شکل مختلف تعمیم می‌یابند. برای

---

1. Multiple networks system  
4. Haykin

2. Ensemble  
5. Committee machine

3. Modular

رسیدن به این منظور می‌توانیم یکی از موارد زیر را تغییر دهیم: وزن‌های اولیه، داده‌های آموزش، ساختار شبکه‌ها و الگوریتم آموزش.

سوال اصلی در اینجا این است که چه روشی برای آنسامبل کردن شبکه‌ها مفید است و آیا اصولاً روشی وجود دارد که به طور حتم نسبت به روش‌های دیگر نتیجه‌ی بهتری داشته باشد. در حقیقت روش‌های مختلفی برای ایجاد ساختار آنسامبلی پیشنهاد شده است. یک روش ساده این است که شرایط اولیه شبکه‌های جزء را تغییر دهیم. به عنوان مثال می‌توان به تغییر تعداد لایه‌های نهان یا الگوریتم آموزشی اشاره کرد. روش‌های دیگری نیز در غالب ایجاد تغییر در داده‌های آموزشی پیشنهاد شده است. این روش‌ها شامل نمونه‌گیری<sup>۱</sup>، استفاده از داده‌های متفاوت، استفاده از روش‌های پیش-پردازش مختلف و اعوجاج<sup>۲</sup> می‌باشد.

معروفترین روش نمونه‌گیری، روش بگینگ<sup>۳</sup> می‌باشد. در این روش سری آموزشی جدید با استفاده از نمونه‌گیری و جایگزینی از مجموعه آموزشی اصلی انتخاب می‌شوند. نتیجه سری‌های آموزشی متفاوت می‌باشند که با روش نمونه‌گیری از مجموعه آموزشی اصلی انتخاب شده‌اند. وقتی سیستم‌های پیشگو با این سری‌های آموزشی، آموزش می‌بینند و سپس با فرایند رأی‌گیری با هم ادغام می‌شوند عملکرد بهتری نسبت به سیستمی که با کل مجموعه آموزشی، آموزش دیده است خواهد داشت. در این روش می‌توان سری‌های آموزشی را که از سری آموزشی اصلی نمونه‌گیری می‌شوند به صورت مجزا در نظر گرفت، یعنی نمونه‌گیری بدون جایگزین.

روش دیگر زمانی ممکن می‌باشد که داده‌های بیش از یک گیرنده در دسترس باشند. در این شرایط یک ترکیب می‌تواند از شبکه‌هایی که هر یک با استفاده از داده‌های یک گیرنده خاص آموزش دیده، ایجاد شود. همچنین می‌توانیم داده‌ها را به روش‌های مختلف پیش پردازش کنیم و از هر کدام به شکل مجزا برای ساختن شبکه‌های جزء استفاده کنیم.

---

1. Sampling

2. Distortion

3. Bagging

از روش‌های دیگری نیز می‌توانیم استفاده کنیم که عبارت اعوجاج را می‌توانیم برای آن‌ها استفاده کنیم. در این روش‌ها اعوجاج را در مورد ورودی‌ها یا خروجی‌های مجموعه‌ی آموزشی برای ساخت ترکیبات آنسامبلی به کار می‌بریم.

آخرین روش که معمولاً از آن برای ساخت ترکیبات آنسامبلی استفاده می‌شود، روش آدابوست<sup>۱</sup> می‌باشد. آموزش با استفاده از این رهیافت شامل چند مرحله می‌باشد، در مرحله اول تمام اعضای مجموعه‌ی آموزشی از شانس یکسانی برای انتخاب شدن در سری‌های آموزشی برخوردارند. در مرحله بعدی آموزش، سری اول آموزش برای تمام مجموعه آموزشی آزموده می‌شود و احتمال انتخاب نمونه‌هایی که بد کلاس‌بندی شده‌اند تغییر می‌کند (با اگر مثال‌های آموزشی قابلیت این را داشته باشند که وزن دهی شوند، وزن این نمونه‌ها افزایش پیدا می‌کند). از آنجا که این احتمال اصلاح می‌شود، در مرحله بعد نمونه‌هایی که بد کلاس‌بندی شده‌اند شانس بیشتری برای حضور در سری آموزشی جدید را دارند. این روند برای تعداد مراحل خاصی ادامه پیدا می‌کند. نتایج نشان داده روش آدابوست نسبت به روش بگینگ نتایج بهتری را به همراه داشته است. نظر بریمن<sup>۲</sup> (۱۹۹۸) [۴۰] این است که روش آدابوست از اصول متفاوتی نسبت به روش بگینگ و روش‌های اعوجاج برخوردار می‌باشد [۴۰].

علی‌رغم اینکه روش‌های زیادی برای ترکیب آنسامبلی پیشنهاد شده است هنوز تصویر روشنی از اینکه کدام روش بهترین می‌باشد وجود ندارد. تا قسمتی دلیل این امر محدودیت مطالعه و مقایسه بین این روش‌ها می‌باشد، ضمن اینکه بیشتر این مطالعات و مقایسه‌ها بین روش‌های آدابوست و بگینگ صورت گرفته است. واضح است که مطالعات بیشتر در آینده ابهامات موجود در این زمینه را از بین می‌برد. برای مثال، با ارزیابی شرایطی که در آن کدام روش مناسب‌تر است درک بهتری نسبت به ترکیبات آنسامبلی ممکن ایجاد می‌شود [۱۷، ۱۸ و ۴۰].

---

1. Adaboost

2. Breiman

### ۳-۲-۲- ترکیب مدوله‌ای

قابلیت یک شبکه یکپارچه وقتی به اجزاء کوچکتری تقسیم می شود افزایش می یابد. مسأله در این روش به تعدادی زیر مسأله تقسیم می شود که هر کدام توسط یک شبکه عصبی یا ترکیبی از آن ها حل می شود. در بسیاری موارد این روش به بهبود حل مسأله می انجامد. اگر یک شبکه بزرگ جایگزین چند شبکه جزء شود پارامترهای شبکه به اندازه قابل توجهی افزایش می یابند و به تبع آن احتمال بیش برارزش افزایش می یابد. همچنین زمان آموزش چنین شبکه ای از زمان آموزش این زیر شبکه ها که به طور موازی عمل می کنند بیشتر خواهد بود. در این حالت هر قسمت از مسأله می تواند با ساختار متفاوتی از شبکه عصبی حل شود به گونه ای که قابلیت کار بالا رود. پیش پردازشی که روی داده ها صورت می گیرد را می توان نوعی تجزیه‌ی مدوله ای در جهت ساده تر کردن مسأله قلمداد کرد. در هر مؤلفه از سیستم مدوله ای به جز شبکه عصبی از الگوریتم های دیگر نیز می توان استفاده کرد و هیچ لزومی ندارد همه مؤلفه ها از الگوریتم یکسانی استفاده کنند. دلیل دیگر که منجر به استفاده از روش مدوله ای می شود این است که پیچیدگی های مدل مورد مطالعه کاهش می یابد. در نتیجه فهم، بهبود و گسترش مسأله راحت تر می شود.

مهمترین فاکتور سیستم مدوله ای چگونگی تشکیل مؤلفه ها و به عبارتی تفکیک مسأله می باشد. به عنوان مثال یکی از روش های تفکیک روش کلاسی<sup>۱</sup> می باشد که توسط آناند و همکارانش (۱۹۹۵) پیشنهاد شد. در این روش مسأله  $k$  کلاسه به  $k$  مسأله دو کلاسه تقسیم می شود به گونه ای که برای هر طبقه بندی دو کلاسه از تعداد داده های یکسان با مسأله اصلی استفاده می شود [۱۸].

### ۳-۳- روش های ترکیب

از منظر شبکه های عصبی ترکیب خروجی های یک سری شبکه آموزش دیده مانند ایجاد یک شبکه ی بزرگ است که این شبکه های آموزش دیده اجزای آن هستند که به صورت موازی عمل می کنند به

---

1. Classic decomposition

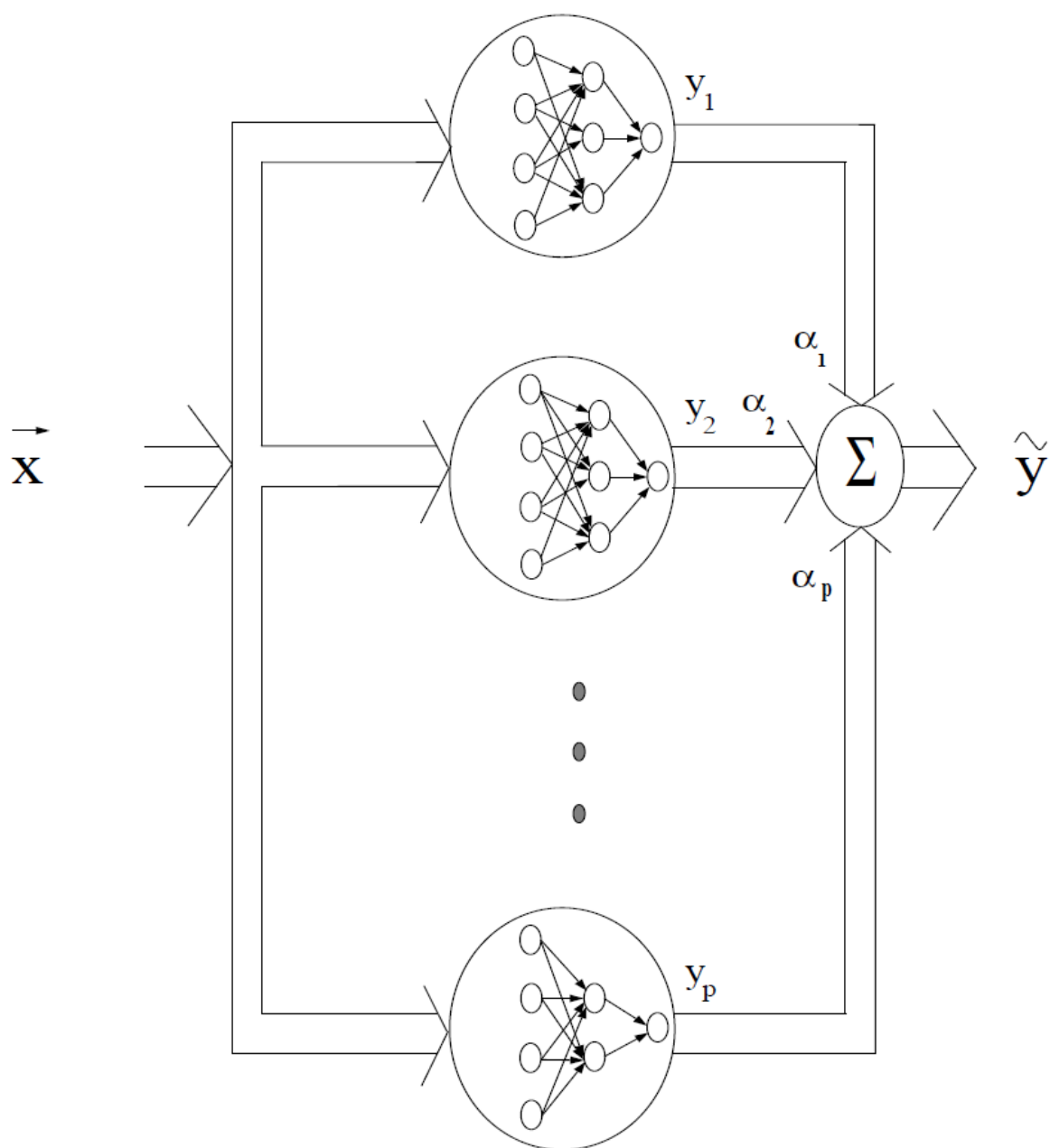
گونه‌ای که ضرایب ترکیب، ضرایب ارتباطی لایه خروجی می‌باشند. در شکل (۱-۳) نمایی از ترکیب خطی شبکه‌ها نشان داده شده است.

برای بردار ورودی  $\vec{x}$ ، خروجی نهایی  $\vec{y}$ ، جمع وزنی خروجی شبکه‌های جزء  $y_j$  می‌باشند که ضرایب ترکیب در اینجا با  $\alpha_j$  نشان داده شده‌اند. تفاوت اصلی یک شبکه عصبی بزرگ معادل و ترکیب شبکه‌های جزء در این است که در مورد ترکیب این شبکه‌ها، ضرایب ارتباطی شبکه‌های آموزش دیده ثابت می‌باشند. این در حالیست که وقتی با یک شبکه‌ی بزرگ روبرو هستیم تعداد پارامترهایی که باید تعیین شوند بیشتر می‌شود، بنابراین زمان آموزش شبکه افزایش و احتمال بیش برآزش زیاد می‌شود. ما در اینجا به ازای چند ورودی یک خروجی داریم. در مسائلی که با چند خروجی سر و کار داریم بردار وزنی را به ازای هر خروجی به طور جداگانه محاسبه می‌کنیم. برای یک شبکه به ازای ورودی  $\vec{x}$  خروجی  $y(\vec{x})$  را داریم و خطای تخمین به صورت زیر محاسبه می‌شود

[۲۶]:

$$\delta(\vec{x}) = r(\vec{x}) - y(\vec{x}) \quad (۱-۳)$$

$r(\vec{x})$  در این رابطه بردار جواب مطلوب می‌باشد.



شکل ۳-۱: ترکیب خطی شبکه‌های آموزش دیده [۲۶].

در ترکیب  $p$  شبکه عصبی، خروجی به صورت زیر در خواهد آمد [۲۶]:

$$\tilde{y}(\vec{x}; \vec{\alpha}) = \sum_{j=1}^p \alpha_j y_j(\vec{x}) \quad (۲-۳)$$

که خطای آن [۲۶]:

$$\delta(\bar{x}) = r(\bar{x}; \bar{\alpha}) - \tilde{y}(\bar{x}; \bar{\alpha}) \quad (3-3)$$

$\gamma_j$  و  $\alpha_j$  به ترتیب خروجی و ضرایب مربوط به شبکه  $j$  ام می باشد.

در اینجا باید ضرایب  $\alpha_j$  مناسب را به دست آوریم. راه‌های مختلفی برای این کار وجود دارد:

یک روش این است که یک شبکه را به عنوان بهترین شبکه  $NN_b$  انتخاب و ضریب  $\alpha_b = 1$  را به آن نسبت دهیم. به این ترتیب به باقی شبکه‌ها ضریب صفر تعلق می‌گیرد. تنها مزیت این روش، سادگی آن و ایراد اصلی، از دست دادن اطلاعات در  $(p-1)$  شبکه دیگر است.

روش دوم، میانگین‌گیری ساده<sup>۱</sup> است که به طور گسترده مورد استفاده قرار می‌گیرد. در اینجا فرض می‌شود شبکه‌ها به یک اندازه خوب هستند و توانایی یکسانی دارند.

روش سوم، یافتن وزن‌های بهینه<sup>۲</sup> می‌باشد به طوری که تابع عملکرد را که معمولاً میانگین مربعات خطا<sup>۳</sup>، می‌باشد کمینه کند. در این زمینه هاشم و همکاران روشی تحلیلی  $MSE - OLC$  را برای یافتن این ضرایب پیشنهاد کردند که با یک سری عملیات ساده ماتریسی حاصل می‌شوند. همچنین می‌توان از خروجی‌های شبکه‌های مختلف به عنوان ورودی یک شبکه عصبی دیگر استفاده کنیم که در این صورت ترکیب غیر خطی است. شبکه عصبی ترکیب‌کننده آموزش می‌بیند، تا شبکه‌ها را با وزن‌هایی که در فضای ورودی تغییر می‌کنند، ترکیب کند [۱۲ و ۲۶].

از آنجا که می‌خواهیم ضرایب ترکیب را با روش الگوریتم ژنتیک به دست آوریم از ذکر جزئیات در مورد روش‌های قبل خودداری می‌کنیم.

---

1. Simple averaging  
3. Mean Square Error (MSE)

2. Optimal Linear Combination (OLC)

## ۳-۴- الگوریتم ژنتیک

### ۳-۴-۱- مقدمه

در چهار دهه‌ی گذشته ، الگوریتم‌های زیادی برای حل مسائل بهینه‌سازی ارائه شده است. اغلب این الگوریتم‌ها بر اساس روش‌های برنامه ریزی خطی و غیر خطی که نیازمند اطلاعات جامع و ریاضیات قوی هستند می‌باشند. در گذر فرآیندهای بهینه‌سازی، افراد متعددی برای گریز از روابط سنگین ریاضی و به دلایل مختلف دیگر ، به روش‌های نوین و فرا ذهنی بهینه‌سازی روی آوردند. امروزه روش‌های نوین بهینه‌سازی که بر پایه الهام از فرآیندهای طبیعی بنا شده است، در حل بسیاری از مسائل مختلف مورد استفاده قرار می‌گیرند که عمده‌ترین آن‌ها عبارتند از:

1. Simulated Annealing
2. Ant Colony
3. Random Cost
4. Evolution Strategy
5. Genetic Algorithm
6. Cellular Automata
7. Harmony Search
8. Tabu Search
9. Swarm Optimization

روش Simulated Annealing که تقلیدی از پدیده سرد شدن فلزات ، برای ساخت یک روال جستجو است و یا روش Ant Colony که با الهام از زندگی دسته جمعی حشرات به خصوص مورچه پی‌ریزی شده است و در حل مشکلات ترافیک شبکه‌ها و مسیریابی در سیستم‌های مخابراتی شلوغ استفاده می‌-



شود. این روش اولین بار توسط دنیوبورگه<sup>۱</sup> مطرح و بعدها توسط دوریگو<sup>۲</sup> در سال ۱۹۹۹ توسعه یافت [۱۱]. توضیح بیشتر در پیوست ب ارائه شده است.

در این فصل، از میان روش‌های نوین بهینه‌سازی روش الگوریتم ژنتیک را برای بررسی بیشتر انتخاب نموده‌ایم؛ که به طور اجمالی در ذیل به شرح آن پرداخته‌ایم.

### ۳-۴-۲- تاریخچه الگوریتم‌های ژنتیک

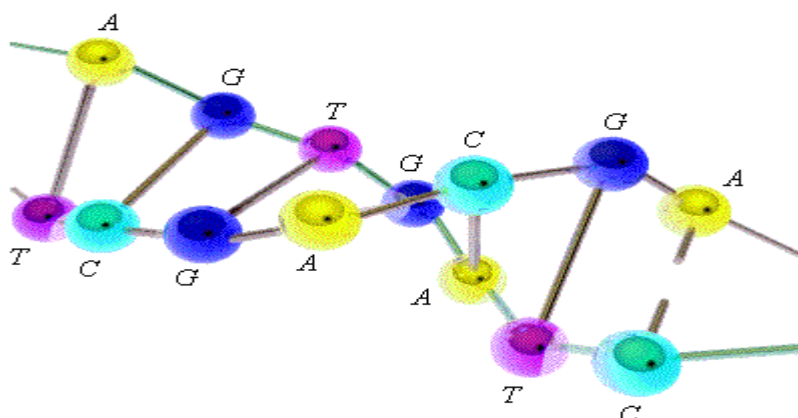
ایده اصلی الگوریتم‌های تکاملی<sup>۳</sup> در سال ۱۹۶۰ توسط ریچنبرگ<sup>۴</sup> مطرح گردید [۱۱]. الگوریتم‌های ژنتیک که منشعب از این نوع الگوریتم‌ها است، در حقیقت روش جستجوی کامپیوتری بر پایه الگوریتم‌های بهینه‌سازی و بر اساس ساختار ژن<sup>۵</sup> ها و کروموزوم<sup>۶</sup> ها است که توسط پروفیسور هولند<sup>۷</sup> در دانشگاه میشیگان مطرح شد و پس از وی توسط جمعی از دانشجویانش مثل گلدبرگ<sup>۸</sup> و آن آربر<sup>۹</sup> توسعه یافت. تا به امروز کتاب‌های متعددی توسط افرادی چون گلدبرگ، دب<sup>۱۰</sup> و جنکینز<sup>۱۱</sup> به رشته تحریر در آمده است. و پس از آن بسیاری از محققین مقالات متنوعی درباره بسط دامنه کاربرد GA، بهبود نتایج و افزایش سرعت محاسبات، ارائه نمودند. این الگوریتم امروزه در بسیاری از علوم مختلف مثل زیست‌شناسی، علوم فنی و مهندسی ( شبکه‌های عصبی، پردازش تصویر و تشخیص الگو و ... ) علوم پایه، علوم اجتماعی و ... کاربرد دارد و به دلیل کاربرد فراوان، از جایگاه ویژه‌ای در میان روش‌های نوین بهینه‌سازی برخوردار است [۱۰ و ۱۱].

---

1. Deneubourge    2. Dorigo    3. Evolutionary    4. Rechenberg    5. Gene    6. Chromosome  
7. Holland    8. Goldberg    9. Ann Arbor    10. Deb    11. Jenkins

### ۳-۴-۳- پیش زمینه بیولوژیکی ژن‌ها و کروموزوم‌ها

همه‌ی ارگانیسم‌های زنده از سلول تشکیل شده‌اند. در هر سلول مجموعه‌ای از کروموزوم‌ها به شکل رشته‌ای از DNA تشکیل شده که در شکل (۲-۳) تعدادی از ژن‌های سازنده‌ی کروموزوم نشان داده شده است.



شکل ۲-۳: یک رشته DNA [۱۱].

هر ژن یک الگوی خاصی را رمز گشایی<sup>۱</sup> می‌کند. به عبارت دیگر هر ژن یک صفت<sup>۲</sup> را رمز گشایی می‌نماید. مثلاً رنگ چشم یک فرد به عنوان یک صفت است. مجموعه‌ای از این صفت‌ها را آلل<sup>۳</sup> می‌گوییم. از سوی دیگر هر ژن دارای موقعیت مشخصی در کروموزوم است که به این موقعیت لوکوس<sup>۴</sup> می‌گویند. مجموعه کامل ماده ژنتیکی را ژنوم<sup>۵</sup> می‌گویند و یک مجموعه به خصوصی از ژن‌ها را در ژنوم، ژنوتیپ<sup>۶</sup> می‌نامند که این ژنوتیپ اساس فنوتیپ<sup>۷</sup> بوده و ویژگی فیزیکی و فکری مثل رنگ چشم و هوش و ... را بوجود می‌آورد.

در هنگام تولید سلول‌های جدید یک تلفیق<sup>۸</sup> توسط عمل ادغام<sup>۹</sup> صورت می‌گیرد. در این فرآیند ژن‌های والد، کروموزوم‌های جدید را تشکیل می‌دهند. احتمال دارد، این مولودهای جدید جهش<sup>۱۰</sup> یابند.

1. Decode  
5. Genome  
9. CrossOver

2. Trait  
6. Genotype  
10. Mutation

3. Alleles  
7. Phenotype

4. Locus  
8. Recombination

یعنی DNA آن‌ها دستخوش تحول و تغییرات شود. معیار مناسب بودن<sup>۱</sup> یک ارگانیسم با توجه به موفقیت این ارگانیسم در ادامه حیات آن تعیین می‌شود [۱۱].

### ۳-۴-۴- بهینه‌سازی بالگوریتم‌های ژنتیک

الگوریتم ژنتیک یا به اختصار GA، یک روش جستجوی مؤثر در فضاهای بسیار وسیع و بزرگ است که در نهایت منجر به جهت‌گیری به سمت یافتن یک جواب بهینه می‌گردد. این الگوریتم با الهام از تئوری داروین<sup>۲</sup> درباره حیات بهترین‌ها شکل گرفته است که بر اساس آن بهترین‌ها حق بقا دارند، بنابراین می‌توانیم بگوییم [۱۱]:

« GA بر اساس اصل " ادامه حیات بهترین‌ها " و " تکثیر نوع برتر " پی‌ریزی شده است.»

این روش هوشمند بطور موفقیت آمیزی طرح بهینه‌ی عمومی<sup>۳</sup> را بدون در نظر گرفتن فرضیات محدود کننده‌ای از قبیل پیوسته بودن فضای جستجو و یا وجود مشتقات پیدا می‌کند [۲۷]. در این روش فضای طراحی<sup>۴</sup> به فضای ژنتیک<sup>۵</sup> تبدیل شده و عملیات بر روی متغیرهای طراحی رمز شده که رشته‌ای از کاراکترها با طول محدود هستند، انجام می‌گیرد. مزیت کار با متغیرهای کد شده در این است که اصولاً کدها قابلیت تبدیل فضای پیوسته به فضای گسسته را دارند.

نکته جالب در GA آن است که اساساً GA بر پردازش تصادفی<sup>۶</sup> یا به تعبیر صحیح‌تر پردازش تصادفی هدایت شده<sup>۷</sup> استوار است؛ بنابراین عملگرهای تصادفی فضای جستجو را به صورت تطبیقی<sup>۸</sup> مورد بررسی قرار می‌دهند [۱۱].

---

1. Fitness  
4. Design Space  
7. Guided Random

2. Darwin  
5. Genetic Space  
8. Adaptive

3. Global Optimum  
6. Random

الگوریتم GA در ابتدا با مجموعه ای از جواب های تصادفی (کروموزوم ها) که به آن جمعیت<sup>۱</sup> گفته می شود، آغاز می گردد. از این جواب ها برای ساخت جمعیت بعدی استفاده می شود، به این امید که جمعیت های جدید بهتر از جمعیت های قدیم باشند. زیرا روش هایی که برای انتخاب جمعیت جدید استفاده شده با توجه به مناسب بودن آن ها صورت گرفته است. پس بهترین ها شانس بیشتری برای تولید مثل خواهند داشت. این فرآیند آنقدر تکرار می شود تا شرایط خاتمه (برای دستیابی به بهترین راه حل) محقق شود [۳۴].

در بهینه یابی از طریق GA هفت مرحله ی اساسی انجام می پذیرد [۱۰]:

۱. تولید تصادفی مجموعه ای از طرح ها، که بدان جمعیت اولیه گفته می شود. هر عضو این جمعیت یک کروموزوم می باشد که بصورت کد<sup>۲</sup> بوده و رشته<sup>۳</sup> نامیده می شود.
۲. ارزیابی و مرتب سازی جواب ها و حذف اعضای اضافی جمعیت (در صورت وجود).
۳. انتخاب عده ای از بهترین اعضای جمعیت به عنوان والد ها بر حسب شایستگی، انجام عمل تقاطع میان آن ها و ایجاد جمعیت جدیدی از فرزندان.
۴. انتخاب عده ای از اعضای جمعیت به صورت تصادفی، انجام عمل جهش بر روی آن ها و ایجاد جمعیتی از جهش یافتگان.
۵. ادغام جمعیت اصلی، جمعیت فرزندان و جمعیت جهش یافتگان.
۶. در صورتی که شرایط خاتمه محقق شده اند، رفتن به مرحله ی ۷، در غیر این صورت بازگشت به مرحله ی ۲.
۷. پایان.

برای بررسی بیشتر اصطلاحات و نکات GA، بشرح ذیل به طور اجمالی هریک را مورد پردازش قرار می دهیم.

---

1. Population  
3. String

2. Code

### ۳-۴-۵- تولید مثل<sup>۱</sup>

تولید مثل معمولاً اولین عملی است که بر روی جمعیت اعمال می‌شود. در این روش یک سری کروموزوم از میان جمعیت به عنوان والد انتخاب شده که در نهایت با عمل ادغام منجر به تولید فرزندان می‌شوند.

براساس نظریه "حیات بهترین‌ها" باید بهترین موارد انتخاب شوند تا نسل بعدی بهتری را تولید کنند. به همین دلیل گاه به عملگر تولید مثل، عملگر انتخاب<sup>۲</sup> نیز گفته می‌شود؛ روش‌های انتخابی مختلفی در GA برای انتخاب کروموزوم‌ها وجود دارد، اما هدف اصلی در همه‌ی آن‌ها انتخاب رشته‌هایی با میانگین شایستگی بالا از جمعیت فعلی و قرار دادن آن‌ها در یک مکان به نام استخر تولید مثل<sup>۳</sup> است. پس کروموزوم‌های با مقدار شایستگی بالاتر، از شانس بیشتری برای حضور در استخر تولید مثل برخوردار خواهند بود. عکس این مطلب نیز صادق است، یعنی کروموزوم‌هایی با مقدار شایستگی کمتر، احتمال کمتری برای کپی شدن در استخر تولید مثل خواهند داشت.

### ۳-۴-۶- انواع روش‌های انتخاب

روش‌های مختلفی برای انتخاب کروموزوم‌ها و ادغام آن‌ها وجود دارد که مهم‌ترین این روش‌ها عبارتند از [۲۷]:

۱. روش چرخ گردان<sup>۴</sup>

۲. روش مسابقه<sup>۵</sup>

۳. روش رتبه بندی<sup>۶</sup>

<sup>۱</sup> . Reproduction  
4. Roulette Wheel

2. Selection Operator  
5. Tournament

3. Mating Pool  
6. Rank

۴. روش حالت پایدار<sup>۱</sup>

۵. مفهوم برگزیده

در اینجا به توضیح مختصر روش چرخ گردان می‌پردازیم.

### ۳-۴-۶-۱- روش چرخ گردان

یکی از روش‌های متداول در انتخاب یک رشته، استفاده از عملگری است که مبتنی بر یک احتمالی از تابع شایستگی باشد. بنابراین رشته  $i$  ام در جمعیت با یک احتمالی از  $F_i$  انتخاب می‌شود که باید حاصل جمع این احتمال‌ها نیز برابر یک شود. احتمال رشته انتخابی  $i$  ام برابر است با [۱۱]:

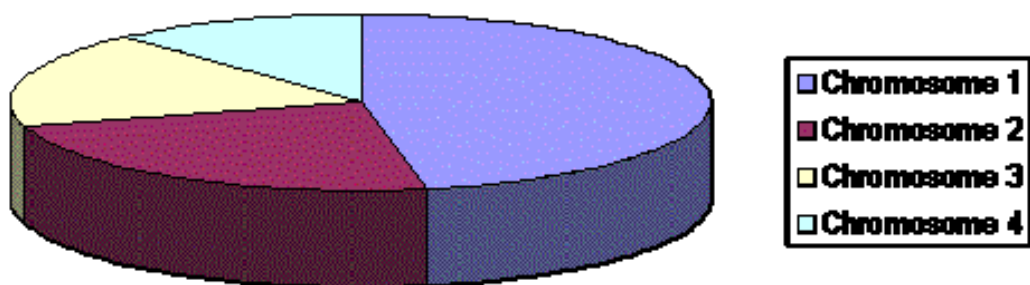
$$P_i = \frac{F_i}{\sum_{j=1}^n F_j} \quad (۳-۴)$$

که  $n$  تعداد جمعیت و  $F$  مقدار شایستگی هر کروموزوم می‌باشد.

حال اگر درصد احتمال این مقادیر را به صورت یک چرخ رولت نشان دهیم. آن گاه مقدار مناسب بودن جمعیت پس از  $n$  بار چرخیدن بدست خواهد آمد. در هر زمان یک مورد از رشته‌ها توسط اشاره گر چرخ رولت انتخاب می‌شود. در این شکل کروموزوم‌های ۱ و ۲ دارای برزندگی بیشتری نسبت به کروموزوم‌های ۳ و ۴ هستند و بنابراین در مرحله‌ی انتخاب شانس بیشتری دارند. این شیوه-ی انتخاب سبب می‌شود که با گذشت زمان، تعداد کروموزوم‌های مطلوب در جمعیت افزایش یابد به-طوری که میانگین مقدار برزندگی جمعیت، در مقایسه با جمعیت مرحله‌ی قبل، بیشتر می‌شود [۱۱].

---

1. Steady State



شکل ۳-۳: مدلی از چرخ گردان [۳۱]

### ۳-۴-۷- عملگرهای GA

عملگرهای وراثتی متعددی برای تولید رشته های بهتر وجود دارد که هدف این عملگرها جستجوی فضای پارامترها و تا حد امکان حفظ اطلاعات نهفته در رشته ها است. چرا که این والد ها بهترین موارد انتخاب شده توسط عملگرهای فاز تولیدمثل هستند و نباید از دست بروند. از مهمترین عملگرهای GA می توان به موارد زیر اشاره کرد [۴۱]:

۱. عمل معکوس کردن<sup>۱</sup>
۲. عمل حذف کردن<sup>۲</sup>
۳. عمل جداسازی<sup>۳</sup>
۴. عمل نقل مکان<sup>۴</sup>
۵. عمل بخش بندی<sup>۵</sup>
۶. عمل جفت گیری<sup>۶</sup> یا ادغام
۷. عمل غالب شدن یا تسلط<sup>۷</sup>
۸. عمل کپی کردن<sup>۸</sup>

1. Inversion  
5. Sharing

2. Deletion  
6. Mating

3. Segregation  
7. Dominance

4. Migration  
8. Duplication

که معمولاً در یک الگوریتم ژنتیک ساده تنها از سه عملگر اصلی زیر استفاده می‌شود:

۱. تولید مثل

۲. ادغام

۳. جهش

در قسمت قبل فرآیند تولید مثل به تفصیل مورد بحث قرار گرفت. در این بخش به بررسی ادغام و چند عملگر دیگر می‌پردازیم.

### ۳-۴-۷-۱- عملگر ادغام

پس از اینکه مرحله تولید مثل تمام شد، جمعیتی از بهترین‌ها بوجود آمده است. در حقیقت عمل تولید مثل یک مجموعه‌ای از بهترین رشته‌ها را انتخاب کرده، اما رشته جدیدی را به وجود نیاورده است. به همین دلیل عملگر ادغام با هدف تولید رشته‌های بهتر بر روی استخر تولید مثل اعمال می‌شود. هدف از ادغام جستجوی فضای پارامتر و تا حد امکان حفظ اطلاعات نهفته در رشته‌ها است.

عملگر ادغام یک عملگر ترکیبی است که شامل سه عمل است. اول عملگر تولید مثل یک جفت رشته را به صورت تصادفی انتخاب می‌کند. دوم یک محلی را برای عمل ادغام به طور تصادفی در طول رشته انتخاب کرده و سرانجام در سومین مرحله مقدار رشته را با توجه به محل ادغام (که مشخص کردیم) جابه جا می‌کند.

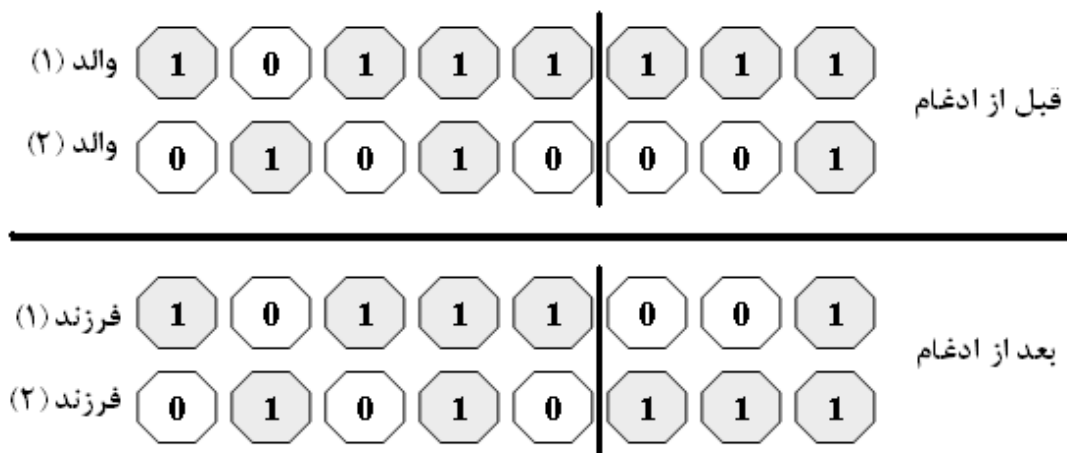
این روش را اصطلاحاً روش ادغام تک مکانی می‌نامند. بر اساس نظر دب (۱۹۹۵) نمی‌توان گفت که کدامیک از روش‌های ادغام بهتر است. بنابراین انتخاب یک روش ادغام مناسب با توجه به سلیقه‌ی افراد و شرایط مسئله به دلخواه صورت می‌گیرد.

روش‌های مختلف دیگری برای عمل ادغام وجود دارد که به شرح هر یک از آن‌ها می‌پردازیم [۱۱]:



### ۳-۴-۷-۱-۱- روش ادغام تک نقطه‌ای یا مکانی<sup>۱</sup>

همان‌طور که در قسمت قبل بیان نمودیم، در این روش یک مکان تصادفی در طول رشته انتخاب می‌شود و بیت‌های پس از این مکان جابجا می‌شوند. شکل ۳-۴ نمونه‌ای از روش ادغام تک نقطه‌ای می‌باشد.



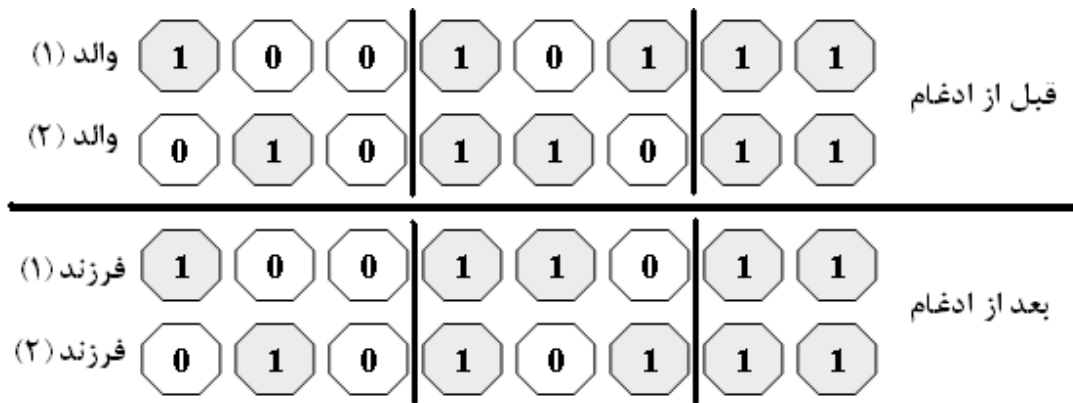
شکل ۳-۴: روش ادغام تک نقطه‌ای [۱۱].

از آنجایی که مکان جابجایی به صورت تصادفی انتخاب شده است لذا از مناسب بودن این مکان اطلاعی نداریم. اگر این نقطه مکان مناسبی باشد منجر به تولید فرزندان مناسبی می‌شود در غیر این صورت باعث از بین رفتن کیفیت رشته و حذف شدن رشته در نسل بعدی می‌گردد.

1. Single Point Crossover

### ۳-۴-۷-۱-۲- روش ادغام دو نقطه‌ای<sup>۱</sup>

عملگر ادغام دو نقطه‌ای، دو مکان را به صورت تصادفی در طول رشته انتخاب کرده و مقادیر بین این دو نقطه را جابجا می‌کند. در شکل (۳-۵) روش ادغام دو نقطه‌ای نشان داده شده است [۲۷].



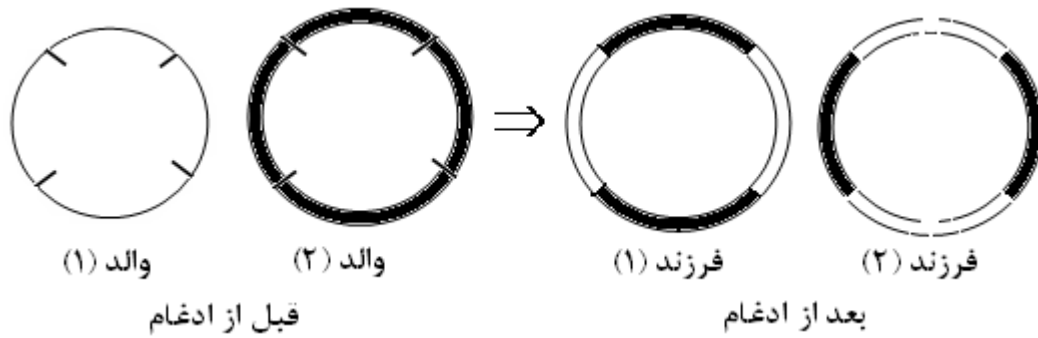
شکل ۳-۵: روش ادغام دو نقطه‌ای [۱۱].

### ۳-۴-۷-۱-۳- روش ادغام چند نقطه‌ای<sup>۲</sup>

در روش چند نقطه‌ای ممکن است تعداد مکان‌ها زوج و یا فرد شود. در حالتی که تعداد زوج باشد، رشته‌ها به صورت یک حلقه‌ای که بدون ابتدا و انتها است، خواهند بود. این مکان‌ها به صورت تصادفی در اطراف دایره انتخاب شده‌اند؛ بنابراین اطلاعات بین این مکان‌ها به صورت شکل (۳-۶) جابجا می‌شود.

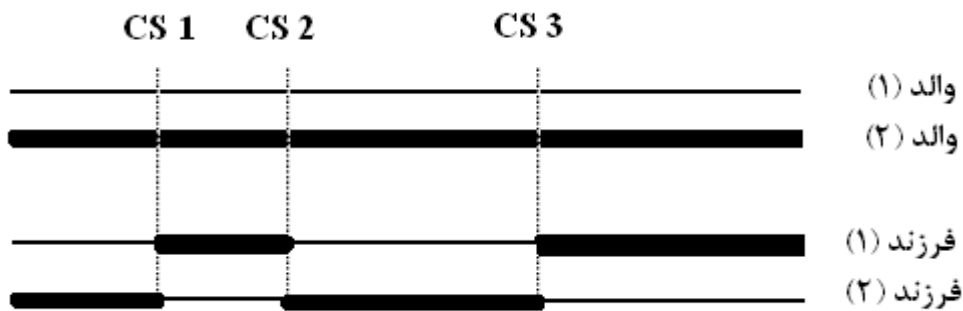
1. Two-Point Crossover

2. Multi-Point Crossover



شکل ۳-۶: روش ادغام چند نقطه‌ای (حالتی که تعداد مکان‌ها زوج باشد) [۱۱].

اما اگر تعداد آن‌ها فرد باشد یک مکان متفاوتی در رشته در نظر گرفته شده و اطلاعات آن‌ها به صورت شکل (۳-۷) جایجا می‌شود [۳۷ و ۴۱].



شکل ۳-۷: روش ادغام چند نقطه‌ای (حالتی که تعداد مکان‌ها فرد باشد) [۱۱].

### ۳-۴-۷-۲- جهش:

از دیگر مراحل رایج در عملیات GA که تکامل جمعیت برای نسل بعد را به دنبال خواهد داشت، استفاده از عملگر جهش است. پس از عمل ادغام رشته‌ها، نوبت به عمل جهش یا موتاسیون می‌رسد. هدف از بکارگیری اپراتور جهش، جستجوی بهتر فضای طراحی است.

### ۳-۴-۷-۲-۱- عمل جهش یا موتاسیون

عمل جهش یک بیت، شامل تبدیل عدد صفر به یک و بالعکس است که بر اساس یک احتمال کوچک مثل  $P_m$  به صورت بیت به بیت صورت می‌گیرد. عمل جهش به این ترتیب است که یک عدد تصادفی بین صفر تا یک تولید می‌شود؛ اگر عدد تولید شده کوچکتر از  $P_m$  باشد مقدار خروجی را برابر درست<sup>۱</sup> و در غیر این صورت برابر غلط<sup>۲</sup> در نظر می‌گیریم. اگر برای هر بیت مقدار خروجی درست باشد، بیت تغییر می‌کند و گرنه بیت بدون تغییر باقی خواهد ماند.

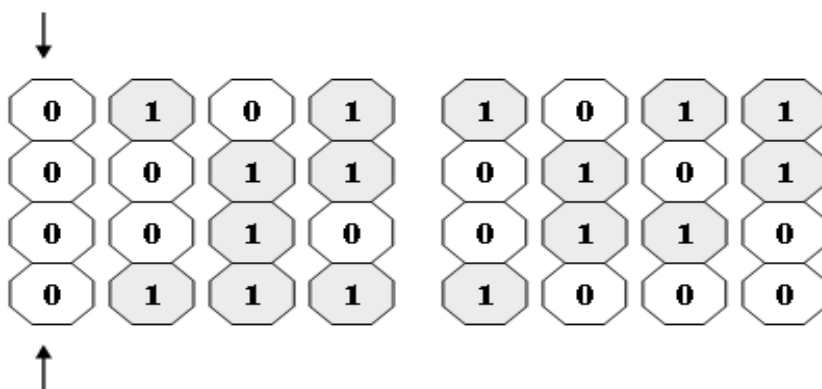
بیت‌های یک رشته بصورت مستقل جهش می‌یابند، به این معنا که جهش یک بیت بر روی احتمال سایر بیت‌ها تأثیر نمی‌گذارد. این عمل در یک الگوریتم ژنتیک ساده به منزله‌ی یک عملگر ثانویه و به منظور حفظ اطلاعاتی که در حال از دست رفتن است، تلقی می‌گردد. برای مثال، فرض کنید مقادیر بیت‌های رشته‌های یک جمعیت در یک محدوده‌ی خاص برابر صفر شده است و حل بهینه نیاز به عدد یک در آن نقطه دارد. این در حالی است که عملگر ادغام نمی‌تواند عدد یک را در آن موقعیت تولید نماید، بنابراین برای تولید عدد یک از عمل جهش یا موتاسیون استفاده می‌کنیم.

این عمل برای جلوگیری از همگرایی سریع و کمک به الگوریتم جستجو برای فرار از به دام افتادن در مینیمم‌های موضعی<sup>۳</sup> مفید است. از سوی دیگر این عمل برای حفظ حالت متفاوت و متمایز بودن کروموزوم‌ها در یک جمعیت به کار می‌رود. به عنوان مثال جمعیت زیر را در نظر می‌گیریم:

---

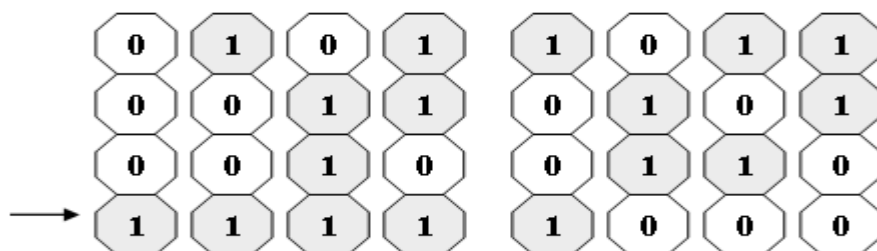
1. True  
3. Local Optimum

2. False



شکل ۳-۸: نیاز یک جمعیت به بیت ۱ [۱۱].

همان طور که ملاحظه می‌شود، همهی بیت‌های سمت چپ این رشته‌ها صفر است. اگر راه حل بهینه نیاز به یک در این موقعیت داشته باشد، در این صورت هیچکدام از عملیات‌های ادغام و تولید مثل قادر به این کار نخواهد بود اما عمل جهش این کار را برای ما انجام خواهد داد و جمعیت پس از این عمل با توجه به یک احتمال مشخص به صورت زیر در می‌آید [۱۱]:



شکل ۳-۹: تاثیر عملگر جهش [۱۱].

پس به طور خلاصه می‌توان گفت:

«عمل جهش باعث جنبش و حرکت در فضای جستجو و ذخیره‌ی اطلاعات از دست رفته جمعیت می‌شود.»

### ۳-۴-۸- شرط توقف الگوریتم:

سؤال مهمی که در اینجا وجود دارد، این است که چه زمانی باید الگوریتم متوقف شود؟

باید اذعان کرد که هیچ جواب جامعی و مناسبی برای این سؤال وجود ندارد. اما بعضی از شرط-

های توقف متداول عبارتند از:

۱. رسیدن به حد مطلوبی از جواب.
  ۲. سپری شدن تعداد معینی از تکرارها.
  ۳. سپری شدن تعداد معینی از تکرارها، بدون مشاهده‌ی بهبود خاصی در نتیجه.
- در پایان باید اذعان داشت که ممکن است در GA نیز، همانند دیگر روش‌های نوین بهینه‌سازی، به جواب بهینه نرسیم و یا در نقطه کمینه‌ی محلی گرفتار شویم. در هر صورت اگر الگوریتم به یک حل بهینه همگرا نشود، می‌بایست پارامترهای آن را (مثل اندازه جمعیت، درصد جهش، درصد ادغام و امثال آن را) تغییر داده و مجدداً الگوریتم را اجرا نماییم [۱۱].

## فصل چهارم

تخمین اشباع از آب سازند سروک واقع در  
یکی از میادین جنوب غربی ایران، توسط  
شبکه‌های عصبی کمیته ای و الگوریتم ژنتیک

#### ۴-۱- مقدمه

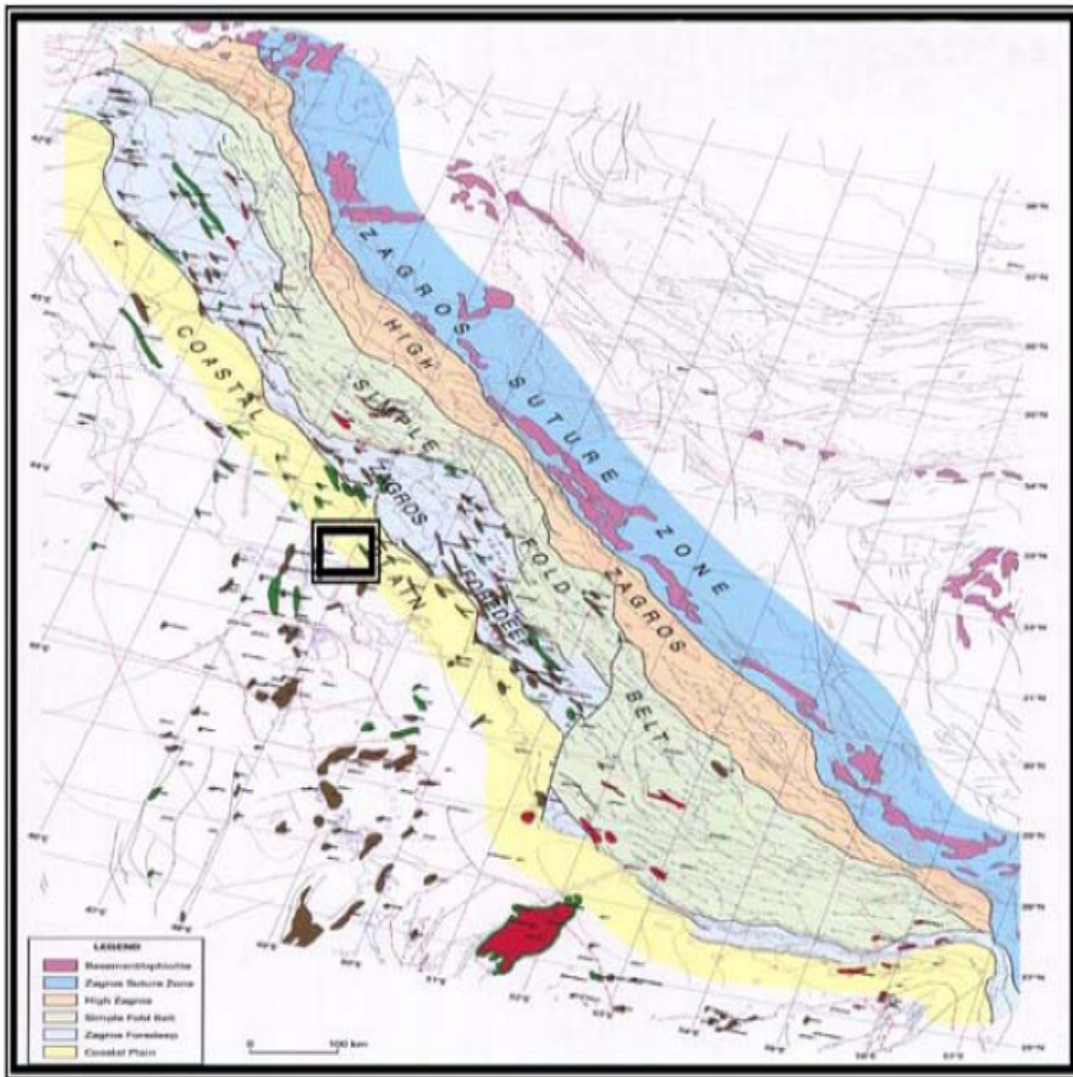
در این فصل به منظور بررسی قابلیت و توانایی شبکه‌های عصبی مصنوعی کمیته‌ای در بهبود نتایج حاصل از شبکه‌های عصبی منفرد، مطالعه‌ی موردی در یکی از میادین جنوب غربی انجام شد. در این مهم از داده‌های مربوط به محدوده‌ی عمقی سازند سروک استفاده شد و اشباع از آب ابتدا با استفاده از شبکه‌های عصبی مصنوعی منفرد و سپس با استفاده از ترکیب آنسامبلی نتایج آن‌ها تخمین زده شد.

#### ۴-۲- زمین‌شناسی میدان مورد مطالعه

در زمین‌های پست در امتداد مرز ایران و عراق واقع در صفحه‌ی آبادان و غرب دزفول، رخنمون سنگی وجود ندارد و اطلاعات زمین‌شناسی از کاوش‌های زیر سطحی حاصل شده است.

مخازن اصلی جنوب غربی ایران در حوزه زاگرس، در سازند آسماری (گروه فارس)، و سازندهای گروه بنگستان وجود دارند. حوزه‌ی زاگرس یک حوزه‌ی جلو آمده از رشته کوه زاگرس می‌باشد. در شکل ۴-۱ موقعیت میدان مورد نظر توسط کادر مستطیلی نشان داده شده است.





شکل ۴-۱: کادر مستطیلی موقعیت میدان مورد نظر را نشان می‌دهد [۲۴].

Time Unit			Lithology Unit			
Era	Series	Stage	Group	Formation		
Tertiary	Miocene			Aghajari		
				Gachsaran		
Mesozoic	Oligocene-Miocene		Fars	Asmari		
	Eocene-Paleocene			Pabdeh		
	Late Cretaceous	Maastrichtian			Gurpi	
		Campanian			Ilam	
		Santonian			Lafan	
	Middle Cretaceous	Cenomanian			Sarvak	
		Albian			Bangestan	Kazhdumi
		Early Cretaceous		Aptian		Dariyan
				Neocomian		Gadvan
					Fahliyan	
					Garau	
					Khami	

جدول ۴-۱: چینه شناسی میدان مورد نظر [۲۴].

این میدان یک ساختار بسته در افق‌های عمیقتر چینه‌شناسی می‌باشد که در قسمت شمال شرقی دزفول واقع شده است. طول و عرض این میدان به ترتیب حدود ۲۱ و ۵ کیلومتر است. قسمت شمال ساختار مزبور از قسمت جنوبی شیب بیشتری دارد.

۶ چاه در این میدان حفر شده‌اند. این میدان از مخازنی که در گروه‌های خامی و بنگستان وجود دارند تشکیل شده است. ۳ واحد زمین‌شناسی به عنوان اهداف اصلی در ارتباط با استخراج و تولید نفت در این میدان مشخص شده‌اند: ماسه سنگ‌های سازند گادوان و سازند کژدمی و سنگ آهک سازند سروک.

طبقه‌بندی چین‌شناسی میدان مورد مطالعه در جدول (۴-۱) ارائه شده است. باید توجه داشت که از لحاظ منطقه‌ای دگر شییبی‌هایی بین سازندهای داریان و کژدمی وجود دارند که این دگر شییبی‌ها در بالا و وسط شیل لافان و بین سازندهای گورپی و پابده ایجاد شده‌اند.

#### ۴-۳- چین‌شناسی میدان مورد نظر

۶ چاه در میدان مورد نظر حفر شده‌اند. این چاه‌ها ۱۳ سازند را قطع نموده‌اند که در غالب ۳ گروه تقسیم‌بندی می‌شوند. جوان‌ترین آنها گروه فارس است که شامل سازندهای آجاجاری، گچساران، آسماری، پابده و گورپی می‌باشد. گروه میانی بنگستان نام دارد و از سازندهای ایلام، لافان، سروک و کژدمی تشکیل شده است. مسن‌ترین گروه نیز گروه خامی است و شامل سازندهای داریان، گادوان و فهلیان است. (جدول ۴-۱).

ضخامت سازندهای آجاجاری و گچساران از جنوب به شمال میدان مورد مطالعه افزایش می‌یابد. از طرف دیگر سازندهای گورپی و پابده روند معکوسی دارند و از شمال به جنوب بر ضخامتشان افزوده می‌شود. سازندهای دیگر نیز تقریباً از ضخامت یکسانی برخوردار می‌باشند.

دگرشییبی‌های<sup>۱</sup> منطقه‌ای بین سازندهای گورپی و پابده، ایلام و لافان، لافان و سروک، کژدمی و داریان به دلیل عدم انطباق بین صفحه‌ی آبادان و میدان مزبور به وجود آمده‌اند (جدول ۴-۱).

---

1. Unconformities

مهمترین واحد مخزنی در میدان مورد مطالعه سازند سروک است که از لحاظ سن زمین‌شناسی مربوط به دوره‌ی آلبین - سنومانین<sup>۱</sup> می‌باشد. این سازند معادل سازندهای مودود (آلبین - سنومانین)، احمدی (سنومانین)، رومیلا (سنومانین) و میشریف (سنومانین) در عراق می‌باشد.

#### ۴-۳-۱- سازند سروک

نام‌گذاری این سازند به دلیل نوع ساختار تنگه سروک واقع در دامنه‌ی جنوبی کوه بنگستان در استان خوزستان است. در این ساختار، سازند مزبور از سه بخش با ضخامت کلی ۸۲۲ متر تشکیل شده است (شکل ۴-۲). سازند سروک از رسوبات کربناته که در یک محیط دریایی کم عمق ته‌نشین شده‌اند تشکیل شده است. مرز پایینی سازند سروک قابل انطباق با سازند کژدمی می‌باشد و جایی تعریف می‌شود که سنگ آهک تمیز سروک روی شیل کژدمی قرار می‌گیرد.

مرز بالای سازند سروک تغییرات زیادی دارد (شکل ۴-۳). در منطقه‌ی فارس، این سازند توسط سنگ آهک روتالی (سنومانین) سازند ایلام پوشیده شده است. در استان لرستان مرز بالای سازند سورگاه ظاهر شده و مورد حفاری قرار گرفته است. سازند سروک رخساره آب‌های عمیق می‌باشد که ناپیوستگی هم شیب واقع در بالای آن به راحتی قابل تشخیص نیست. در بیشتر مناطق استان‌های فارس و خوزستان نقطه‌ی تماس سازندهای سروک و ایلام به دلیل وجود سنگ آهک در هر دو سازند قابل تشخیص نیست. در خوزستان و فارس ناپیوستگی هم شیب بین سازند ایلام و گورپی وجود دارد. در قسمت شمالی شیل‌های دریایی مربوط به عمق کم می‌باشند و کنگلومراهای مربوط به سازند لافان، سازندهای سروک و ایلام را از هم جدا می‌کند.

---

1. Albian - Cenomanian

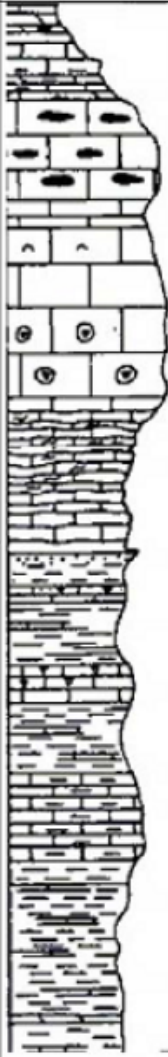
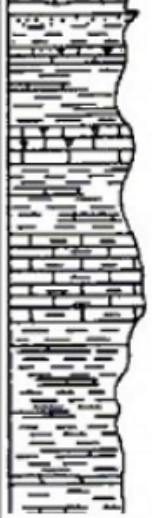
سن سازند سروک در استان‌های خوزستان، فارس و لرستان آلبین تا تورونین<sup>۱</sup> گزارش شده است. این سازند از واحدهای زیستی مختلفی تشکیل شده است و ویژگی‌های نهشته‌ای متعددی را از محیط‌های دریایی نشان می‌دهد.

در منطقه‌ی این میدان نفتی قسمت بالای سازند سروک توسط یک لایه نازک شیلی که به آن سازند شیلی لافان گفته می‌شود به راحتی قابل شناسایی است (جدول ۴-۱). مرز پایینی نیز بر سازند کژدمی منطبق<sup>۲</sup> می‌باشد. ضخامت متوسط سازند سروک در این میدان ۶۲۷ متر است [۲۴].

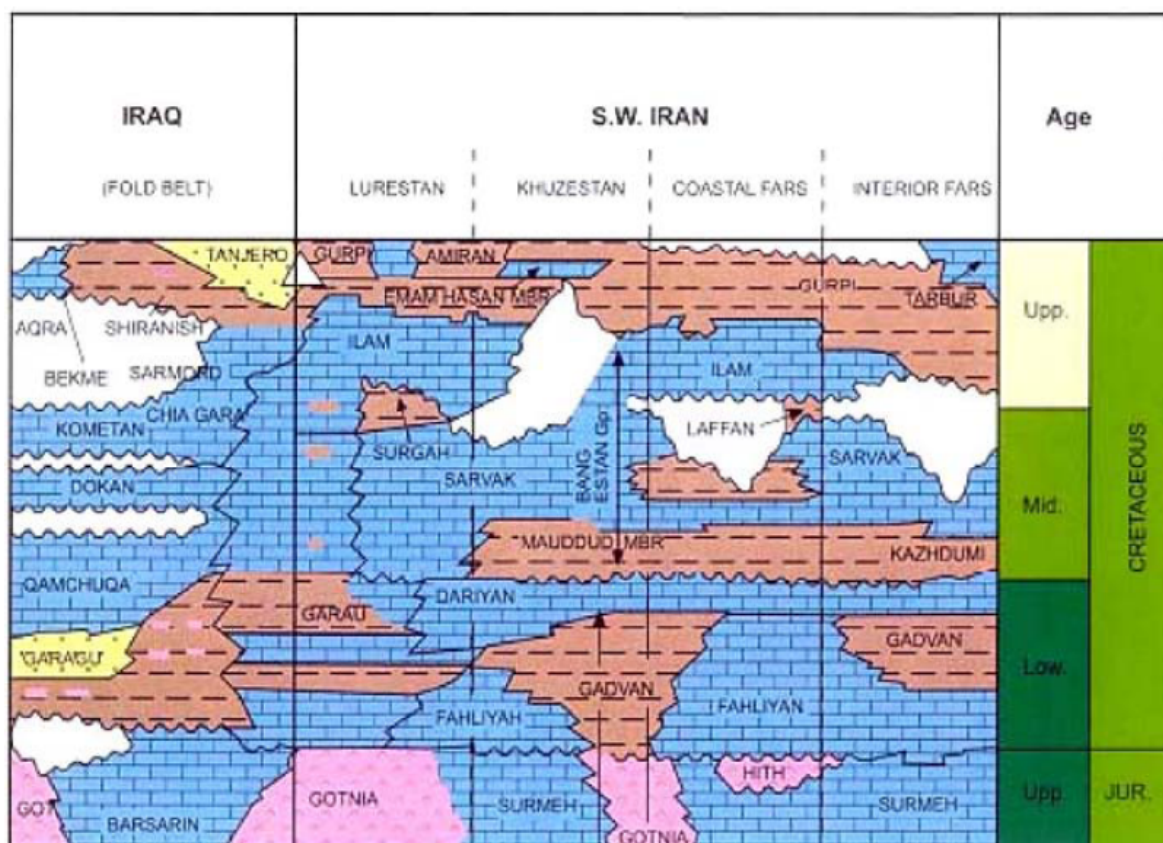
---

1. Turonian

2. Conformable

AGE	FORMATION	THICKNESS (m)	LITHOLOGY	GENERAL DESCRIPTION	TYPE SECTION	
Albian Cenomanian	SARVAK	820		Limestone with ferruginous zones of breccia	Unit 4	Tang-e Sarvak in Khuzestan Province
				Massive limestone with abundant rudist debris	Unit 3	
				Massive chalk, limestone with numerous siliceous nodules		
				Fine-grained nodular-bedded argillaceous limestone with thin marl parting	Unit 1	
Neocomian Aptian	GARAU	823		Sandy limestone (Unit 8) Shale and thin-bedded shaly limestone with minor sand and glauconite (Unit 7)  Limestone with chert nodules (Unit 5) Shale with minor limestone (Unit 4) Thinly-bedded limestone (Unit 3)  Limestone with carbonaceous shale partings (Unit 2) Carbonaceous shale and argillaceous pyritic limestone (Unit 1)	Kabir Kuh in northwestern Lorestan	

شکل ۴-۲: واحدهای پایینی و میانی چینش‌شناسی دوره‌ی کرتاسه در جنوب غربی ایران [۲۴].

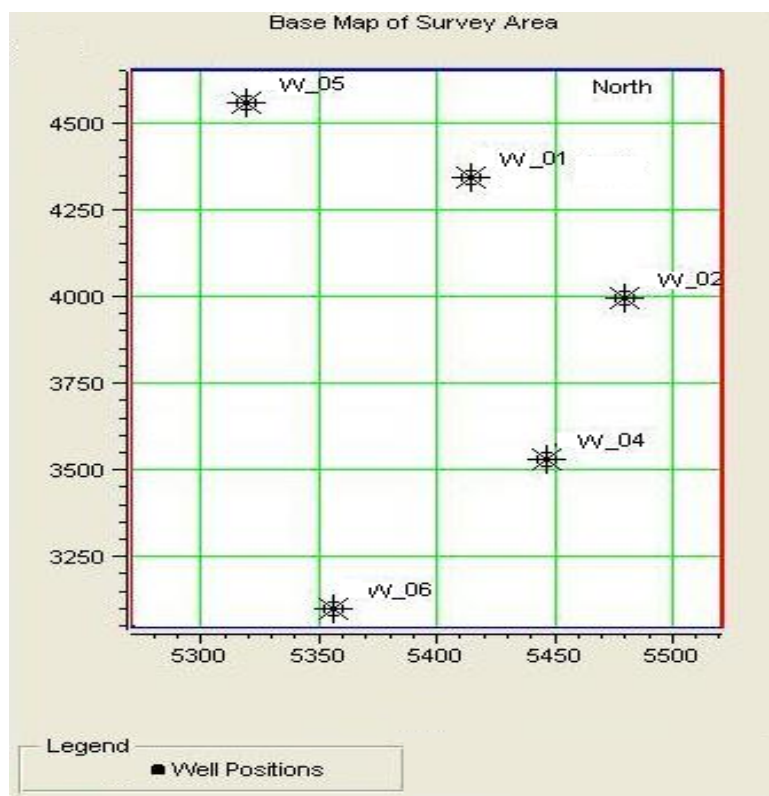


شکل ۴-۳: همبستگی سازندهای دوره‌ی کرتاسه شمال شرقی صفحه‌ی عربی از عراق تا جنوب غربی ایران [۲۴].

#### ۴-۴ - چاه‌های مورد مطالعه

شکل ۴-۴ موقعیت جغرافیایی ۵ حلقه چاه از چاه‌های میدان مورد نظر را در مختصات UTM محلی نشان می‌دهد. از داده‌های چاه‌های W\_01، W\_02، W\_04، W\_05 و W\_06 در بازه‌ی عمقی سازند سروک برای آموزش و آزمون شبکه‌ها استفاده شد. عمق شروع سازند سروک در چاه‌های مورد نظر حدوداً ۲۶۰۰ و ضخامت آن حدوداً ۶۰۰ متر است.

این داده‌ها شامل داده‌های نگارهای صوتی، چگالی، تخلخل مؤثر، پرتو گاما، مقاومت ویژه و مقاومت ویژه نرمال بلند با نرخ نمونه‌برداری حدود ۱۵ سانتیمتر است. برای هر نقطه مقدار اشباع از آب نیز با استفاده از نگارها و روابط تجربی و پس از اعمال تصحیحات محاسبه شده است.



شکل ۴-۴: موقعیت چاه‌های مورد استفاده [۲۴].

#### ۴-۵- انتخاب داده‌های آموزش و آزمون

از داده‌های چاه‌های W\_01، W\_02 و W\_04 برای آموزش شبکه‌ها و از تلفیق چاه‌های W\_05 و W\_06 برای آزمون توان تعمیم شبکه‌ها استفاده شد. مراحل طی شده برای انتخاب چاه‌ها به این شکل بود که ابتدا هر بار یکی از چاه‌ها برای آزمون انتخاب شد ولی تفاوت معناداری در رابطه با میزان خطا مشاهده نشد. از طرف دیگر برای آزمون توان تعمیم تعداد داده‌های پرت قابل توجه بود و امکان حذف این داده‌ها برای ۱ چاه وجود نداشت بنابراین از چاه‌های W\_01، W\_02 و W\_04 که بهترین نتیجه را در آموزش داشتند برای آموزش و از ترکیب چاه‌های W\_05 و W\_06 برای آزمون توان تعمیم شبکه‌ها استفاده شد و داده‌های پرت حذف شدند. نگارهای صوتی، چگالی، تخلخل مؤثر، پرتو



گاما، مقاومت ویژه و مقاومت ویژه نرمال بلند به عنوان ورودی و نگار اشباع از آب به عنوان خروجی در نظر گرفته شد. در کل از هر چاه ۷۵ ردیف داده در عمق مورد نظر در اختیار بود. داده‌های ۵ چاه در محیط اکسل قرار گرفتند. برای ۳ چاه ۱، ۲ و ۴ داده‌ها بر حسب عمق زیر یکدیگر قرار گرفتند به گونه‌ای که در مجموع ۲۲۵ ردیف داده حاصل شد. هر ردیف نیز دارای ۸ ستون عمق، نگارهای صوتی، دانسیته، تخلخل مؤثر، پرتو گاما، مقاومت ویژه، مقاومت ویژه نرمال بلند و اشباع از آب بود. به طور مشابه برای چاه‌های ۵ و ۶ نیز داده‌ها در محیط Excel قرار گرفتند و در مجموع ۹۵ ردیف داده از ۱۵۰ ردیف برای آزمون توان تعمیم انتخاب شدند. به دلیل کم بودن داده‌ها عملیات مرتب‌سازی<sup>۱</sup> برای داده‌ها صورت نگرفت و از تمام آنها برای آموزش استفاده شد. برای داده‌های آزمون نیز به دلیل وجود داده‌های نویز تعداد ۵۵ ردیف از آنها حذف شد.

در شرایط واقعی محاسبه پارامترهای پتروفیزیکی از جمله اشباع از آب با محدودیت همراه می‌باشد و مستلزم محاسبات فراوان، استفاده از داده‌های زمین‌شناسی و مغزه‌ها و انجام تصحیحات و به طور کلی صرف زمان و هزینه زیادی است. به همین دلیل برای صرفه‌جویی در صرف هزینه و زمان در تعیین مقادیر این پارامترها از روی نگارها از شبکه‌های عصبی استفاده می‌کنیم.

#### ۴-۶- برنامه نویسی رایانه‌ای

در تمام مراحل این تحقیق از نرم افزار MATLAB استفاده شد و تمام برنامه‌ها در محیط این نرم افزار نوشته شد. طراحی شبکه‌های پس انتشار خطا به ۳ روش فرا آموزش<sup>۲</sup>، مرتب‌سازی<sup>۳</sup> و توقف سریع<sup>۴</sup> انجام شد. در هر ۳ مورد از تابع Prestd به منظور پیش‌پردازش داده‌ها استفاده شد. این تابع، داده‌ها را به توزیع گوسی استاندارد تبدیل می‌کند. اصولاً پیش‌پردازش داده‌ها باعث می‌شود که شبکه

---

1. Sorting  
4. Early stopping

2. Overtraining

3. Regularization

الگوی بین ورودی‌ها و خروجی‌ها را بهتر بیاموزد. تعداد دوره‌های تکرار<sup>۱</sup> برای آموزش شبکه‌ها در هر ۳ روش ۴۰۰ در نظر گرفته شد. بعد از سعی و خطای طولانی تابع انتقال یا فعال‌سازی نرون‌های لایه-های میانی تانژانت هیپربولیک و تابع انتقال نرون‌های لایه‌ی خروجی تابع خطی انتخاب شد. مراحل کلی این تحقیق که در ادامه به شرح آنها می‌پردازیم در نمای شماتیک نشان داده شده در شکل ۴-۵ ارائه شده است. می‌توان کارهای انجام شده را در ۳ بخش کلی ترسیم کرد: ۱- طراحی شبکه‌های پس انتشار خطا به روش فرا آموزش و ترکیب نتایج، ۲- طراحی شبکه‌های پس انتشار خطا به روش مرتب‌سازی و ترکیب نتایج، ۳- طراحی شبکه‌های پس انتشار خطا به روش توقف سریع و ترکیب نتایج. در هر بخش، شبکه‌هایی با یک و دو لایه‌ی میانی با تعداد نرون‌های مختلف ایجاد شدند و برای هر ساختار خاص (تعداد مشخص لایه‌ها و نرون‌ها) یک مرحله طولانی آزمون و خطا طی و بهترین نتیجه با معیار کمترین میانگین مربعات خطا (MSE) برای داده‌های آزمون، ثبت شد (رابطه ۴-۱)). سپس در هر یک از روش‌های آموزش، ۷ شبکه با بهترین نتایج (از لحاظ داشتن کمترین خطای تخمین برای داده‌های آزمون)، برای ساختن ترکیبات آنسامبلی انتخاب شدند. این ترکیبات به صورت خطی و به ۲ روش ساخته شدند: ۱- الگوریتم ژنتیک (GA) و ۲- میانگین‌گیری ساده.

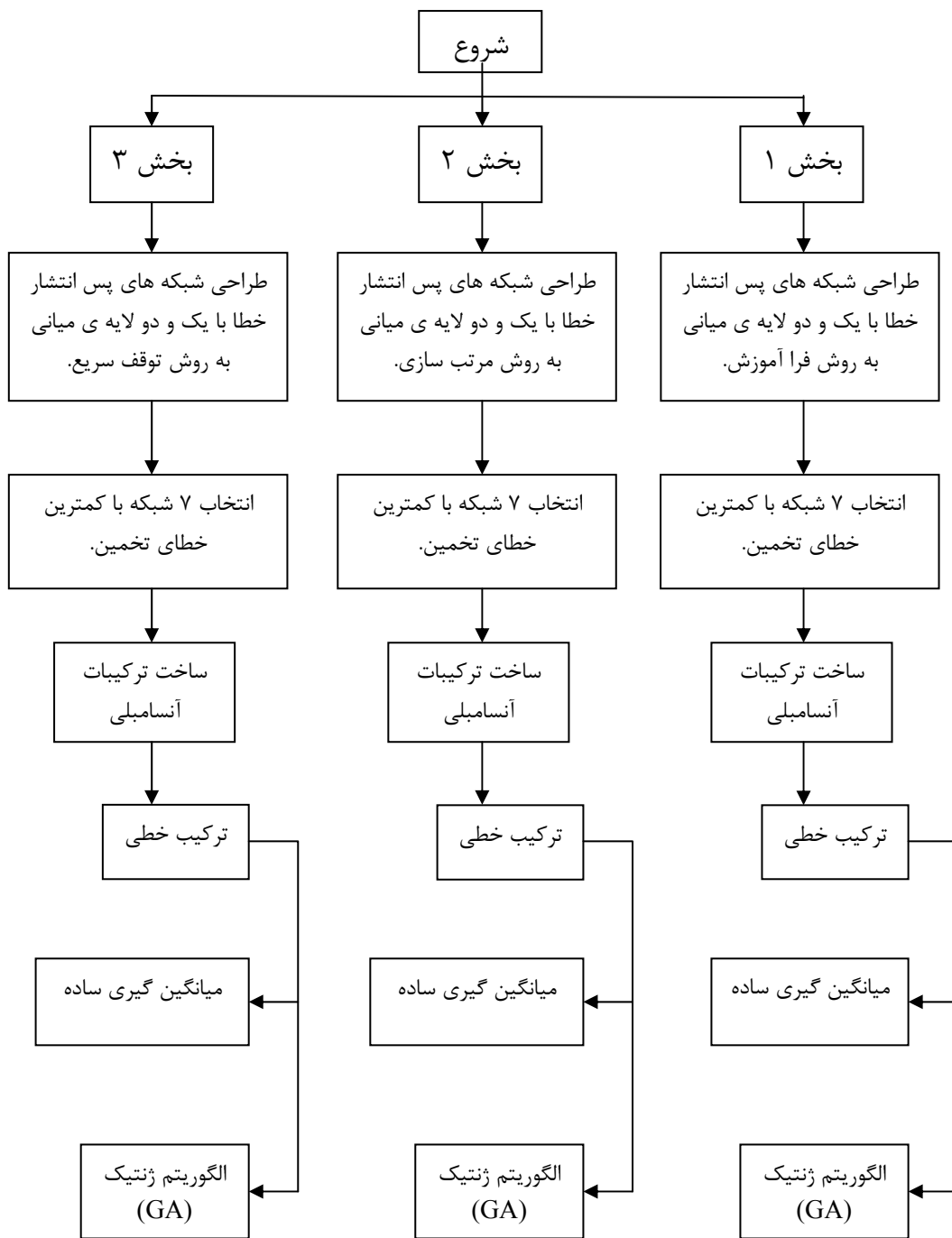
در هر یک از این ۲ روش ۱۲۰ ترکیب آنسامبلی ممکن این ۷ شبکه ساخته شدند: ۲۱ ترکیب ۲ شبکه‌ای، ۳۵ ترکیب ۳ شبکه‌ای، ۳۵ ترکیب ۴ شبکه‌ای، ۲۱ ترکیب ۵ شبکه‌ای، ۷ ترکیب ۶ شبکه‌ای و ۱ ترکیب ۷ شبکه‌ای. از بین این ترکیبات بهترین ترکیب انتخاب شد.

$$MSE = \frac{1}{N} \sum_{i=1}^N (t(i) - o(i))^2 \quad \text{رابطه (۴-۱)}$$

در این رابطه  $t$  بردار مطلوب،  $O$  بردار خروجی یا تخمین زده شده و  $N$  تعداد بردارهای موجود می‌باشد.

---

1. Epochs



شکل ۴-۵: نمای شماتیک مراحل طی شده در این تحقیق.

#### ۴-۷- ایجاد شبکه‌های پس انتشار خطا به روش فرا آموزش و ترکیب نتایج

در این قسمت از تابع لونبرگ - مارکوت<sup>۱</sup> (Trainlm) برای آموزش شبکه‌ها استفاده شد. این شبکه‌ها با یک و دو لایه‌ی میانی و تعداد نرون‌های مختلف، با شرایط اولیه گوناگون طی یک مرحله‌ی طولانی آزمون و خطا مورد آزمون قرار گرفتند و برای هر ساختار خاص بهترین نتیجه ثبت شد. افزایش تعداد نرون‌ها تا جایی بود که بیش برآزش رخ می‌داد. در جدول ۴-۲ نتایج به طور کامل ارائه شده است. ستون‌های این جدول از سمت راست عبارتند از شماره‌ی ردیف، ساختار شبکه، میانگین مربعات خطای تخمین شبکه برای داده‌های آموزش، ضریب همبستگی بین پارامتر اشباع از آب و تخمین شبکه برای داده‌های آموزش، میانگین مربعات خطای تخمین شبکه برای داده‌های آموزش، ضریب همبستگی بین پارامتر اشباع از آب و تخمین شبکه برای داده‌های آزمون و تعداد کل پارامترهای شبکه.

جدول ۴-۲: نتایج مربوط به شبکه‌های ساخته شده با یک و دو لایه‌ی میانی به روش فرا آموزش.

Epochs=400, goal=0, min\_grad=1e-10, mem\_reduc=1, mu=0.001, mu\_dec=0.1, mu\_inc=10, mu\_max=1e10.

تعداد کل پارامترهای شبکه	ضریب همبستگی داده‌های آزمون	(MSE) داده‌های آزمون (%)	ضریب همبستگی داده‌های آموزش	(MSE) داده‌های آموزش (%)	ساختار شبکه	ردیف
۴۹	۰/۹۵۱۷	۰/۶۶۴۷	۰/۹۵۷۰	۰/۳۳۵۴	۶-۶-۱	۱
۵۷	۰/۹۵۰۹	۰/۶۸۳۹	۰/۹۶۰۴	۰/۳۰۹۴	۶-۷-۱	۲
۶۵	۰/۹۶۷۱	۰/۴۲۹۶	۰/۹۷۵۰	۰/۱۹۶۴	۶-۸-۱	۳
۷۳	۰/۹۷۱۷	۰/۴۰۸۸	۰/۹۸۴۲	۰/۱۲۴۶	۶-۹-۱	۴
۸۱	۰/۹۷۷۴	۰/۳۰۹۵	۰/۹۸۲۲	۰/۱۴۰۱	۶-۱۰-۱	۵
۸۹	۰/۹۷۶۷	۰/۳۴۳۳	۰/۹۸۲۵	۰/۱۳۷۷	۶-۱۱-۱	۶

1. Levenberg - marquardt

97	•/9723	•/3744	•/9850	•/1187	6-12-1	7
105	•/9751	•/3418	•/9845	•/1219	6-13-1	8
113	•/9719	•/3867	•/9855	•/1141	6-14-1	9
121	•/9635	•/4844	•/9892	•/0854	6-15-1	10
47	•/9483	•/6784	•/9584	•/3244	6-4-3-1	11
53	•/9465	•/6977	•/9568	•/3363	6-4-4-1	12
59	•/9512	•/6715	•/9656	•/2692	6-4-5-1	13
65	•/9555	•/6057	•/9670	•/2584	6-4-6-1	14
71	•/9668	•/4387	•/9707	•/2301	6-4-7-1	15
77	•/9699	•/4132	•/9792	•/1635	6-4-8-1	16
83	•/9733	•/3553	•/9821	•/1415	6-4-9-1	17
89	•/9702	•/3901	•/9827	•/1362	6-4-10-1	18
95	•/9601	•/5844	•/9887	•/0893	6-4-11-1	19
101	•/9499	•/6731	•/9903	•/0766	6-4-12-1	20
107	•/9596	•/6818	•/9925	•/0594	6-4-13-1	21
113	•/9404	•/8046	•/9912	•/1481	6-4-14-1	22
57	•/9543	•/5939	•/9666	•/2616	6-5-3-1	23
64	•/9589	•/5330	•/9614	•/3011	6-5-4-1	24
71	•/9611	•/5179	•/9642	•/2801	6-5-5-1	25
78	•/9687	•/4146	•/9788	•/1668	6-5-6-1	26
85	•/9664	•/4431	•/9827	•/1360	6-5-7-1	27
92	•/9649	•/4556	•/9864	•/1074	6-5-8-1	28
99	•/9589	•/6119	•/9892	•/0854	6-5-9-1	29
106	•/9490	•/6916	•/9870	•/1028	6-5-10-1	30
113	•/9443	•/8201	•/9920	•/0629	6-5-11-1	31
67	•/9684	•/4192	•/9811	•/1490	6-6-3-1	32
75	•/9697	•/4093	•/9796	•/1609	6-6-4-1	33
83	•/9775	•/3105	•/9809	•/1508	6-6-5-1	34
91	•/9714	•/3802	•/9833	•/1318	6-6-6-1	35
99	•/9675	•/4231	•/9858	•/1120	6-6-7-1	36
107	•/9628	•/4958	•/9928	•/0566	6-6-8-1	37
115	•/9629	•/5046	•/9947	•/0415	6-6-9-1	38

۱۲۳	۰/۹۵۰۷	۰/۷۹۴۳	۰/۹۹۳۴	۰/۰۵۲۴	۶-۶-۱۰-۱	۳۹
۷۷	۰/۹۵۶۰	۰/۶۰۹۶	۰/۹۶۴۶	۰/۲۷۷۱	۶-۷-۳-۱	۴۰
۸۶	۰/۹۶۲۷	۰/۵۲۰۵	۰/۹۸۲۵	۰/۱۳۷۹	۶-۷-۴-۱	۴۱
۹۵	۰/۹۷۲۵	۰/۳۶۱۷	۰/۹۸۳۴	۰/۱۳۱۲	۶-۷-۵-۱	۴۲
۱۰۴	۰/۹۷۰۷	۰/۳۹۴۳	۰/۹۸۶۸	۰/۱۰۳۱	۶-۷-۶-۱	۴۳
۱۱۳	۰/۹۶۵۰	۰/۵۱۰۱	۰/۹۸۶۷	۰/۱۰۴۸	۶-۷-۷-۱	۴۴
۱۲۲	۰/۹۵۶۲	۰/۷۰۸۵	۰/۹۹۳۵	۰/۰۵۱۶	۶-۷-۸-۱	۴۵
۱۳۱	۰/۹۵۳۸	۰/۷۰۸۵	۰/۹۹۵۴	۰/۰۳۶۶	۶-۷-۹-۱	۴۶
۱۴۰	۰/۹۴۲۸	۰/۸۳۲۸	۰/۹۹۶۱	۰/۰۳۰۴	۶-۷-۱۰-۱	۴۷
۱۴۹	۰/۹۳۴۹	۰/۹۴۸۲	۰/۹۹۷۱	۰/۰۲۳۰	۶-۷-۱۱-۱	۴۸
۸۷	۰/۹۵۳۵	۰/۶۲۹۸	۰/۹۶۸۴	۰/۲۴۷۹	۶-۸-۳-۱	۴۹
۹۷	۰/۹۶۰۴	۰/۵۲۴۴	۰/۹۷۶۹	۰/۱۸۱۳	۶-۸-۴-۱	۵۰
۱۰۷	۰/۹۷۰۹	۰/۳۹۷۶	۰/۹۸۴۶	۰/۱۲۱۵	۶-۸-۵-۱	۵۱
۱۱۷	۰/۹۶۵۷	۰/۴۷۷۷	۰/۹۸۶۱	۰/۱۰۹۹	۶-۸-۶-۱	۵۲
۱۲۷	۰/۹۶۵۲	۰/۴۷۳۷	۰/۹۸۷۱	۰/۱۰۲۱	۶-۸-۷-۱	۵۳
۱۳۷	۰/۹۶۲۴	۰/۵۱۳۱	۰/۹۹۷۱	۰/۰۲۳۱	۶-۸-۸-۱	۵۴
۱۴۷	۰/۹۴۶۳	۰/۸۱۹۵	۰/۹۹۸۱	۰/۰۱۵۱	۶-۸-۹-۱	۵۵
۹۷	۰/۹۶۶۵	۰/۴۷۲۳	۰/۹۸۳۶	۰/۱۲۹۷	۶-۹-۳-۱	۵۶
۱۰۸	۰/۹۵۴۳	۰/۶۷۸۳	۰/۹۹۳۲	۰/۰۵۳۹	۶-۹-۴-۱	۵۷
۱۱۹	۰/۹۵۴۰	۰/۶۵۴۳	۰/۹۹۳۹	۰/۰۴۸۱	۶-۹-۵-۱	۵۸
۹۵	۰/۹۴۸۷	۰/۷۱۲۶	۰/۹۸۸۲	۰/۰۹۳۰	۶-۱۰-۲-۱	۵۹

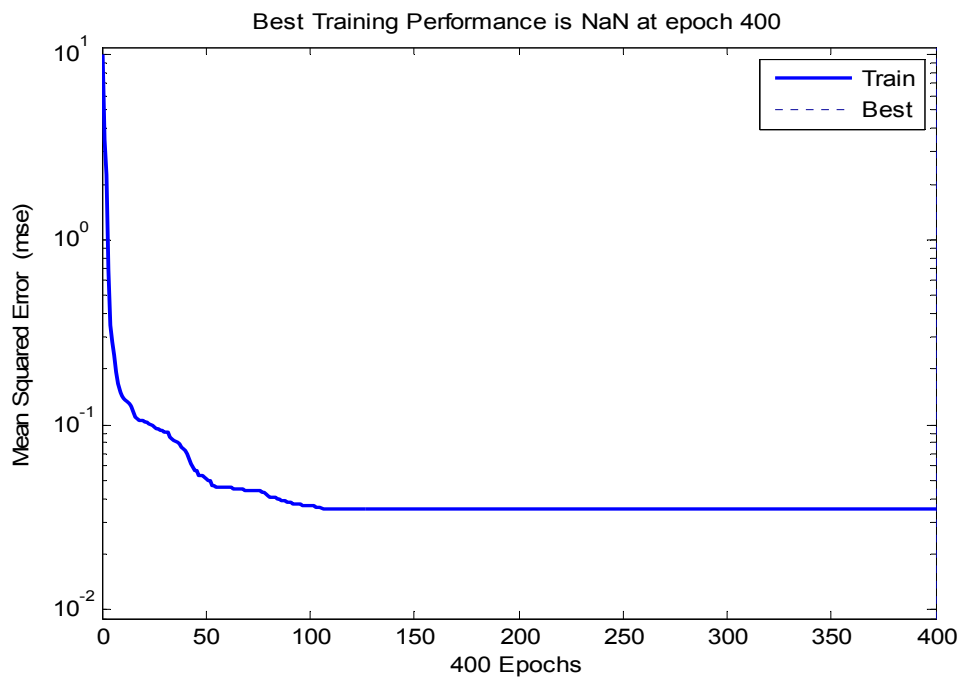
همانطور که در جدول مشاهده می‌شود شبکه مربوط به ردیف ۵ (شبکه ۱ لایه با ۱۰ نرون در لایه‌ی نهان) بهترین شبکه از لحاظ داشتن کمترین میزان خطای تخمین برای داده‌های آزمون است. نتایج مربوط به ۷ شبکه با بهترین نتایج در جدول ۴-۳ ارائه شده‌اند. از این نتایج برای مراحل بعدی و ایجاد ترکیبات استفاده شده است. همانطور که مشاهده می‌شود شبکه‌ها بر اساس خطای آزمون مرتب شده‌اند.

جدول ۳-۴: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش فرا آموزش بر اساس داشتن کمترین خطای آزمون.

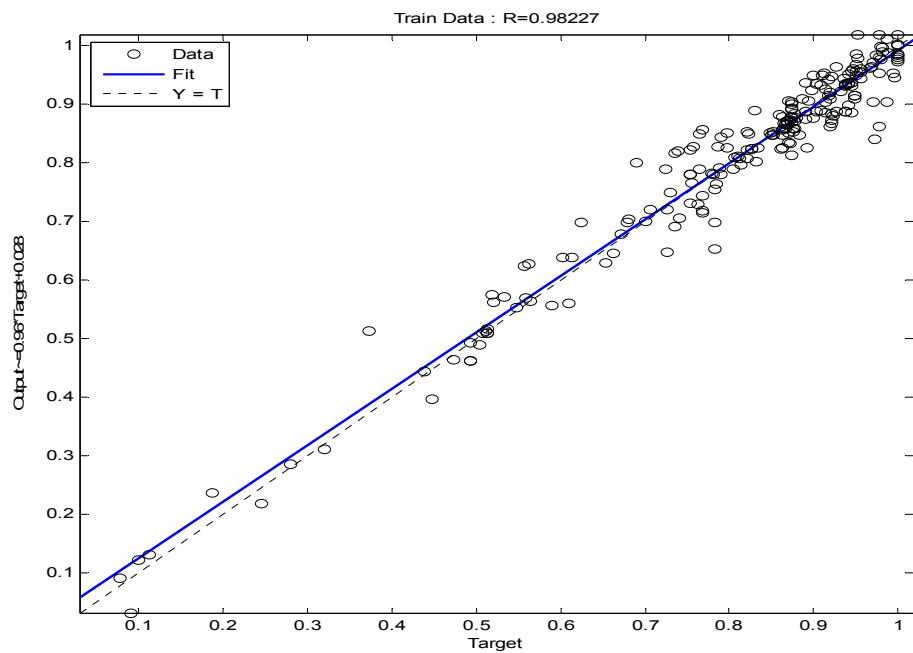
ردیف	ساختار شبکه	(MSE) داده های آموزش (%)	ضریب همبستگی داده های آموزش	(MSE) داده های آزمون (%)	ضریب همبستگی داده های آزمون	تعداد کل پارامترهای شبکه
۱	۶-۱۰-۱	۰/۱۴۰۱	۰/۹۸۲۲	۰/۳۰۹۵	۰/۹۷۷۴	۸۱
۲	۶-۶-۵-۱	۰/۱۵۰۸	۰/۹۸۰۹	۰/۳۱۰۵	۰/۹۷۷۵	۸۳
۳	۶-۱۳-۱	۰/۱۲۱۹	۰/۹۸۴۵	۰/۳۴۱۸	۰/۹۷۵۱	۱۰۵
۴	۶-۱۱-۱	۰/۱۳۷۷	۰/۹۸۲۵	۰/۳۴۳۳	۰/۹۷۶۷	۸۹
۵	۶-۴-۹-۱	۰/۱۴۱۵	۰/۹۸۲۱	۰/۳۵۵۳	۰/۹۷۳۳	۸۳
۶	۶-۷-۵-۱	۰/۱۳۱۲	۰/۹۸۳۴	۰/۳۶۱۷	۰/۹۷۲۵	۹۵
۷	۶-۱۲-۱	۰/۱۱۸۷	۰/۹۸۵۰	۰/۳۷۴۴	۰/۹۷۲۳	۹۷

در شکل ۶-۴ نتایج مربوط به بهترین شبکه منفرد نمایش داده شده است (شبکه شماره ۱ جدول ۴-۴).

(۳)

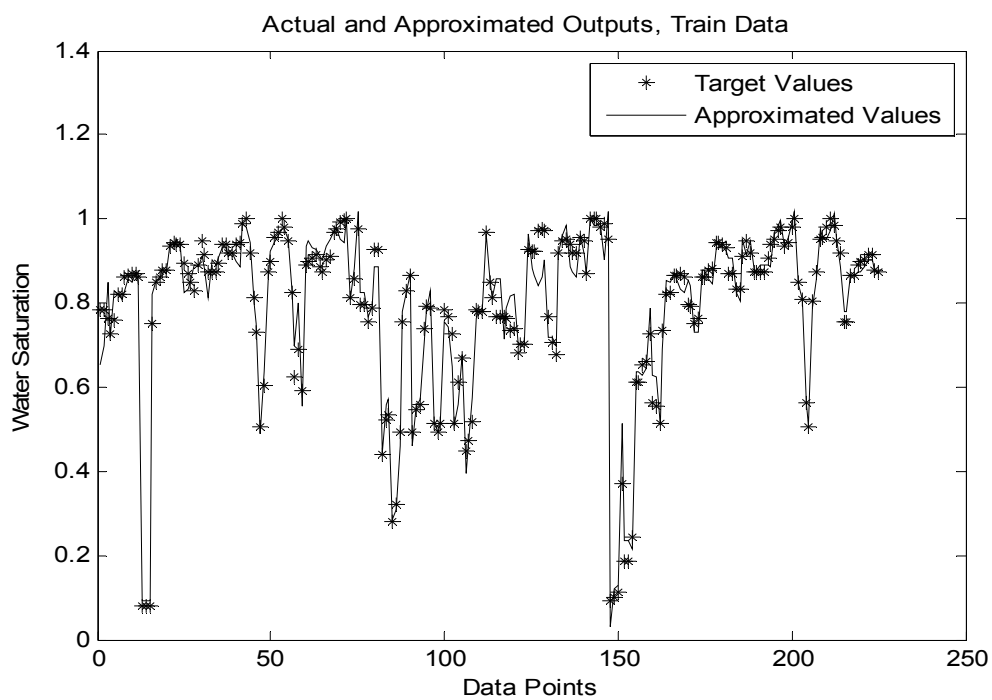


شکل ۴-۶-الف: روند کاهش خطا برای مرحله‌ی آموزش شبکه‌ی شماره ۱ جدول ۳-۴.

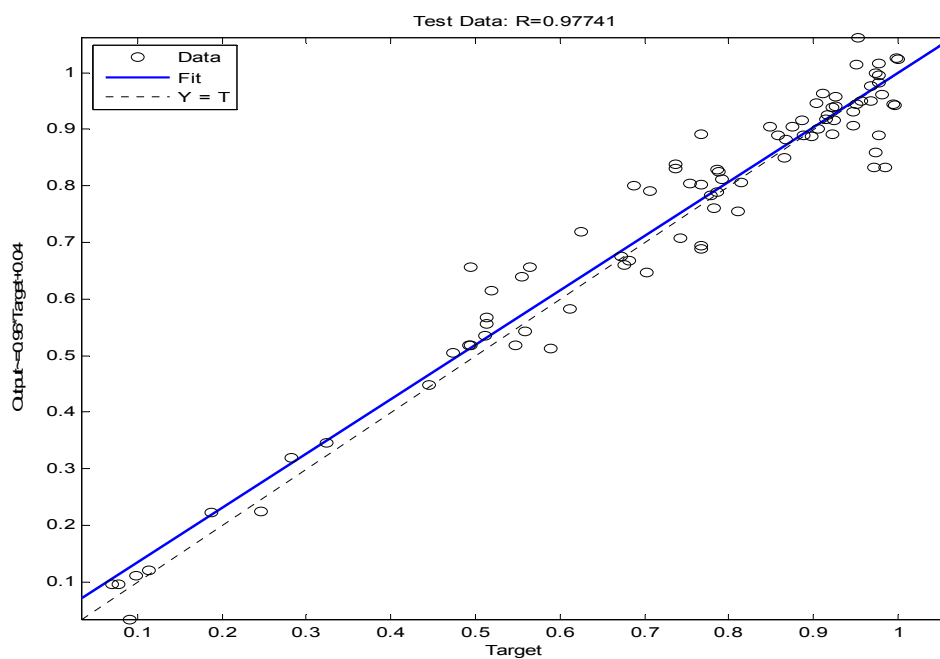


شکل ۴-۶-ب: نمودار همبستگی بین اشباع از آب و تخمین شبکه‌ی شماره ۱ جدول ۳-۴ برای داده‌های آموزش.

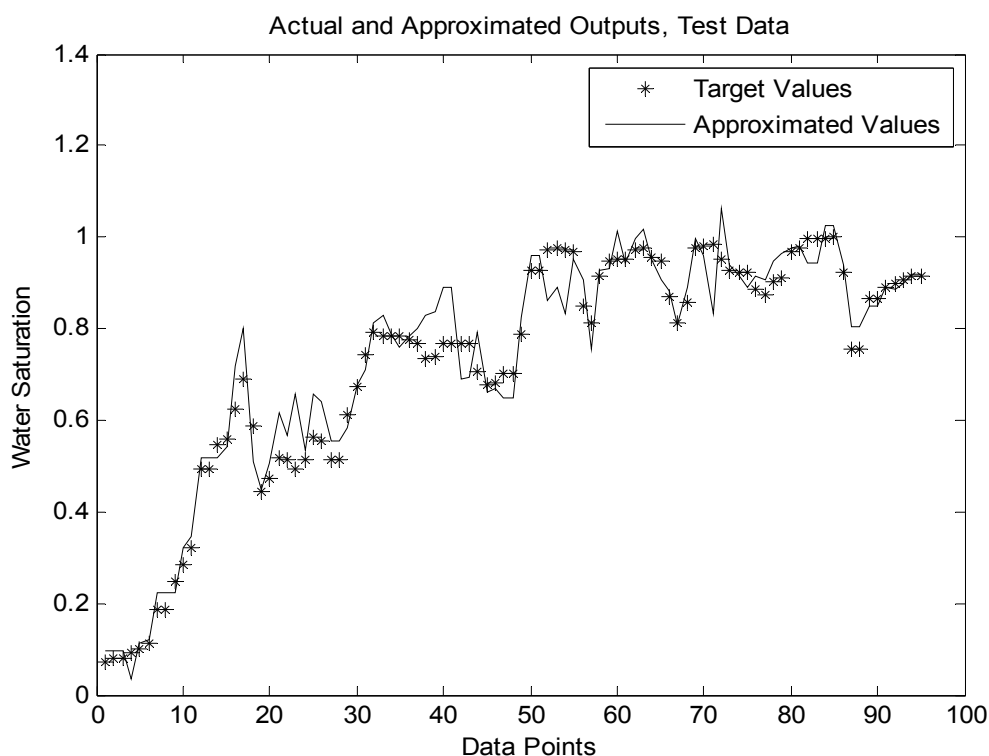




شکل ۴-۶-ج: مقادیر اشباع از آب و تخمین شبکه شماره ۱ جدول ۳-۴ برای داده‌های آموزش.



شکل ۴-۶-د: نمودار همبستگی بین اشباع از آب و تخمین شبکه‌ی شماره ۱ جدول ۳-۴ برای داده‌های آزمون.



شکل ۴-۶-۵: مقادیر اشباع از آب و تخمین شبکه‌ی شماره ۱ جدول ۴-۳ برای داده‌های آزمون.

#### ۴-۷-۱- ترکیبات آنسامبلی خطی

همان‌طور که قبلاً اشاره کردیم ۱۲۰ ترکیب خطی آنسامبلی شبکه‌های منفرد جدول ۴-۳ به دو روش الگوریتم ژنتیک و میانگین‌گیری ساده ایجاد شدند. نتایج این ۱۲۰ ترکیب در جدول ۴-۴ ارائه شده است. ستون اول از سمت چپ (NO)، شماره ردیف و ستون دوم (Ensemble)، نوع ترکیب آنسامبلی را نشان می‌دهد. ستون‌های دیگر میانگین مربعات خطای تخمین را برای داده‌های آموزش و آزمون برای هر دو روش نشان می‌دهد. این خطاها تا ۴ رقم اعشار گرد شده‌اند.

برای الگوریتم ژنتیک جمعیت اولیه و همچنین تعداد مراحل تکرار ۱۰۰ در نظر گرفته شد. محدوده ضرایب نیز بین ۰ و ۱ تعیین شد. ۸۰ درصد جمعیت برای عمل تقاطع و ۲۰ درصد برای جهش در نظر گرفته شد. همچنین برای انتخاب از روش چرخ رولت استفاده کردیم.

جدول ۴-۴: نتایج ترکیبات آنسامبلی خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده.

NO.	Ensemble	GA-train(%)	GA-test(%)	Simple Avg-train(%)	Simple Avg-test(%)
1	1,2	0.1028	0.1846	0.1029	0.1841
2	1,3	0.0880	0.2153	0.0885	0.2123
3	1,4	0.0999	0.2079	0.0999	0.2076
4	1,5	0.0917	0.2153	0.0917	0.2154
5	1,6	0.0848	0.2159	0.0849	0.2145
6	1,7	0.0861	0.2338	0.0868	0.2279
7	2,3	0.0955	0.1951	0.0969	0.1875
8	2,4	0.1001	0.2320	0.1004	0.2302
9	2,5	0.1127	0.204	0.1129	0.2018
10	2,6	0.0988	0.2292	0.0994	0.2246
11	2,7	0.0965	0.2333	0.0983	0.2203
12	3,4	0.0906	0.1830	0.0910	0.1814
13	3,5	0.0884	0.2270	0.0890	0.2262
14	3,6	0.0944	0.2618	0.0945	0.2620
15	3,7	0.0902	0.2606	0.0902	0.2601
16	4,5	0.0895	0.1969	0.0896	0.1969
17	4,6	0.0901	0.2476	0.0901	0.2471
18	4,7	0.0900	0.2109	0.0906	0.2065
19	5,6	0.0959	0.2574	0.0962	0.2568
20	5,7	0.0901	0.2768	0.0909	0.2735
21	6,7	0.0905	0.2592	0.0908	0.2577
22	1,2,3	0.0810	0.1617	0.0823	0.1527
23	1,2,4	0.0871	0.1707	0.0872	0.1694
24	1,2,5	0.0870	0.1680	0.0886	0.1589
25	1,2,6	0.0791	0.1766	0.0807	0.1679
26	1,2,7	0.0807	0.1851	0.0825	0.1706
27	1,3,4	0.0790	0.1593	0.0798	0.1568
28	1,3,5	0.0746	0.1797	0.0748	0.1788
29	1,3,6	0.0753	0.1923	0.0754	0.1936
30	1,3,7	0.0757	0.1986	0.0757	0.1973
31	1,4,5	0.0782	0.1642	0.0784	0.1635
32	1,4,6	0.0761	0.1878	0.0768	0.1847
33	1,4,7	0.0782	0.1781	0.0792	0.1712
34	2,3,4	0.0817	0.1546	0.0825	0.1557
35	2,3,5	0.0841	0.1761	0.0867	0.1616
36	2,3,6	0.0840	0.1914	0.0844	0.1870
37	2,3,7	0.0825	0.1931	0.0833	0.1828
38	2,4,5	0.0849	0.1715	0.0868	0.1674
39	2,4,6	0.0815	0.2039	0.0822	0.1991
40	2,4,7	0.0824	0.1801	0.0833	0.1778
41	2,5,6	0.0887	0.2000	0.0900	0.1895
42	2,5,7	0.0858	0.2177	0.0886	0.1936
43	3,4,5	0.0752	0.1538	0.0752	0.1531
44	3,4,6	0.0791	0.1896	0.0791	0.1906
45	3,4,7	0.0787	0.1735	0.0788	0.1703
46	3,5,6	0.0801	0.2122	0.0805	0.2135

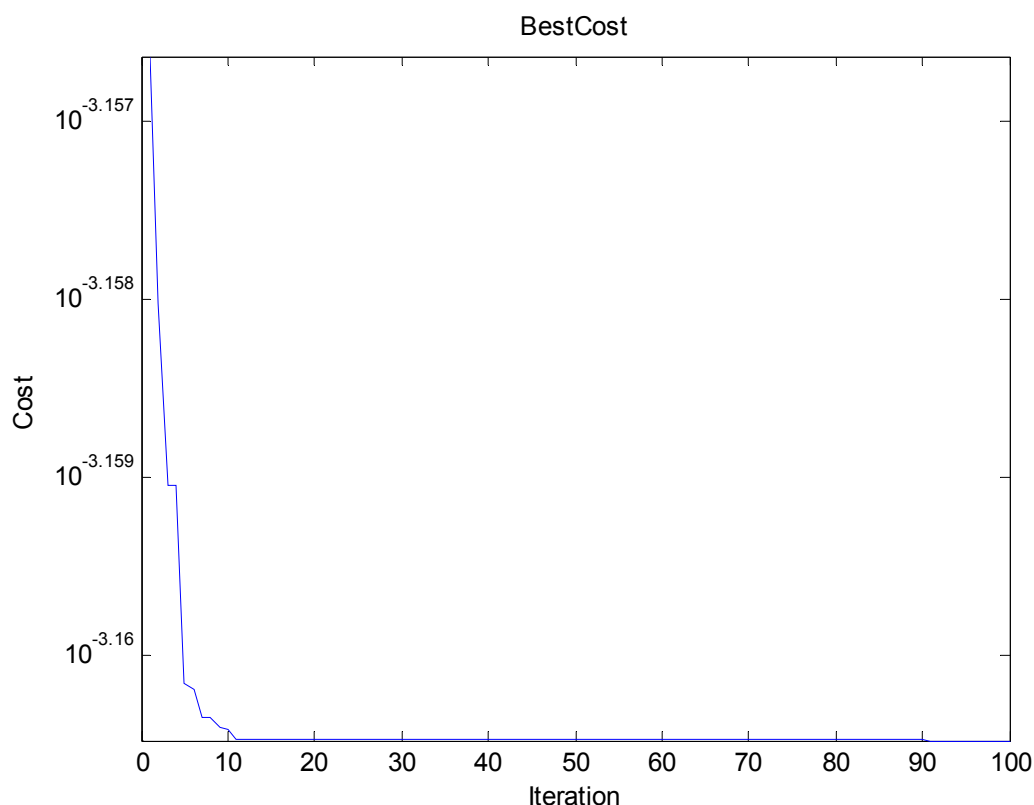
47	3,5,7	0.0775	0.2189	0.0776	0.2186
48	4,5,6	0.0769	0.1931	0.0770	0.1937
49	4,5,7	0.0762	0.1832	0.0763	0.1816
50	1,5,6	0.0750	0.1900	0.0754	0.1912
51	1,5,7	0.0750	0.2057	0.0753	0.2031
52	1,6,7	0.0733	0.1948	0.0733	0.1950
53	2,6,7	0.0828	0.2041	0.0837	0.1960
54	3,6,7	0.0810	0.2275	0.0812	0.2268
55	4,6,7	0.0776	0.1965	0.0776	0.1962
56	5,6,7	0.0796	0.2309	0.0800	0.2289
57	4,5,6,7	0.0708	0.1786	0.0709	0.1803
58	3,5,6,7	0.0736	0.2055	0.0737	0.2049
59	3,4,6,7	0.0731	0.1752	0.0731	0.1761
60	3,4,5,7	0.0703	0.1591	0.0703	0.1593
61	3,4,5,6	0.0709	0.1631	0.0711	0.1674
62	2,5,6,7	0.0776	0.1994	0.0793	0.1834
63	2,4,6,7	0.0744	0.1766	0.0751	0.1729
64	2,4,5,7	0.0750	0.1680	0.0771	0.1594
65	2,4,5,6	0.0752	0.1768	0.0770	0.1680
66	2,3,6,7	0.0768	0.1865	0.0772	0.1795
67	2,3,5,7	0.0760	0.1886	0.0779	0.1696
68	2,3,5,6	0.0777	0.1792	0.0790	0.1686
69	1,5,6,7	0.0686	0.1844	0.0689	0.1863
70	1,4,6,7	0.0694	0.1700	0.0698	0.1667
71	1,4,5,7	0.0699	0.1641	0.0702	0.1592
72	1,4,5,6	0.0692	0.1654	0.0693	0.1634
73	1,3,6,7	0.0697	0.1829	0.0699	0.1852
74	1,3,5,7	0.0690	0.1810	0.0690	0.1812
75	1,3,5,6	0.0692	0.1738	0.0694	0.1758
76	1,3,4,7	0.0716	0.1588	0.0720	0.1528
77	1,3,4,6	0.0707	0.1647	0.0709	0.1617
78	1,2,6,7	0.0716	0.1731	0.0732	0.1628
79	1,2,5,7	0.0742	0.1850	0.0770	0.1620
80	1,2,5,6	0.0739	0.1706	0.0766	0.1572
81	1,2,4,7	0.0752	0.1582	0.0763	0.1520
82	1,2,4,6	0.0732	0.1677	0.0744	0.1614
83	1,2,3,7	0.0734	0.1683	0.0745	0.1560
84	1,2,3,6	0.0727	0.1644	0.0738	0.1558
85	2,3,4,7	0.0751	0.1533	0.0757	0.1503
86	2,3,4,6	0.0749	0.1658	0.0754	0.1636
87	2,3,4,5	0.0739	0.1410	0.0759	0.1371
88	1,3,4,5	0.0695	0.1433	0.0698	0.1412
89	1,2,4,5	0.0767	0.1492	0.0781	0.1442
90	1,2,3,5	0.0737	0.1589	0.0762	0.1422
91	1,2,3,4	0.0751	0.1392	0.0761	0.1377
92	3,4,5,6,7	0.0678	0.1621	0.0679	0.1657
93	2,4,5,6,7	0.0703	0.1688	0.0719	0.1610
94	2,3,5,6,7	0.0726	0.1842	0.0738	0.1700
95	2,3,4,6,7	0.0711	0.1596	0.0715	0.1566
96	2,3,4,5,7	0.0699	0.1498	0.0715	0.1425
97	2,3,4,5,6	0.0702	0.1532	0.0716	0.1488
98	1,4,5,6,7	0.0655	0.1608	0.0656	0.1593

99	1,3,5,6,7	0.0659	0.1733	0.0662	0.1759
100	1,3,4,6,7	0.0670	0.1594	0.0672	0.1567
101	1,3,4,5,7	0.0660	0.1495	0.0661	0.1463
102	1,3,4,5,6	0.0657	0.1504	0.0658	0.1499
103	1,2,5,6,7	0.0683	0.1750	0.0709	0.1589
104	1,2,4,6,7	0.0684	0.1585	0.0697	0.1514
105	1,2,4,5,7	0.0696	0.1562	0.0716	0.1431
106	1,2,4,5,6	0.0688	0.1576	0.0707	0.1470
107	1,2,3,6,7	0.0687	0.1655	0.0698	0.1564
108	1,2,3,5,7	0.0688	0.1704	0.0709	0.1505
109	1,2,3,5,6	0.0688	0.1623	0.0708	0.1482
110	1,2,3,4,7	0.0701	0.1443	0.0710	0.1373
111	1,2,3,4,6	0.0692	0.1504	0.0699	0.1442
<b>112</b>	<b>1,2,3,4,5</b>	<b>0.0691</b>	<b>0.1356</b>	<b>0.0710</b>	<b>0.1277</b>
113	1,2,3,4,5,6	0.0656	0.1459	0.0672	0.1363
114	1,2,3,4,5,7	0.0660	0.1455	0.0677	0.1330
115	1,2,3,4,6,7	0.0664	0.1498	0.0672	0.1425
116	1,2,3,5,6,7	0.0658	0.1676	0.0678	0.1524
117	1,2,4,5,6,7	0.0654	0.1573	0.0673	0.1456
118	1,3,4,5,6,7	0.0636	0.1515	0.0638	0.1511
119	2,3,4,5,6,7	0.0675	0.1552	0.0688	0.1495
120	1,2,3,4,5,6,7	0.0636	0.1498	0.0652	0.1387

همان‌طور که مشاهده می‌شود ترکیب مربوط به ردیف ۱۱۲ (ترکیب شبکه‌های ۱، ۲، ۳، ۴ و ۵) برای هر دو روش بهترین نتایج را از جهت کمتر بودن خطای آزمون به همراه دارد. این نتایج به شکل کامل‌تری در جدول ۴-۵ ارائه شده‌اند. در شکل ۴-۷ نیز روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک ارائه شده است.

جدول ۴-۵: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک.

روش ترکیب خطی		نتایج
الگوریتم ژنتیک	میانگین گیری ساده	
۵ و ۴، ۳، ۲، ۱	۵ و ۴، ۳، ۲، ۱	ترکیب آنسامبلی
۰/۹۹۱۴	۰/۹۹۱۲	ضریب همبستگی داده های آموزش
۰/۰۶۹۱	۰/۰۷۱۰	MSE داده های آموزش (%)
%۵۰/۶۸	%۴۹/۳۲	کاهش در MSE داده های آموزش
۰/۹۸۹۷	۰/۹۹۰۳	ضریب همبستگی داده های آزمون
۰/۱۳۵۶	۰/۱۲۷۷	MSE داده های آزمون (%)
%۵۶/۱۹	%۵۷/۷۴	کاهش در MSE داده های آزمون
۰/۲۱۰۷	۰/۲	$W_1$
۰/۰۶۳۶	۰/۲	$W_2$
۰/۲۷۷۱	۰/۲	$W_3$
۰/۲۰۸۵	۰/۲	$W_4$
۰/۲۴۰۱	۰/۲	$W_5$
۰	۰	$W_6$
۰	۰	$W_7$
۱	۱	$\sum W$



شکل ۴-۷: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک.

در جدول ۴-۵ ردیف دوم بهترین ترکیب از بین ۱۲۰ ترکیب را برای هر دو روش میانگین گیری ساده و الگوریتم ژنتیک نشان می‌دهد. ردیف سوم ضریب همبستگی بین مقادیر اشباع از آب و تخمین ترکیب مورد نظر برای داده‌های آموزش، ردیف چهارم میانگین مربعات خطای تخمین برای داده‌های آموزش ترکیب مورد نظر و ردیف پنجم درصد کاهش میانگین مربعات خطا برای داده‌های آموزش را نسبت به بهترین شبکه منفرد (شبکه ۱ جدول ۴-۳) نشان می‌دهد. میزان کاهش میانگین مربعات خطای داده‌های آموزش از طریق تقسیم حاصل تفریق میانگین مربعات خطای ترکیب و بهترین شبکه‌ی منفرد بر میانگین مربعات خطای بهترین شبکه‌ی منفرد محاسبه شده است. ردیف‌های ششم، هفتم و هشتم به ترتیب ضریب همبستگی بین مقادیر اشباع از آب و تخمین ترکیب، میانگین مربعات خطای ترکیب و درصد کاهش میانگین مربعات خطا نسبت به بهترین شبکه‌ی منفرد (شبکه ۱ جدول ۴-۳) را برای داده‌های آزمون نشان می‌دهند. در ادامه‌ی جدول ضرایب مربوط به این شبکه‌ها برای هر دو روش ارائه شده است. همانطور که مشاهده می‌شود شبکه‌ی شماره ۲ جدول ۴-۳ کمترین ضریب و

شبهه شماره ۳ بیشترین ضریب را به خود اختصاص داده است. از طرف دیگر شاهد این هستیم که در روش میانگین گیری ساده میزان کاهش خطای تخمین ترکیب برای داده های آزمون نسبت به الگوریتم ژنتیک بیشتر می باشد. این بدین معنی است که در اینجا و با این شرایط روش میانگین گیری ساده برای بهترین ترکیب نتیجه بهتری را به همراه دارد.

#### ۴-۷-۲- واریانس و بایاس خطا

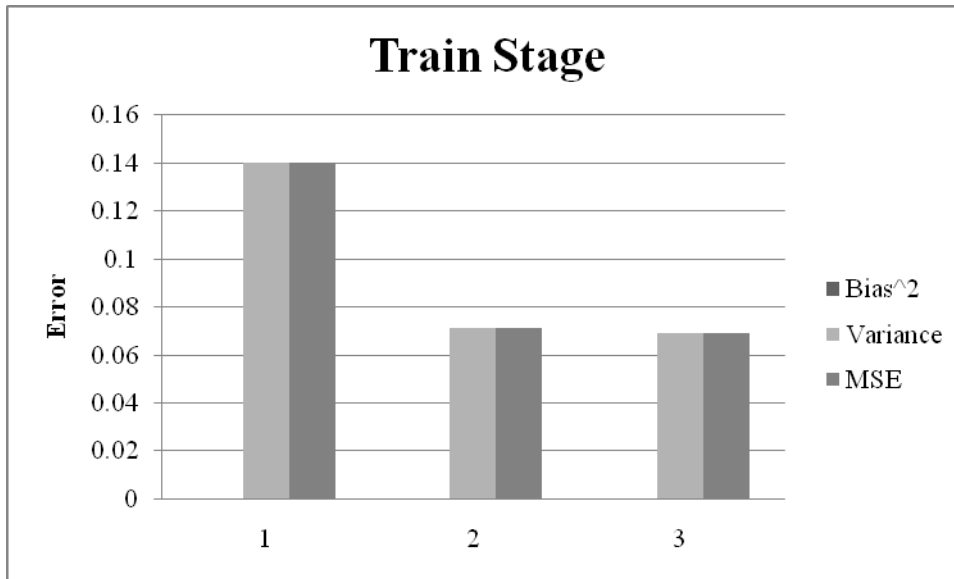
در جدول ۴-۶ واریانس، بایاس و میانگین مربعات خطای تخمین برای داده های آموزش و آزمون بهترین شبکه ی منفرد (۱) (شبهه ۱ جدول ۴-۳)، بهترین ترکیب با هر دو روش میانگین گیری ساده (۲) و الگوریتم ژنتیک (۳) ارائه شده است.

شکل ۴-۸ نمودار میله ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین برای داده های آموزش و آزمون حالات مربوط به جدول ۴-۶ را نشان می دهد. با توجه به شکل ۴-۸ الف برای داده های آموزش می بینیم که واریانس مؤلفه ی اصلی خطا است و بایاس تقریباً نقشی در خطا ندارد. برای داده های آزمون اگر چه نقش بایاس نسبت به داده های آموزش بیشتر است ولی باز هم می توان گفت که تقریباً نقشی در خطا ندارد و واریانس مؤلفه ی اصلی خطاست.

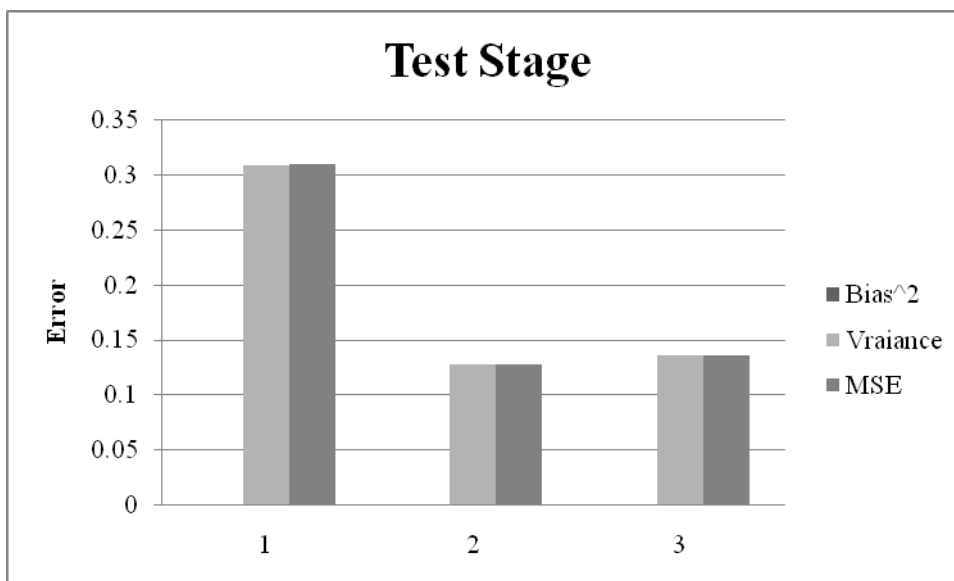
جدول ۴-۶: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه ۱ جدول ۴-۳ و بهترین ترکیب آن جدول.

ردیف	بایاس تخمین داده های آموزش	واریانس تخمین داده های آموزش	میانگین مربعات خطای تخمین داده های آموزش	بایاس تخمین داده های آزمون	واریانس تخمین داده های آزمون	میانگین مربعات خطای تخمین داده های آزمون
۱	$-1/13 \times 10^{-7}$	۰/۱۴۰۱	۰/۱۴۰۱	۰/۱۱۱۸	۰/۳۰۹۳	۰/۳۰۹۵
۲	$-1/05 \times 10^{-6}$	۰/۰۷۱۰	۰/۰۷۱۰	۰/۰۰۱۸	۰/۱۲۷۷	۰/۱۲۷۷
۳	$-7/59 \times 10^{-7}$	۰/۰۶۹۱	۰/۰۶۹۱	۰/۰۰۳۵	۰/۱۳۵۶	۰/۱۳۵۶





شکل ۴-۸- الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).



شکل ۴-۸- ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آزمون (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).

#### ۴-۸- ایجاد شبکه‌های پس انتشار خطا به روش مرتب‌سازی و ترکیب نتایج

در این مرحله برای آموزش شبکه‌ها از تابع مرتب‌سازی بیزین<sup>۱</sup> استفاده شد. این شبکه‌ها با یک و دو لایه‌ی میانی و تعداد نرون‌های مختلف، با شرایط اولیه گوناگون طی یک مرحله‌ی طولانی آزمون و خطا مورد آزمون قرار گرفتند و برای هر ساختار خاص بهترین نتیجه ثبت شد. ساختارهای شبکه‌ها در اینجا مانند ساختارهای مورد استفاده در روش فرا آموزش است. نتایج مربوط به شبکه‌های منفرد این بخش در جدول ۴-۷ ارائه شده است. ستون‌های این جدول مانند جدول ۴-۲ می‌باشند که قبلاً توضیح داده شده‌اند. تنها ستون اضافه شده مربوط به تعداد پارامترهای مؤثر<sup>۲</sup> شبکه می‌باشد.

جدول ۴-۷: نتایج مربوط به شبکه‌های پس انتشار خطای آموزش دیده با ۱ و ۲ لایه‌ی میانی به روش مرتب‌سازی.

Epochs=400, goal=0, min\_grad=1e-10, mem\_reduc=1, mu=0.005, mu\_dec=0.1, mu\_inc=10, mu\_max=1e10.

ردیف	ساختار شبکه	(MSE) داده‌های آموزش (%)	ضریب همبستگی داده‌های آموزش	(MSE) داده‌های آزمون (%)	ضریب همبستگی داده‌های آزمون	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۱	۶-۶-۱	۰/۴۴۶۶	۰/۹۵۲۹	۰/۶۳۲۲	۰/۹۴۲۳	۴۹	۴۵
۲	۶-۷-۱	۰/۳۶۰۴	۰/۹۵۷۲	۰/۵۶۱۹	۰/۹۵۳۷	۵۷	۵۳
۳	۶-۸-۱	۰/۲۶۸۱	۰/۹۶۵۸	۰/۵۲۴۲	۰/۹۶۱۸	۶۵	۶۱
۴	۶-۹-۱	۰/۲۲۶۱	۰/۹۷۱۲	۰/۵۱۱۵	۰/۹۶۱۳	۷۳	۶۹
۵	۶-۱۰-۱	۰/۱۶۷۸	۰/۹۷۸۷	۰/۴۱۱۵	۰/۹۶۹۹	۸۱	۷۶
۶	۶-۱۱-۱	۰/۱۵۹۹	۰/۹۷۹۷	۰/۴۵۸۲	۰/۹۶۴۷	۸۹	۸۲
۷	۶-۱۲-۱	۰/۱۵۹۹	۰/۹۸۳۴	۰/۴۵۸۲	۰/۹۷۰۷	۹۷	۸۹
۸	۶-۱۳-۱	۰/۱۲۴۱	۰/۹۸۴۳	۰/۳۴۹۸	۰/۹۷۵۴	۱۰۵	۹۳
۹	۶-۱۴-۱	۰/۱۲۳۷	۰/۹۸۴۳	۰/۳۶۸۴	۰/۹۷۲۳	۱۱۳	۹۹
۱۰	۶-۱۵-۱	۰/۱۰۱۲	۰/۹۸۷۲	۰/۴۲۰۳	۰/۹۷۰۵	۱۲۱	۱۰۴

1. Bayesian Regularization Function

2. Effective parameters

۴۴	۴۷	•/۹۴۶۴	•/۷۲۲۰	•/۹۴۸۹	•/۳۹۶۴	۶-۴-۳-۱	۱۱
۴۹	۵۳	•/۹۵۱۸	•/۶۹۳۵	•/۹۶۶۴	•/۲۶۳۲	۶-۴-۴-۱	۱۲
۵۵	۵۹	•/۹۶۳۳	•/۵۰۷۹	•/۹۶۸۹	•/۲۴۳۴	۶-۴-۵-۱	۱۳
۶۱	۶۵	•/۹۶۳۰	•/۵۱۵۷	•/۹۷۲۱	•/۲۱۸۸	۶-۴-۶-۱	۱۴
۶۸	۷۱	•/۹۷۲۵	•/۳۶۶۷	•/۹۸۱۹	•/۱۴۲۷	۶-۴-۷-۱	۱۵
۷۱	۷۷	•/۹۷۳۱	•/۳۹۲۶	•/۹۷۶۴	•/۱۸۵۷	۶-۴-۸-۱	۱۶
۷۶	۸۳	•/۹۷۵۷	•/۳۲۷۴	•/۹۸۱۶	•/۱۴۵۲	۶-۴-۹-۱	۱۷
۸۳	۸۹	•/۹۷۷۴	•/۳۱۳۸	•/۹۸۲۲	•/۱۴۰۶	۶-۴-۱۰-۱	۱۸
۸۷	۹۵	•/۹۷۳۰	•/۳۹۳۲	•/۹۸۶۱	•/۱۰۹۸	۶-۴-۱۱-۱	۱۹
۹۲	۱۰۱	•/۹۷۱۶	•/۴۰۲۴	•/۹۸۷۶	•/۰۹۸۰	۶-۴-۱۲-۱	۲۰
۱۰۲	۱۰۷	•/۹۷۱۶	•/۴۳۳۴	•/۹۹۴۳	•/۰۴۴۷	۶-۴-۱۳-۱	۲۱
۱۰۷	۱۱۳	•/۹۷۳۰	•/۴۰۷۱	•/۹۹۵۰	•/۰۳۹۲	۶-۴-۱۴-۱	۲۲
۵۳	۵۷	•/۹۶۸۳	•/۴۳۶۸	•/۹۷۲۱	•/۲۱۹۰	۶-۵-۳-۱	۲۳
۵۹	۶۴	•/۹۶۴۹	•/۴۷۶۱	•/۹۷۰۳	•/۲۳۳۰	۶-۵-۴-۱	۲۴
۶۸	۷۱	•/۹۷۴۵	•/۳۳۶۹	•/۹۸۰۷	•/۱۵۱۹	۶-۵-۵-۱	۲۵
۷۱	۷۸	•/۹۷۶۹	•/۳۱۲۱	•/۹۸۰۱	•/۱۵۶۷	۶-۵-۶-۱	۲۶
۸۰	۸۵	•/۹۷۴۵	•/۳۴۹۹	•/۹۸۴۳	•/۱۲۳۶	۶-۵-۷-۱	۲۷
۸۴	۹۲	•/۹۷۰۷	•/۳۸۸۲	•/۹۸۲۳	•/۱۳۹۴	۶-۵-۸-۱	۲۸
۹۰	۹۹	•/۹۷۰۵	•/۳۸۵۷	•/۹۸۶۶	•/۱۰۵۷	۶-۵-۹-۱	۲۹
۹۸	۱۰۶	•/۹۷۷۱	•/۳۰۲۲	•/۹۸۷۳	•/۱۰۰۷	۶-۵-۱۰-۱	۳۰
۱۰۶	۱۱۳	•/۹۶۶۰	•/۴۸۲۵	•/۹۹۲۸	•/۰۵۷۲	۶-۵-۱۱-۱	۳۱
۶۳	۶۷	•/۹۶۳۶	•/۴۸۵۴	•/۹۷۶۳	•/۱۸۶۵	۶-۶-۳-۱	۳۲
۶۹	۷۵	•/۹۶۵۷	•/۴۴۷۲	•/۹۷۶۵	•/۱۸۴۹	۶-۶-۴-۱	۳۳
۷۶	۸۳	•/۹۷۸۰	•/۳۱۱۹	•/۹۸۱۹	•/۱۴۲۶	۶-۶-۵-۱	۳۴
۸۳	۹۱	•/۹۷۵۵	•/۳۲۰۳	•/۹۸۳۴	•/۱۳۱۰	۶-۶-۶-۱	۳۵
۹۰	۹۹	•/۹۷۴۹	•/۳۵۱۳	•/۹۸۴۲	•/۱۲۴۸	۶-۶-۷-۱	۳۶
۱۰۱	۱۰۷	•/۹۷۰۷	•/۴۲۸۷	•/۹۹۲۳	•/۰۶۰۹	۶-۶-۸-۱	۳۷
۱۱۲	۱۱۵	•/۹۶۹۶	•/۴۵۰۲	•/۹۹۶۴	•/۰۲۷۹	۶-۶-۹-۱	۳۸
۱۱۲	۱۲۳	•/۹۶۶۳	•/۴۹۳۹	•/۹۹۴۱	•/۰۴۶۵	۶-۶-۱۰-۱	۳۹
۷۱	۷۷	•/۹۷۰۱	•/۳۹۵۱	•/۹۷۷۹	•/۱۷۳۷	۶-۷-۳-۱	۴۰
۸۰	۸۶	•/۹۷۱۱	•/۳۸۰۱	•/۹۸۱۲	•/۱۴۸۳	۶-۷-۴-۱	۴۱
۹۰	۹۵	•/۹۷۱۴	•/۳۹۶۱	•/۹۸۵۱	•/۱۱۷۵	۶-۷-۵-۱	۴۲

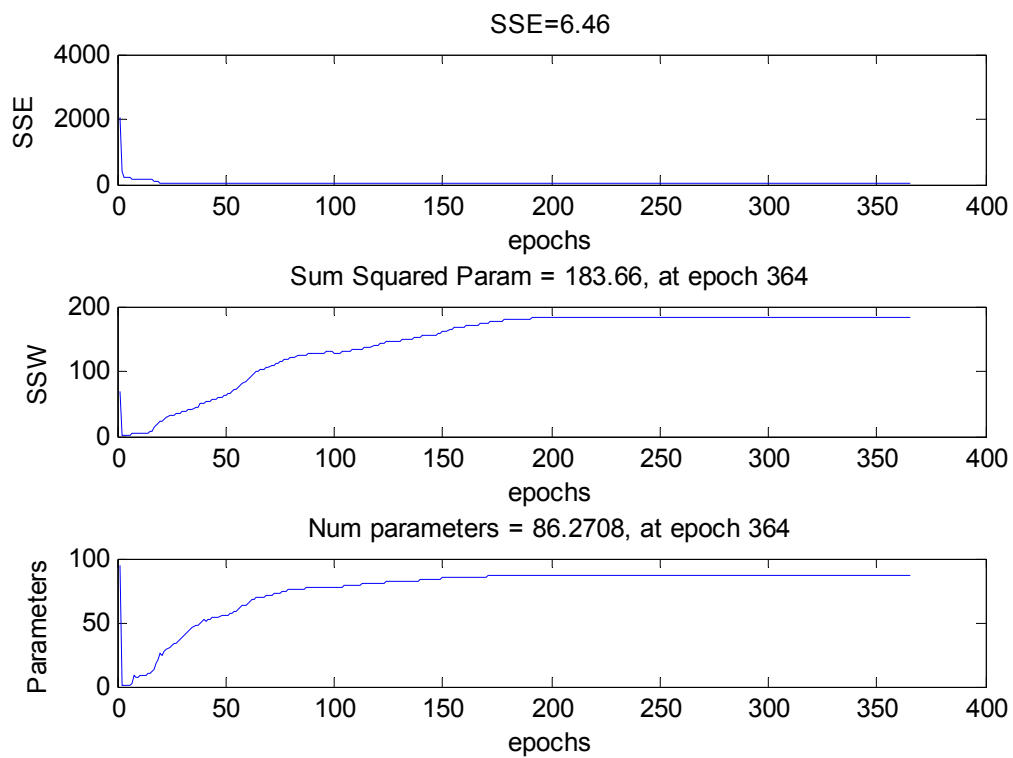
۹۳	۱۰۴	۰/۹۷۶۷	۰/۳۲۲۷	۰/۹۸۶۴	۰/۱۰۷۷	۶-۷-۶-۱	۴۳
۱۰۰	۱۱۳	۰/۹۷۰۱	۰/۴۲۰۷	۰/۹۸۸۷	۰/۰۸۹۶	۶-۷-۷-۱	۴۴
۱۱۷	۱۲۲	۰/۹۶۵۵	۰/۴۵۷۲	۰/۹۹۵۹	۰/۰۳۲۵	۶-۷-۸-۱	۴۵
۱۲۳	۱۳۱	۰/۹۶۷۰	۰/۴۵۴۰	۰/۹۹۵۶	۰/۰۳۴۵	۶-۷-۹-۱	۴۶
۱۲۳	۱۴۰	۰/۹۵۱۱	۰/۶۴۰۴	۰/۹۹۳۸	۰/۰۴۸۹	۶-۷-۱۰-۱	۴۷
۱۳۱	۱۴۹	۰/۹۵۱۶	۰/۸۵۹۹	۰/۹۹۶۲	۰/۰۲۹۶	۶-۷-۱۱-۱	۴۸
۷۹	۸۷	۰/۹۷۴۰	۰/۳۴۱۹	۰/۹۸۰۵	۰/۱۵۳۸	۶-۸-۳-۱	۴۹
۸۵	۹۷	۰/۹۷۰۵	۰/۳۹۸۰	۰/۹۸۶۰	۰/۱۱۰۷	۶-۸-۴-۱	۵۰
۹۸	۱۰۷	۰/۹۷۳۷	۰/۳۶۵۲	۰/۹۸۶۹	۰/۱۰۳۱	۶-۸-۵-۱	۵۱
۱۰۷	۱۱۷	۰/۹۶۸۱	۰/۴۷۳۳	۰/۹۹۲۱	۰/۰۶۲۳	۶-۸-۶-۱	۵۲
۱۱۵	۱۲۷	۰/۹۶۷۵	۰/۴۳۰۸	۰/۹۹۳۴	۰/۰۵۲۳	۶-۸-۷-۱	۵۳
۱۲۱	۱۳۷	۰/۹۶۵۳	۰/۵۴۲۱	۰/۹۹۴۸	۰/۰۴۱۰	۶-۸-۸-۱	۵۴
۱۳۵	۱۴۷	۰/۹۶۸۳	۰/۵۲۳۷	۰/۹۹۷۴	۰/۰۲۰۷	۶-۸-۹-۱	۵۵
۸۹	۹۷	۰/۹۷۳۰	۰/۳۵۶۵	۰/۹۸۶۴	۰/۱۰۷۱	۶-۹-۳-۱	۵۶
۹۳	۱۰۸	۰/۹۷۷۱	۰/۳۰۱۹	۰/۹۸۶۱	۰/۱۰۹۵	۶-۹-۴-۱	۵۷
۱۰۸	۱۱۹	۰/۹۷۳۰	۰/۳۵۵۴	۰/۹۹۰۴	۰/۰۷۵۷	۶-۹-۵-۱	۵۸
۸۶	۹۵	۰/۹۷۷۸	۰/۳۰۱۶	۰/۹۸۵۴	۰/۱۱۵۰	۶-۱۰-۲-۱	۵۹

همان‌طور که در جدول بالا مشاهده می‌شود شبکه مربوط به ردیف ۵۹ (شبکه ۲ لایه با ۱۰ نرون در لایه‌ی اول و ۲ نرون در لایه‌ی دوم) بهترین شبکه از لحاظ داشتن کمترین خطای تخمین برای داده‌های آزمون است. نتایج مربوط به ۷ شبکه با بهترین نتایج به ترتیب خطای آزمون در جدول ۴-۸ ارائه شده‌اند. از این نتایج برای مراحل بعدی و ایجاد ترکیبات استفاده شده است.

جدول ۴-۸: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش مرتب‌سازی بر اساس داشتن کمترین خطای آزمون.

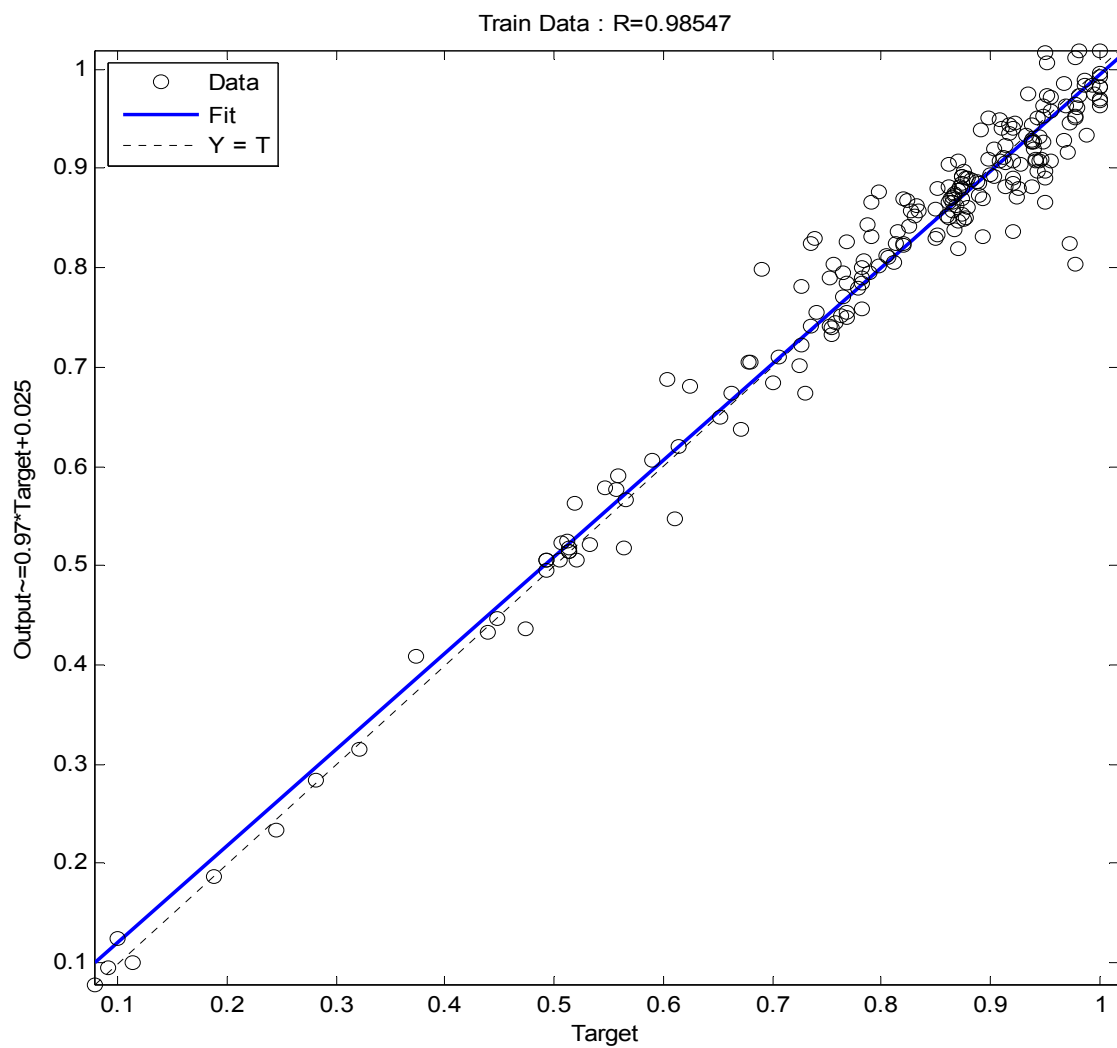
ردیف	ساختار شبکه	(MSE) داده‌های آموزش (%)	ضریب همبستگی داده‌های آموزش	(MSE) داده‌های آزمون (%)	ضریب همبستگی داده‌های آزمون	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۱	۶-۱۰-۲-۱	۰/۱۱۵۰	۰/۹۸۵۴	۰/۳۰۱۶	۰/۹۷۷۸	۹۵	۸۶
۲	۶-۹-۴-۱	۰/۱۰۹۵	۰/۹۸۶۱	۰/۳۰۱۹	۰/۹۷۷۱	۱۰۸	۹۳
۳	۶-۵-۱۰-۱	۰/۱۰۰۷	۰/۹۸۷۳	۰/۳۰۲۲	۰/۹۷۷۱	۱۰۶	۹۸
۴	۶-۶-۵-۱	۰/۱۴۲۶	۰/۹۸۱۹	۰/۳۱۱۹	۰/۹۷۸۰	۸۳	۷۶
۵	۶-۵-۶-۱	۰/۱۵۶۷	۰/۹۸۰۱	۰/۳۱۲۱	۰/۹۷۶۹	۷۸	۷۱
۶	۶-۴-۱۰-۱	۰/۱۴۰۶	۰/۹۸۲۲	۰/۳۱۳۸	۰/۹۷۷۴	۸۹	۸۳
۷	۶-۶-۶-۱	۰/۱۳۱۰	۰/۹۸۳۴	۰/۳۲۰۳	۰/۹۷۵۵	۹۱	۸۳

در شکل ۴-۹ نتایج مربوط به بهترین شبکه (شبکه ۱ جدول ۴-۸) ارائه شده‌اند. در شکل ۴-۹-الف، نمودار بالایی مربوط به مجموع مربعات خطا (SSE)، نمودار وسط مجموع مربعات پارامترهای شبکه (SSW) و نمودار پایینی بیانگر تعداد پارامترهای مؤثر شبکه می‌باشد.

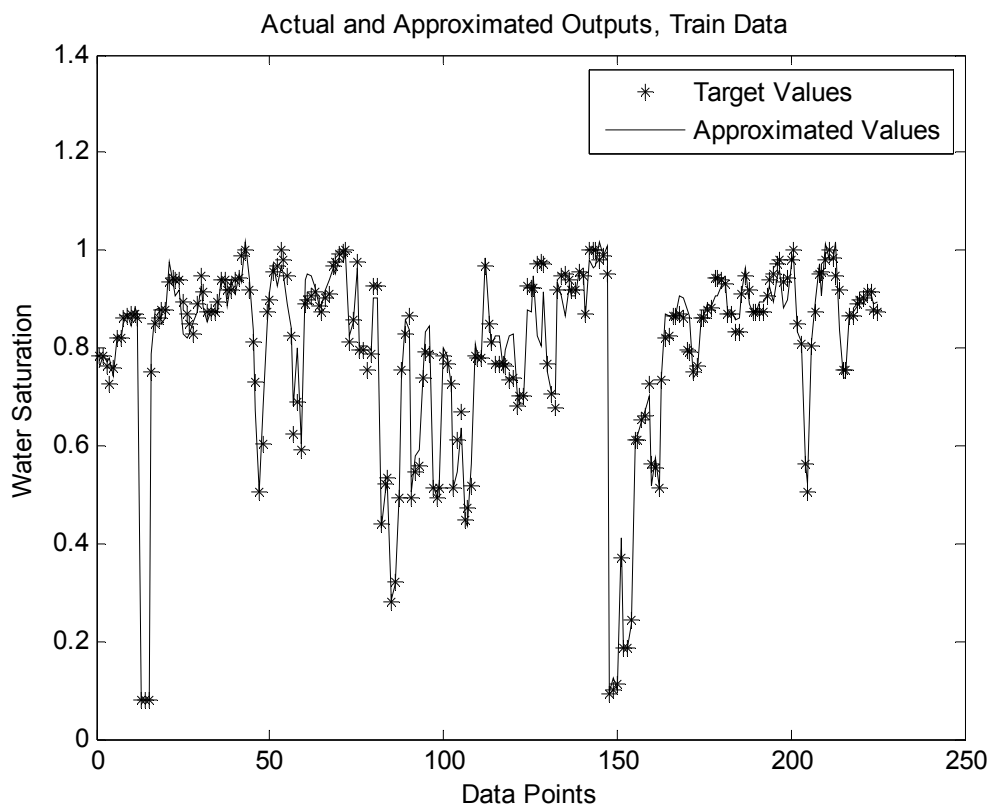


شکل ۴-۹-الف: نمایش روند آموزش و همگرایی SSE و SSW و تعداد پارامترهای مؤثر شبکه‌ی شماره‌ی ۱ جدول

۸-۴

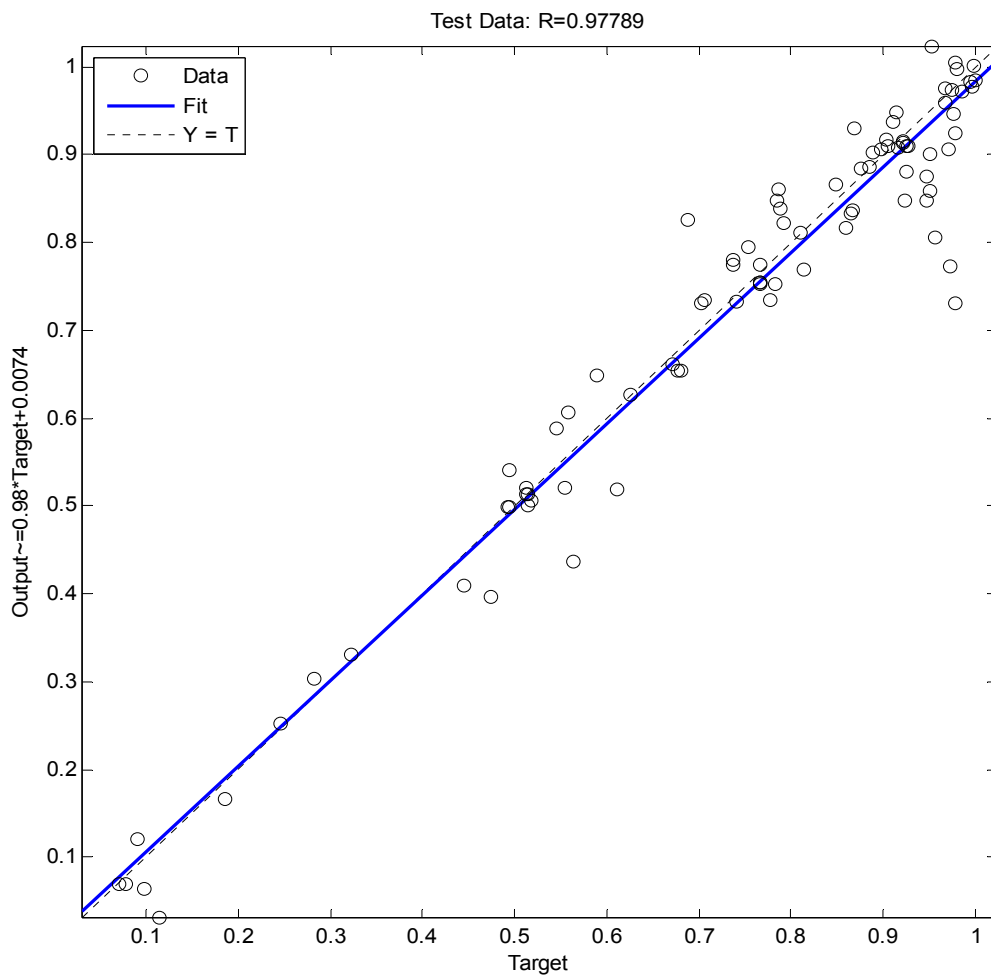


شکل ۴-۹-ب: نمودار همبستگی مقادیر اشباع از آب و تخمین شبکه‌ی ۱ جدول ۴-۸ برای داده‌های آموزش.

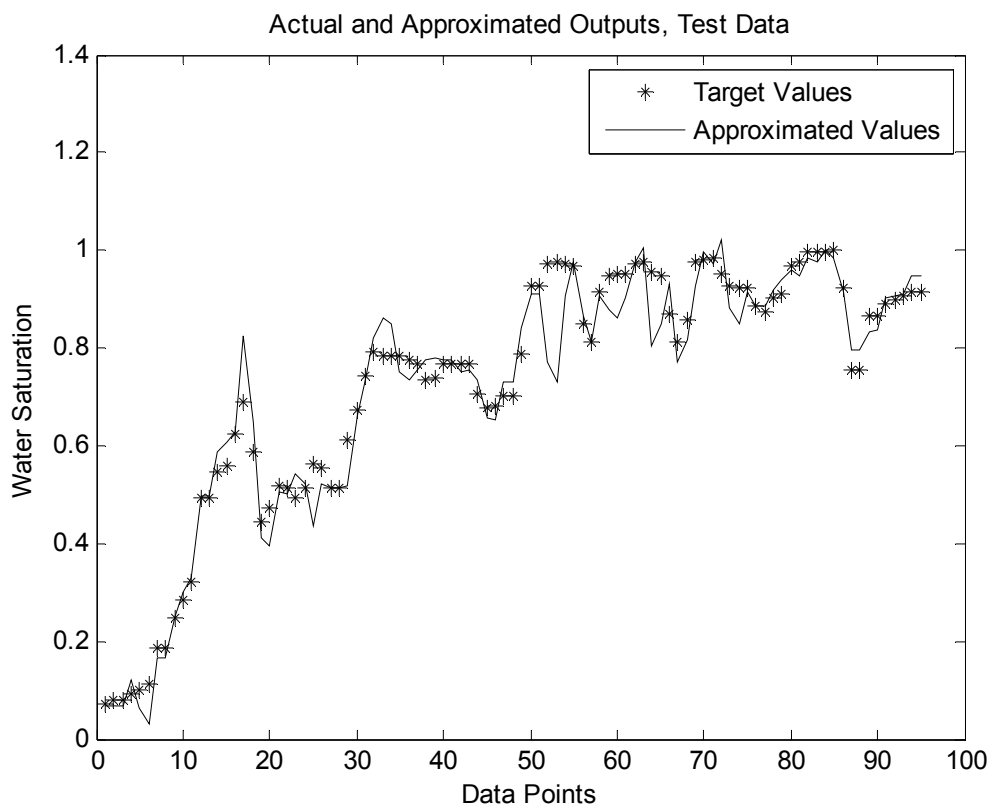


شکل ۴-۹ ج: مقادیر اشباع از آب و تخمین شبکه‌ی شماره ۱ جدول ۴-۸ برای داده‌های آموزش.





شکل ۴-۹-۵: نمودار همبستگی بین مقادیر اشباع از آب و تخمین شبکه‌ی ۱ جدول ۴-۸ برای داده‌های آزمون.



شکل ۴-۹-۵: مقادیر اشباع از آب و تخمین شبکه شماره ۱ جدول ۴-۸ برای داده‌های آزمون.

#### ۴-۸-۱- ترکیبات آنسامبلی خطی

۱۲۰ ترکیب خطی آنسامبلی شبکه‌های منفرد جدول (۳-۴) به دو روش الگوریتم ژنتیک و میانگین گیری ساده ایجاد شد. نتایج این ۱۲۰ ترکیب در جدول ۴-۹ ارائه شده است. ستون اول از سمت چپ (NO)، شماره ردیف و ستون دوم (Ensemble)، نوع ترکیب آنسامبلی را نشان می‌دهد. ستون‌های دیگر میانگین مربعات خطای تخمین را برای داده‌های آموزش و آزمون برای هر دو روش نشان می‌دهد. این خطاها تا ۴ رقم اعشار گرد شده‌اند.

پارامترهای الگوریتم ژنتیک مانند بخش مربوط به فرا آموزش در نظر گرفته شدند.

جدول ۴-۹: نتایج ترکیبات آنسامبلی خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده.

NO.	Ensemble	optimal GA-train(%)	optimal GA-test(%)	Simple Avg-train(%)	Simple Avg-test(%)
1	1,2	0.0942	0.2515	0.0943	0.2512
2	1,3	0.0847	0.2401	0.0852	0.2384
3	1,4	0.1019	0.2173	0.1038	0.2115
4	1,5	0.1070	0.2489	0.1114	0.2388
5	1,6	0.0946	0.2150	0.0959	0.2124
6	1,7	0.1012	0.2463	0.1019	0.2458
7	2,3	0.0814	0.2271	0.0816	0.2264
8	2,4	0.0995	0.2368	0.1024	0.2289
9	2,5	0.1026	0.2547	0.1081	0.2428
10	2,6	0.0887	0.1890	0.0904	0.1841
11	2,7	0.1039	0.2815	0.1059	0.2806
12	3,4	0.08647	0.2067	0.0899	0.1962
13	3,5	0.0915	0.2253	0.0979	0.2067
14	3,6	0.0887	0.2352	0.0922	0.2273
15	3,7	0.0893	0.2437	0.0917	0.2394
16	4,5	0.1204	0.2540	0.1209	0.2532
17	4,6	0.1007	0.1961	0.1007	0.1960
18	4,7	0.1044	0.2151	0.1047	0.2139
19	5,6	0.1044	0.2149	0.1047	0.2140
20	5,7	0.1128	0.2364	0.1142	0.2316
21	6,7	0.1033	0.2117	0.1034	0.2108
22	1,2,3	0.0786	0.2175	0.0799	0.2176
23	1,2,4	0.0912	0.2173	0.0928	0.2057
24	1,2,5	0.0931	0.2369	0.0971	0.2240
25	1,2,6	0.0837	0.1871	0.0841	0.1859
26	1,2,7	0.0931	0.2439	0.0948	0.2430
27	1,3,4	0.0811	0.1990	0.0842	0.1854
28	1,3,5	0.0837	0.2213	0.0896	0.2022
29	1,3,6	0.0802	0.2079	0.0819	0.1994
30	1,3,7	0.0827	0.2243	0.0854	0.2189
31	1,4,5	0.0998	0.2101	0.1033	0.2098
32	1,4,6	0.0886	0.1762	0.0892	0.1725
33	1,4,7	0.0943	0.1984	0.0948	0.1946
34	2,3,4	0.0791	0.1991	0.0826	0.1878
35	2,3,5	0.0809	0.2149	0.0871	0.1986
36	2,3,6	0.0770	0.1902	0.0785	0.1815
37	2,3,7	0.0812	0.2253	0.0862	0.2290
38	2,4,5	0.0974	0.2256	0.1019	0.2193
39	2,4,6	0.0852	0.1702	0.0868	0.1676
40	2,4,7	0.0960	0.2236	0.0965	0.2177
41	2,5,6	0.0865	0.1805	0.0896	0.1817

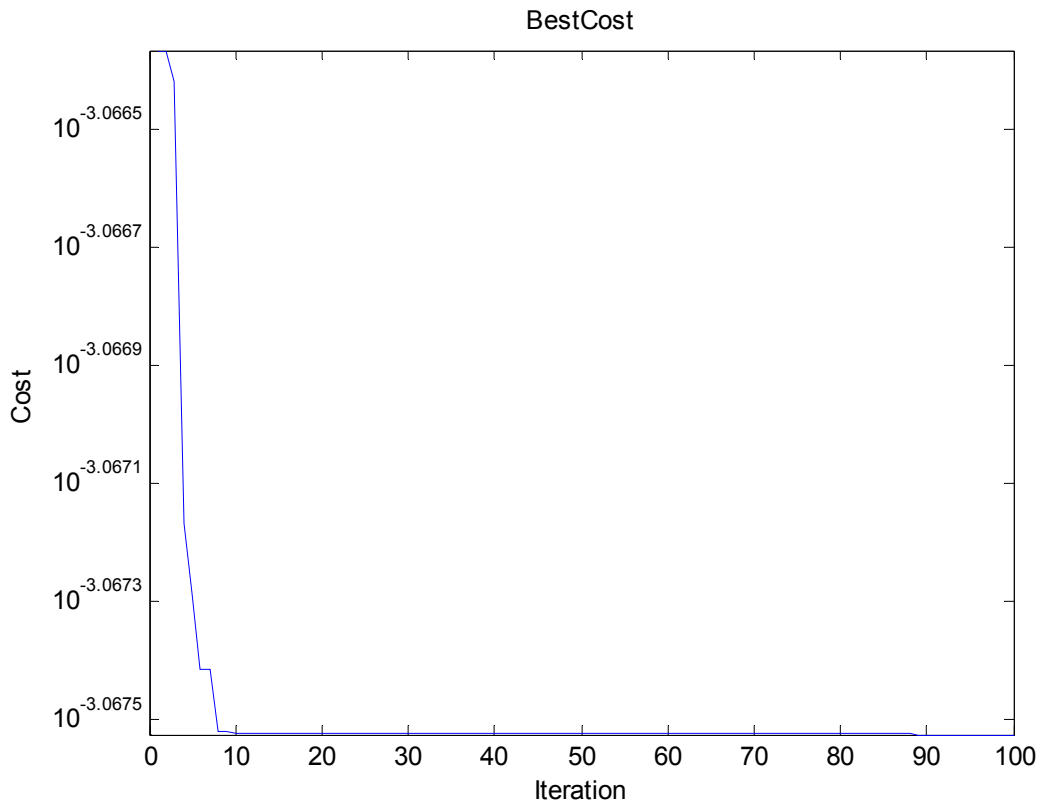
42	2,5,7	0.0995	0.2447	0.1017	0.2317
43	3,4,5	0.0851	0.1955	0.0928	0.1887
44	3,4,6	0.0808	0.1819	0.0831	0.1722
45	3,4,7	0.0827	0.1936	0.0857	0.1848
46	3,5,6	0.0839	0.1964	0.0869	0.1848
47	3,5,7	0.0866	0.2128	0.0919	0.1973
48	4,5,6	0.0946	0.1870	0.0961	0.1905
49	4,5,7	0.1012	0.2044	0.1032	0.2056
50	1,5,6	0.0911	0.1955	0.0929	0.1926
51	1,5,7	0.0985	0.2234	0.1008	0.2145
52	1,6,7	0.0902	0.1958	0.0909	0.1934
53	2,6,7	0.0881	0.1895	0.0909	0.1962
54	3,6,7	0.0838	0.2087	0.0864	0.1971
55	4,6,7	0.0912	0.1707	0.0912	0.1708
56	5,6,7	0.0951	0.1864	0.0957	0.1866
57	4,5,6,7	0.0894	0.1686	0.0908	0.1726
58	3,5,6,7	0.0818	0.1889	0.0849	0.1764
59	3,4,6,7	0.0791	0.1747	0.0813	0.1649
60	3,4,5,7	0.0823	0.1887	0.0884	0.1794
61	3,4,5,6	0.0800	0.1757	0.0840	0.1683
62	2,5,6,7	0.0863	0.1805	0.0895	0.1849
63	2,4,6,7	0.0848	0.1699	0.0864	0.1726
64	2,4,5,7	0.0947	0.2156	0.0966	0.2070
65	2,4,5,6	0.0845	0.1691	0.0881	0.1748
66	2,3,6,7	0.0770	0.1903	0.0811	0.1874
67	2,3,5,7	0.0808	0.2140	0.0876	0.2023
68	2,3,5,6	0.0767	0.1842	0.0803	0.1715
69	1,5,6,7	0.0884	0.1845	0.0900	0.1824
70	1,4,6,7	0.0861	0.1679	0.0864	0.1667
71	1,4,5,7	0.0937	0.1950	0.0960	0.1929
72	1,4,5,6	0.0876	0.1742	0.0900	0.1765
73	1,3,6,7	0.0793	0.1996	0.0817	0.1888
74	1,3,5,7	0.0822	0.2132	0.0877	0.1956
75	1,3,5,6	0.0795	0.1953	0.0827	0.1807
76	1,3,4,7	0.0799	0.1922	0.0831	0.1818
77	1,3,4,6	0.0777	0.1792	0.0796	0.1668
78	1,2,6,7	0.0837	0.1868	0.0860	0.1915
79	1,2,5,7	0.0922	0.2313	0.0949	0.2182
80	1,2,5,6	0.0832	0.1820	0.0860	0.1821
81	1,2,4,7	0.0901	0.2104	0.0910	0.2035
82	1,2,4,6	0.0823	0.1712	0.0834	0.1674
83	1,2,3,7	0.0786	0.2172	0.0832	0.2172
84	1,2,3,6	0.0755	0.1886	0.0767	0.1825
85	2,3,4,7	0.0790	0.1976	0.0836	0.1918
<b>86</b>	<b>2,3,4,6</b>	<b>0.0756</b>	<b>0.1741</b>	<b>0.0777</b>	<b>0.1610</b>
87	2,3,4,5	0.0791	0.1967	0.0866	0.1850

88	1,3,4,5	0.0810	0.1958	0.0879	0.1827
89	1,2,4,5	0.0909	0.2137	0.0948	0.2032
90	1,2,3,5	0.0786	0.2141	0.0844	0.1989
91	1,2,3,4	0.0775	0.1975	0.0809	0.1860
92	3,4,5,6,7	0.0788	0.1715	0.0827	0.1630
93	2,4,5,6,7	0.0842	0.1686	0.0872	0.1737
94	2,3,5,6,7	0.0767	0.1841	0.0818	0.1761
95	2,3,4,6,7	0.0756	0.1738	0.0792	0.1666
96	2,3,4,5,7	0.0790	0.1957	0.0859	0.1853
97	2,3,4,5,6	0.0755	0.1730	0.0802	0.1630
<b>98</b>	<b>1,4,5,6,7</b>	<b>0.0856</b>	<b>0.1668</b>	<b>0.0875</b>	<b>0.1693</b>
99	1,3,5,6,7	0.0789	0.1910	0.0825	0.1764
100	1,3,4,6,7	0.0772	0.1754	0.0795	0.1647
101	1,3,4,5,7	0.0799	0.1909	0.0859	0.1783
102	1,3,4,5,6	0.0776	0.1769	0.0818	0.1661
103	1,2,5,6,7	0.0832	0.1818	0.0865	0.1840
104	1,2,4,6,7	0.0822	0.1707	0.0839	0.1717
105	1,2,4,5,7	0.0899	0.2078	0.0922	0.1980
106	1,2,4,5,6	0.0821	0.1703	0.0855	0.1723
107	1,2,3,6,7	0.0755	0.1888	0.0792	0.1858
108	1,2,3,5,7	0.0786	0.2139	0.0852	0.1997
109	1,2,3,5,6	0.0755	0.1866	0.0792	0.1749
110	1,2,3,4,7	0.0775	0.1970	0.0820	0.1886
111	1,2,3,4,6	0.0747	0.1754	0.0768	0.1638
112	1,2,3,4,5	0.0775	0.1976	0.0844	0.1835
113	1,2,3,4,5,6	0.0747	0.1753	0.0794	0.1649
114	1,2,3,4,5,7	0.0775	0.1965	0.0843	0.1839
115	1,2,3,4,6,7	0.0747	0.1750	0.0783	0.1679
116	1,2,3,5,6,7	0.0755	0.1865	0.0806	0.1776
117	1,2,4,5,6,7	0.0821	0.1698	0.0853	0.1727
118	1,3,4,5,6,7	0.0771	0.1742	0.0813	0.1638
119	2,3,4,5,6,7	0.0755	0.1730	0.0809	0.1655
120	1,2,3,4,5,6,7	0.0747	0.1754	0.0801	0.1670

با توجه به جدول، ترکیب مربوط به ردیف ۹۸ (ترکیب شبکه های ۱، ۴، ۵، ۶ و ۷) برای الگوریتم ژنتیک و ترکیب مربوط به ردیف ۸۶ (ترکیب شبکه های ۲، ۳، ۴ و ۶) برای میانگین گیری ساده، بهترین جواب را از جهت کمترین خطای آزمون به همراه دارند. نتایج کامل تر در جدول ۴-۱۰ ارائه شده‌اند. در شکل ۴-۱۰ نیز روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک ارائه شده است.

جدول ۴-۱۰: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک.

روش ترکیب خطی		نتایج
الگوریتم ژنتیک	میانگین گیری ساده	
۱، ۴، ۵، ۶ و ۷	۲، ۳، ۴ و ۶	ترکیب آنسامبلی
۰/۹۸۹۳	۰/۹۹۰۳	ضریب همبستگی داده های آموزش
۰/۰۸۵۶	۰/۰۷۷۷	MSE داده های آموزش(%)
%۲۵/۵۷	%۳۲/۴۳	کاهش در MSE داده های آموزش
۰/۹۸۷۴	۰/۹۸۷۷	ضریب همبستگی داده های آزمون
۰/۱۶۶۸	۰/۱۶۱۰	MSE داده های آزمون(%)
%۴۴/۶۹	%۴۶/۶۲	کاهش در MSE داده های آزمون
۰/۲۷۵۸	۰/۲۵	$W_1$
۰/۱۹۱۵	۰/۲۵	$W_2$
۰/۰۷۷۴	۰/۲۵	$W_3$
۰/۲۷۹۸	۰/۲۵	$W_4$
۰/۱۷۵۵	۰	$W_5$
۰	۰	$W_6$
۰	۰	$W_7$
۱	۱	$\sum W$



شکل ۴-۱۰: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک.

ساختار جدول ۴-۱۰ نیز مانند جدول ۴-۵ می‌باشد. میزان کاهش میانگین مربعات خطای داده‌های آموزش از طریق تقسیم حاصل تفریق میانگین مربعات خطای ترکیب و بهترین شبکه منفرد بر میانگین مربعات خطای بهترین شبکه منفرد محاسبه شده است. برای روش الگوریتم ژنتیک شبکه‌ی شماره ۶ جدول ۴-۸ بیشترین ضریب و شبکه‌ی شماره ۵ کمترین ضریب را به خود اختصاص داده‌اند. میزان کاهش خطای تخمین بهترین ترکیب برای داده‌های آزمون در روش میانگین گیری بیشتر از روش الگوریتم ژنتیک می‌باشد.

#### ۴-۸-۲- واریانس و بایاس خطا

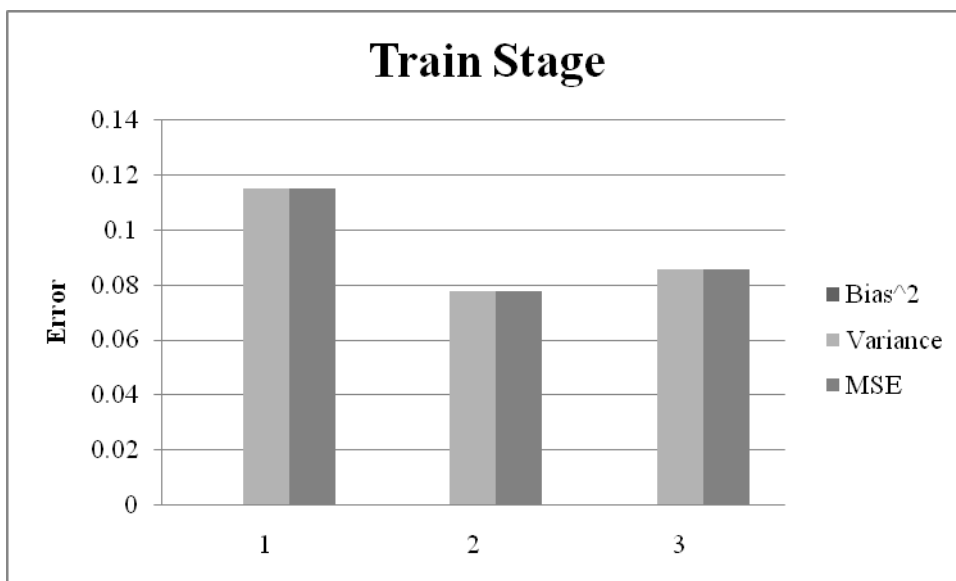
در جدول ۴-۱۱ واریانس، بایاس و میانگین مربعات خطای تخمین برای داده‌های آموزش و آزمون بهترین شبکه‌ی منفرد (۱) (شبکه ۱ جدول ۴-۸)، بهترین ترکیب با هر دو روش میانگین گیری ساده (۲) و الگوریتم ژنتیک (۳) ارائه شده است.

شکل ۴-۱۱ نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین برای داده‌های آموزش و آزمون حالات مربوط به جدول ۴-۱۱ را نشان می‌دهد. با توجه به شکل ۴-۱۱-الف برای داده‌های آموزش می‌بینیم که واریانس مؤلفه‌ی اصلی خطا است و بایاس تقریباً نقشی در خطا ندارد. برای داده‌های آزمون اگر چه نقش بایاس نسبت به داده‌های آموزش بیشتر است ولی باز هم می‌توان گفت که تقریباً نقشی در خطا ندارد و واریانس مؤلفه‌ی اصلی خطاست.

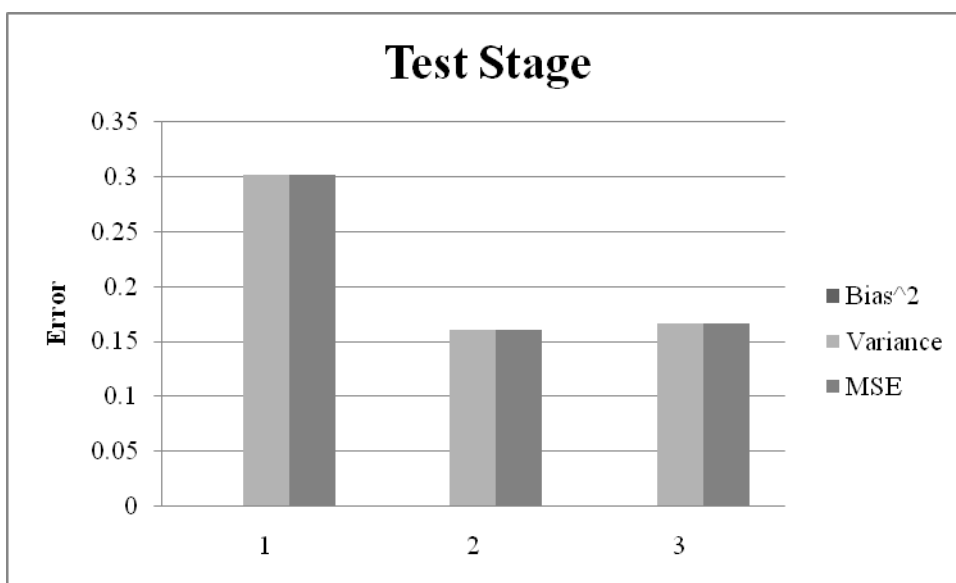
جدول ۴-۱۱: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه‌ی ۱ جدول ۴-۸ و بهترین ترکیب آن جدول.

ردیف	بایاس تخمین داده های آموزش	واریانس تخمین داده- های آموزش	میانگین مربعات خطای تخمین داده‌های آموزش	بایاس تخمین داده های آزمون	واریانس تخمین داده های آزمون	میانگین مربعات خطای تخمین داده های آزمون
۱	$-۲/۰۲ \times ۱۰^{-۶}$	۰/۱۱۵۰	۰/۱۱۵۰	-۰/۰۰۹۸	۰/۳۰۱۵	۰/۳۰۱۶
۲	$-۲/۰۹ \times ۱۰^{-۶}$	۰/۰۷۷۷	۰/۰۷۷۷	۰/۰۰۱۷	۰/۱۶۱۰	۰/۱۶۱۰
۳	$-۳/۵۲ \times ۱۰^{-۶}$	۰/۰۸۵۶	۰/۰۸۵۶	$-۷/۴۶ \times ۱۰^{-۴}$	۰/۱۶۶۸	۰/۱۶۶۸





شکل ۴-۱۱-الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).



شکل ۴-۱۱-ب: نمودار میله ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده های آزمون (۱: شبکه ۱ جدول ۴-۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).

#### ۹-۴- ایجاد شبکه‌های پس انتشار خطا به روش توقف سریع و ترکیب نتایج

برای آموزش به روش توقف سریع از تابع لونبرگ - مارکوت (Trainlm) استفاده شد. برای این تابع در روش توقف سریع پارامترهای تابع آموزش باید به گونه ای تعیین شوند که همگرایی سریع رخ ندهد. بدین منظور پارامترهای تابع آموزش به صورت ( $\mu\_inc=1.2$ ,  $\mu\_max=0.95$ ,  $\mu=1$ ) تعیین شدند. داده‌های آموزش به ۲ دسته تقسیم شدند: ۷۰٪ برای آموزش و ۳۰٪ برای اعتبار سنجی حین آموزش انتخاب شدند. این شبکه‌ها با یک و دو لایه‌ی میانی و تعداد نرون‌های مختلف، با شرایط اولیه گوناگون طی یک مرحله‌ی طولانی آزمون و خطا مورد آزمون قرار گرفتند و برای هر ساختار خاص بهترین نتیجه ثبت شد. ساختارهای شبکه‌ها در اینجا مانند ساختارهای مورد استفاده در روش فرا آموزش است. نتایج مربوط به شبکه‌های منفرد این بخش در جدول ۴-۱۲ ارائه شده است. ستون‌های این جدول مانند جدول ۴-۲ می‌باشند که قبلاً توضیح داده شده‌اند. تنها ستون اضافه شده مربوط به تعداد دوره‌ی توقف آموزش<sup>۱</sup> شبکه می‌باشد.

جدول ۴-۱۲: نتایج مربوط به شبکه‌های ساخته شده با یک و دو لایه‌ی میانی به روش توقف سریع.

Epochs = 400, goal = 0, min\_grad = 1e-10, mem\_reduc = 1, mu = 1, mu\_dec = 0.95,  
mu\_inc = 1.2, mu\_max = 1e10

ردیف	ساختار شبکه	(MSE) داده های آموزش (%)	ضریب همبستگی داده های آموزش	(MSE) داده های آزمون (%)	ضریب همبستگی داده های آزمون	تعداد کل پارامترهای شبکه	دوره‌ی توقف آموزش
۱	۶-۶-۱	۰/۴۳۴۵	۰/۹۴۴۰	۰/۸۱۴۳	۰/۹۴۱۵	۴۹	۵۵
۲	۶-۷-۱	۰/۳۶۴۶	۰/۹۵۳۱	۰/۶۸۲۶	۰/۹۵۰۵	۵۷	۸۸
۳	۶-۸-۱	۰/۳۰۶۸	۰/۹۶۰۸	۰/۶۱۹۵	۰/۹۵۵۴	۶۵	۴۳
۴	۶-۹-۱	۰/۲۸۹۴	۰/۹۶۳۰	۰/۵۶۶۴	۰/۹۵۶۶	۷۳	۴۰
۵	۶-۱۰-۱	۰/۲۴۸۰	۰/۹۶۸۴	۰/۵۰۶۱	۰/۹۶۱۵	۸۱	۵۷

1. Stop epoch

۷۲	۸۹	•/۹۵۹۴	•/۵۳۷۰	•/۹۶۶۶	•/۲۶۱۸	۶-۱۱-۱	۶
۵۱	۹۷	•/۹۶۵۶	•/۴۹۲۷	•/۹۷۱۶	•/۲۲۴۰	۶-۱۲-۱	۷
۳۳	۱۰۵	•/۹۶۲۰	•/۴۹۳۲	•/۹۷۱۰	•/۲۲۷۶	۶-۱۳-۱	۸
۴۳	۱۱۳	•/۹۶۷۰	•/۴۴۴۵	•/۹۷۶۶	•/۱۸۴۶	۶-۱۴-۱	۹
۴۶	۱۲۱	•/۹۶۸۴	•/۴۲۲۲	•/۹۷۴۶	•/۲۰۱۰	۶-۱۵-۱	۱۰
۵۵	۴۷	•/۹۲۹۷	•/۹۲۹۰	•/۹۳۵۵	•/۴۹۸۲	۶-۴-۳-۱	۱۱
۶۴	۵۳	•/۹۲۶۲	۱/۰۳۹۴	•/۹۳۴۳	•/۵۰۹۸	۶-۴-۴-۱	۱۲
۴۲	۵۹	•/۹۲۸۷	•/۹۷۰۸	•/۹۳۹۲	•/۴۷۱۶	۶-۴-۵-۱	۱۳
۲۶	۶۵	•/۹۳۱۰	•/۹۱۸۰	•/۹۳۸۷	•/۴۷۸۱	۶-۴-۶-۱	۱۴
۳۱	۷۱	•/۹۳۹۷	•/۸۵۹۳	•/۹۴۷۷	•/۴۱۰۹	۶-۴-۷-۱	۱۵
۲۹	۷۷	•/۹۴۹۰	•/۶۸۴۴	•/۹۵۸۹	•/۳۲۱۰	۶-۴-۸-۱	۱۶
۴۲	۸۳	•/۹۴۴۱	•/۶۴۸۱	•/۹۵۴۴	•/۴۳۷۵	۶-۴-۹-۱	۱۷
۳۸	۸۹	•/۹۵۲۱	•/۶۹۸۴	•/۹۵۸۴	•/۳۲۴۷	۶-۴-۱۰-۱	۱۸
۲۳	۹۵	•/۹۵۰۶	•/۷۶۰۶	•/۹۵۷۴	•/۳۳۲۶	۶-۴-۱۱-۱	۱۹
۳۶	۱۰۱	•/۹۵۰۵	•/۶۴۲۱	•/۹۵۰۶	•/۳۸۵۳	۶-۴-۱۲-۱	۲۰
۳۰	۱۰۷	•/۹۵۷۵	•/۶۵۴۱	•/۹۵۸۶	•/۳۲۵۰	۶-۴-۱۳-۱	۲۱
۵۶	۱۱۳	•/۹۵۰۸	•/۶۴۵۲	•/۹۵۸۵	•/۳۲۳۹	۶-۴-۱۴-۱	۲۲
۶۳	۵۷	•/۹۴۷۶	•/۷۴۴۲	•/۹۵۰۳	•/۳۸۷۶	۶-۵-۳-۱	۲۳
۳۷	۶۴	•/۹۶۲۰	•/۵۱۷۶	•/۹۶۵۱	•/۲۷۳۷	۶-۵-۴-۱	۲۴
۶۶	۷۱	•/۹۵۵۰	•/۶۴۶۴	•/۹۶۱۳	•/۳۰۳۴	۶-۵-۵-۱	۲۵
۵۲	۷۸	•/۹۶۱۶	•/۵۱۳۶	•/۹۶۱۹	•/۳۰۳۰	۶-۵-۶-۱	۲۶
۶۱	۸۵	•/۹۵۸۲	•/۵۵۱۲	•/۹۶۶۶	•/۲۶۲۷	۶-۵-۷-۱	۲۷
۴۹	۹۲	•/۹۵۹۶	•/۵۵۳۷	•/۹۶۵۷	•/۲۶۸۶	۶-۵-۸-۱	۲۸
۳۲	۹۹	•/۹۶۱۱	•/۵۴۱۶	•/۹۷۱۴	•/۲۲۵۴	۶-۵-۹-۱	۲۹
۴۷	۱۰۶	•/۹۶۵۹	•/۴۶۳۳	•/۹۷۰۴	•/۲۳۲۳	۶-۵-۱۰-۱	۳۰
۲۵	۱۱۳	•/۹۶۰۷	•/۵۲۱۴	•/۹۷۱۶	•/۲۲۴۵	۶-۵-۱۱-۱	۳۱
۵۲	۶۷	•/۹۵۰۲	•/۶۷۸۴	•/۹۵۲۸	•/۳۶۹۷	۶-۶-۳-۱	۳۲
۴۸	۷۵	•/۹۵۴۴	•/۵۸۹۸	•/۹۶۲۹	•/۲۹۱۸	۶-۶-۴-۱	۳۳
۳۵	۸۳	•/۹۵۶۵	•/۵۶۶۷	•/۹۶۲۹	•/۲۸۹۸	۶-۶-۵-۱	۳۴
۶۲	۹۱	•/۹۶۶۹	•/۴۳۷۷	•/۹۷۱۵	•/۲۲۳۴	۶-۶-۶-۱	۳۵
۴۸	۹۹	•/۹۶۳۱	•/۴۹۱۳	•/۹۷۷۰	•/۱۸۱۰	۶-۶-۷-۱	۳۶
۴۶	۱۰۷	•/۹۶۳۰	•/۴۸۴۱	•/۹۶۵۱	•/۲۷۲۸	۶-۶-۸-۱	۳۷

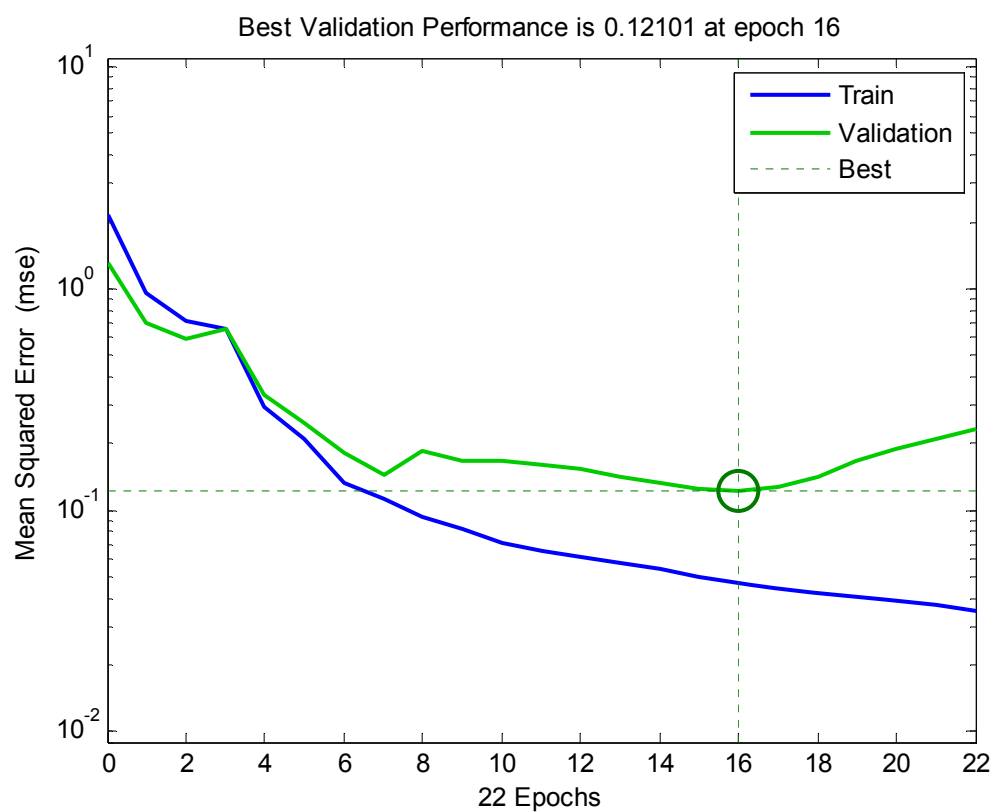
۳۵	۱۱۵	۰/۹۶۸۲	۰/۴۲۷۸	۰/۹۷۲۳	۰/۲۱۷۸	۶-۶-۹-۱	۳۸
۵۸	۱۲۳	۰/۹۶۰۸	۰/۵۴۷۶	۰/۹۷۲۳	۰/۲۱۷۸	۶-۶-۱۰-۱	۳۹
۳۴	۷۷	۰/۹۵۱۸	۰/۶۶۰۱	۰/۹۶۱۸	۰/۲۹۸۹	۶-۷-۳-۱	۴۰
۸۰	۸۶	۰/۹۶۳۷	۰/۴۷۶۵	۰/۹۶۶۷	۰/۲۶۱۱	۶-۷-۴-۱	۴۱
۳۹	۹۵	۰/۹۵۸۹	۰/۵۵۰۹	۰/۹۶۶۰	۰/۲۶۶۸	۶-۷-۵-۱	۴۲
۵۱	۱۰۴	۰/۹۵۷۸	۰/۵۵۸۵	۰/۹۶۰۳	۰/۳۱۰۴	۶-۷-۶-۱	۴۳
۴۷	۱۱۳	۰/۹۶۳۱	۰/۴۹۳۲	۰/۹۷۲۲	۰/۲۱۸۳	۶-۷-۷-۱	۴۴
۳۱	۱۲۲	۰/۹۶۵۶	۰/۴۷۹۲	۰/۹۷۶۷	۰/۱۸۳۱	۶-۷-۸-۱	۴۵
۴۲	۱۳۱	۰/۹۶۶۸	۰/۴۴۹۴	۰/۹۷۶۶	۰/۱۸۴۶	۶-۷-۹-۱	۴۶
۳۱	۱۴۰	۰/۹۶۷۰	۰/۴۳۲۸	۰/۹۷۴۹	۰/۱۹۷۰	۶-۷-۱۰-۱	۴۷
۴۹	۱۴۹	۰/۹۷۰۰	۰/۴۲۴۶	۰/۹۷۸۳	۰/۱۷۱۸	۶-۷-۱۱-۱	۴۸
۵۵	۸۷	۰/۹۴۸۳	۰/۶۶۵۷	۰/۹۵۶۲	۰/۳۴۱۲	۶-۸-۳-۱	۴۹
۶۰	۹۷	۰/۹۵۷۲	۰/۵۸۸۴	۰/۹۶۵۴	۰/۲۷۰۹	۶-۸-۴-۱	۵۰
۲۸	۱۰۷	۰/۹۵۵۸	۰/۵۷۰۴	۰/۹۶۱۱	۰/۳۰۳۷	۶-۸-۵-۱	۵۱
۵۵	۱۱۷	۰/۹۵۷۵	۰/۵۷۱۱	۰/۹۶۰۶	۰/۳۰۸۹	۶-۸-۶-۱	۵۲
۲۴	۱۲۷	۰/۹۵۶۷	۰/۵۹۰۹	۰/۹۶۹۴	۰/۲۴۰۵	۶-۸-۷-۱	۵۳
۲۲	۱۳۷	۰/۹۷۰۲	۰/۴۰۴۹	۰/۹۷۲۵	۰/۲۱۷۱	۶-۸-۸-۱	۵۴
۷۹	۱۴۷	۰/۹۷۰۰	۰/۴۰۵۷	۰/۹۷۳۷	۰/۲۰۶۵	۶-۸-۹-۱	۵۵
۴۵	۹۷	۰/۹۶۳۲	۰/۴۷۸۸	۰/۹۶۴۷	۰/۲۷۶۱	۶-۹-۳-۱	۵۶
۸۴	۱۰۸	۰/۹۶۶۶	۰/۴۳۴۱	۰/۹۶۹۲	۰/۲۴۱۶	۶-۹-۴-۱	۵۷
۴۸	۱۱۹	۰/۹۶۸۵	۰/۴۳۴۷	۰/۹۷۶۶	۰/۱۸۴۴	۶-۹-۵-۱	۵۸
۴۴	۹۵	۰/۹۵۵۷	۰/۵۸۴۰	۰/۹۶۵۱	۰/۲۷۴۱	۶-۱۰-۲-۱	۵۹

همان‌طور که در جدول بالا مشاهده می‌شود شبکه مربوط به ردیف ۵۴ (شبکه ۲ لایه با ۸ نرون در لایه‌ی اول و ۸ نرون در لایه‌ی دوم) بهترین شبکه از لحاظ داشتن کمترین خطای تخمین برای داده‌های آزمون است. نتایج مربوط به ۷ شبکه با بهترین نتایج به ترتیب خطای آزمون در جدول ۴-۱۳ ارائه شده‌اند. از این نتایج برای مراحل بعدی و ایجاد ترکیبات استفاده شده است.

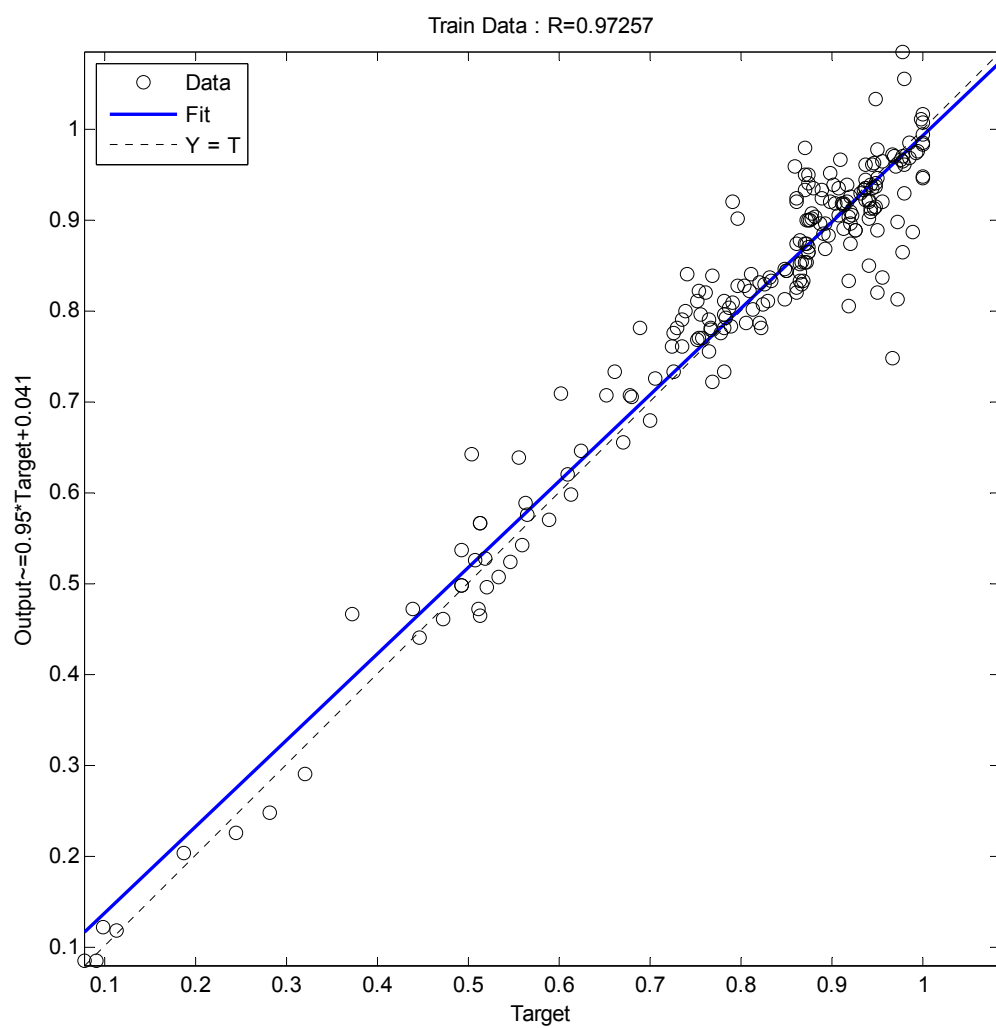
در شکل ۴-۱۲ نتایج مربوط به بهترین شبکه آموزش دیده به روش توقف سریع (شبکه ۱ جدول ۴-۱۳) ارائه شده است.

جدول ۴-۱۳: نتایج مربوط به ۷ شبکه‌ی برتر آموزش دیده به روش توقف سریع بر اساس داشتن کمترین خطای آزمون.

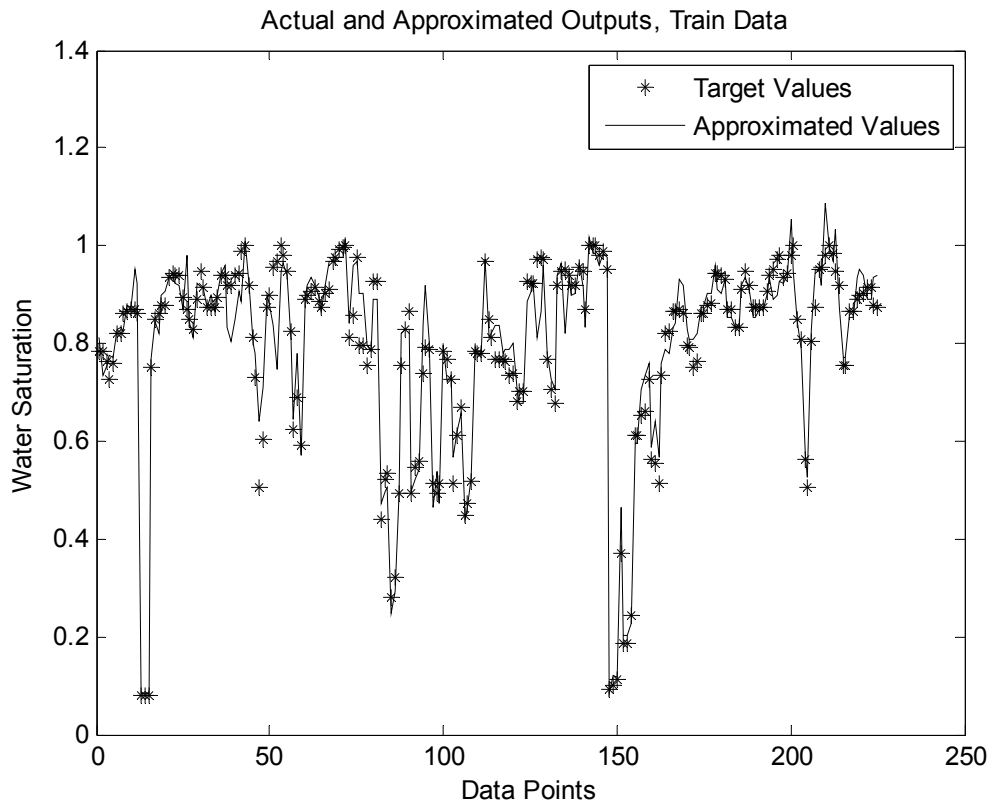
ردیف	ساختار شبکه	(MSE) داده‌های آموزش (%)	ضریب همبستگی داده‌های آموزش	(MSE) داده‌های آزمون (%)	ضریب همبستگی داده‌های آزمون	تعداد کل پارامترهای شبکه	دوره‌ی توقف آموزش
۱	۶-۸-۸-۱	۰/۲۱۷۱	۰/۹۷۲۵	۰/۴۰۴۹	۰/۹۷۰۲	۱۳۷	۲۲
۲	۶-۸-۹-۱	۰/۲۰۶۵	۰/۹۷۳۷	۰/۴۰۵۷	۰/۹۷۰۰	۱۴۷	۷۹
۳	۶-۱۵-۱	۰/۲۰۱۰	۰/۹۷۴۶	۰/۴۲۲۲	۰/۹۶۸۴	۱۲۱	۴۶
۴	۶-۷-۱۱-۱	۰/۱۷۱۸	۰/۹۷۸۳	۰/۴۲۴۶	۰/۹۷۰۰	۱۴۹	۴۹
۵	۶-۶-۹-۱	۰/۲۱۷۸	۰/۹۷۲۳	۰/۴۲۷۸	۰/۹۶۸۲	۱۱۵	۳۵
۶	۶-۷-۱۰-۱	۰/۱۹۷۰	۰/۹۷۴۹	۰/۴۳۲۸	۰/۹۶۷۰	۱۴۰	۳۱
۷	۶-۹-۴-۱	۰/۲۴۱۶	۰/۹۶۹۲	۰/۴۳۴۱	۰/۹۶۶۶	۱۰۸	۸۴



شکل ۴-۱۲-الف: نمایش روند آموزش و کاهش میانگین مربعات خطا برای داده های آموزش و اعتبار سنجی در شبکه شماره ۱ جدول ۴-۱۳.

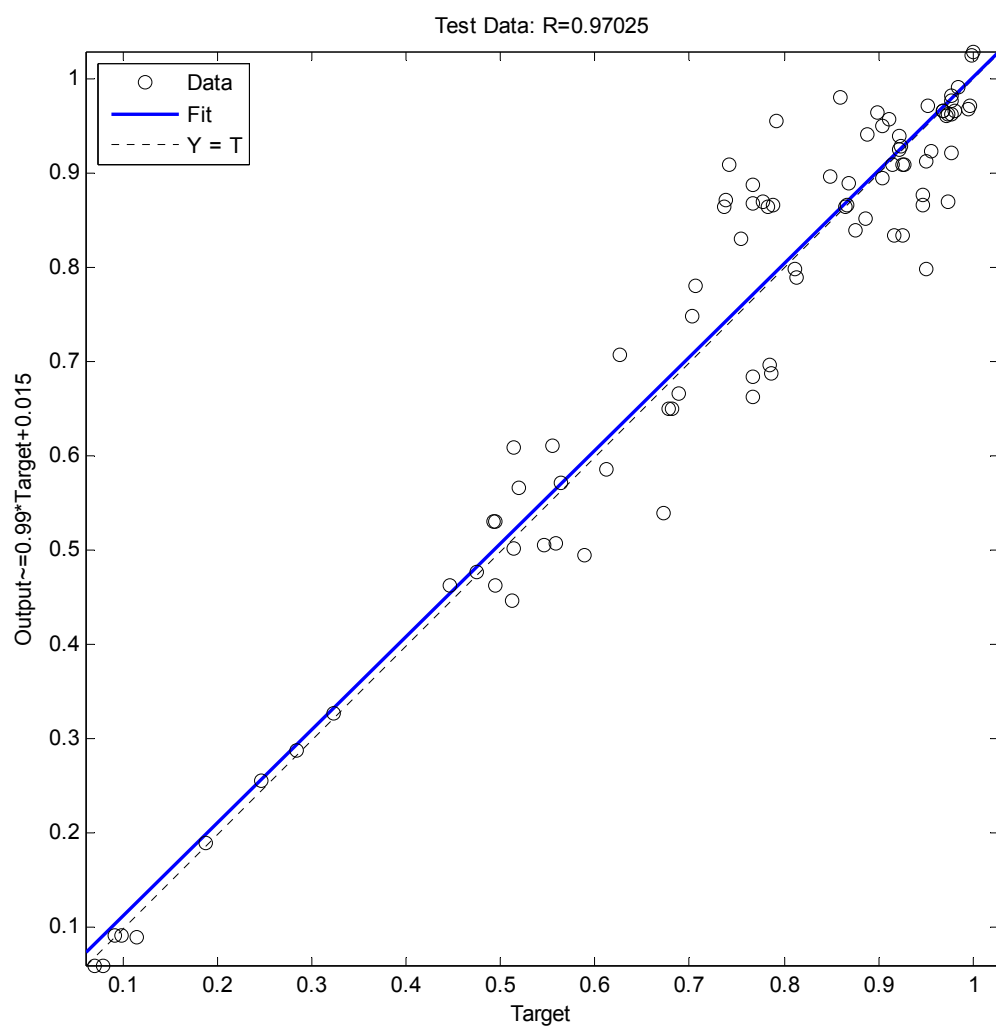


شکل ۴-۱۲-ب: نمودار همبستگی مقادیر اشباع از آب و تخمین شبکه ۱ جدول ۴-۱۳ برای داده های آموزش.

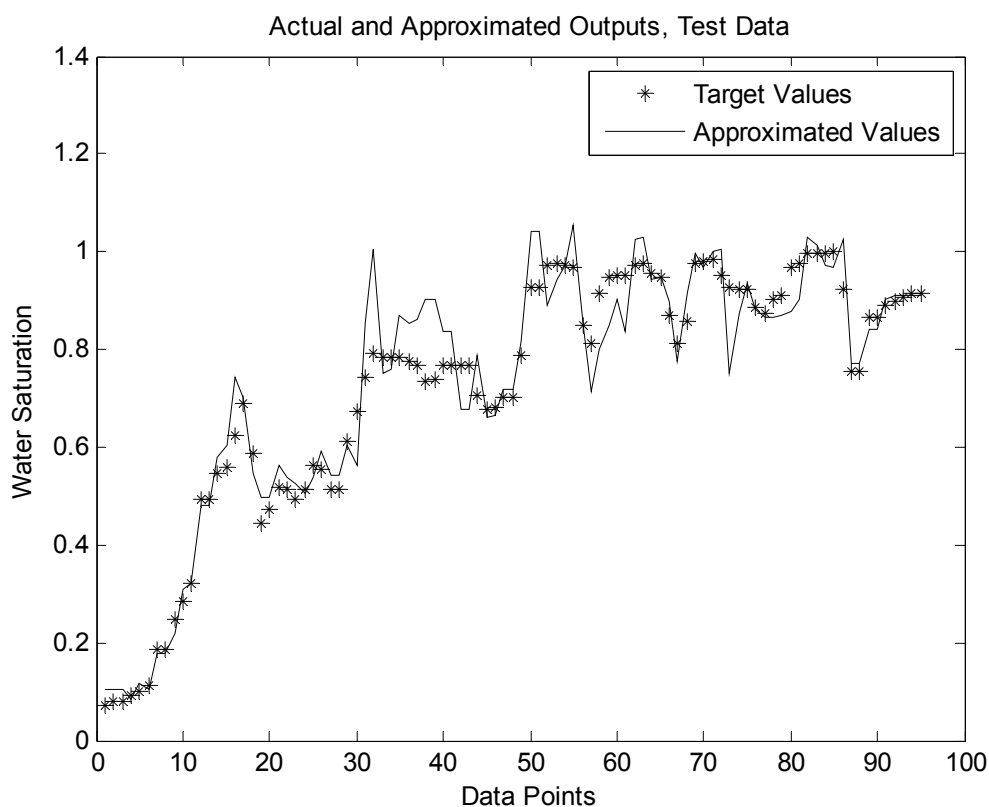


شکل ۴-۱۲: مقادیر اشباع از آب و تخمین شبکه شماره ۱ جدول ۴-۱۳ برای داده های آموزش.





شکل ۴-۱۲-د: نمودار همبستگی بین مقادیر اشباع از آب و تخمین شبکه ۱ جدول ۴-۱۳ برای داده های آزمون.



شکل ۴-۱۲-ه: مقادیر اشباع از آب و تخمین شبکه‌ی شماره‌ی ۱ جدول ۴-۱۳ برای داده‌های آزمون.

#### ۴-۹-۱- ترکیبات آنسامبلی خطی

۱۲۰ ترکیب خطی آنسامبلی شبکه‌های منفرد جدول ۴-۱۳ به دو روش الگوریتم ژنتیک و میانگین‌گیری ساده ایجاد شد. نتایج این ۱۲۰ ترکیب در جدول ۴-۱۴ ارائه شده است. ستون اول از سمت چپ (NO)، شماره ردیف و ستون دوم (Ensemble)، نوع ترکیب آنسامبلی را نشان می‌دهد. ستون‌های دیگر میانگین مربعات خطای تخمین را برای داده‌های آموزش و آزمون برای هر دو روش نشان می‌دهد. این خطاها تا ۴ رقم اعشار گرد شده‌اند.

جدول ۴-۱۴: نتایج ترکیبات آنسامبلی خطی به دو روش الگوریتم ژنتیک و میانگین گیری ساده.

NO.	Ensemble	optimal GA-train(%)	optimal GA-test(%)	Simple Avg-train(%)	Simple Avg-test(%)
1	1,2	0.1582	0.3479	0.1584	0.3478
2	1,3	0.1585	0.3156	0.1588	0.3143
3	1,4	0.1469	0.3407	0.1498	0.3329
4	1,5	0.1799	0.2876	0.1799	0.2877
5	1,6	0.1521	0.2041	0.1526	0.2009
6	1,7	0.1683	0.3571	0.1689	0.3580
7	2,3	0.1625	0.2885	0.1626	0.2881
8	2,4	0.1398	0.3349	0.1414	0.3304
9	2,5	0.1637	0.2989	0.1639	0.2991
10	2,6	0.1453	0.1940	0.1455	0.1930
11	2,7	0.1659	0.3304	0.1673	0.3305
12	3,4	0.1365	0.3208	0.1376	0.3183
13	3,5	0.1686	0.3002	0.1691	0.2991
14	3,6	0.1499	0.2616	0.1499	0.2614
15	3,7	0.1715	0.3179	0.1737	0.3140
16	4,5	0.1524	0.3144	0.1558	0.3043
17	4,6	0.1327	0.2213	0.1335	0.2186
18	4,7	0.1380	0.3335	0.1428	0.3272
19	5,6	0.1541	0.3147	0.1547	0.3134
20	5,7	0.1777	0.3059	0.1784	0.3046
21	6,7	0.1723	0.2639	0.1751	0.2527
22	1,2,3	0.1438	0.2854	0.1438	0.2853
23	1,2,4	0.1316	0.3143	0.1336	0.3122
24	1,2,5	0.1505	0.2864	0.1519	0.2778
25	1,2,6	0.1332	0.1836	0.1339	0.1915
26	1,2,7	0.1451	0.3226	0.1459	0.3223
27	1,3,4	0.1301	0.2964	0.1328	0.2900
28	1,3,5	0.1532	0.2704	0.1550	0.2611
29	1,3,6	0.1363	0.2040	0.1367	0.2052
30	1,3,7	0.1484	0.2967	0.1495	0.2982
31	1,4,5	0.1430	0.2952	0.1484	0.2714
32	1,4,6	0.1263	0.1990	0.1286	0.1942
33	1,4,7	0.1310	0.3144	0.1351	0.3121
34	2,3,4	0.1290	0.2888	0.1319	0.2772
35	2,3,5	0.1505	0.2542	0.1507	0.2544
36	2,3,6	0.1356	0.1860	0.1364	0.1899
37	2,3,7	0.1509	0.2722	0.1517	0.2742
38	2,4,5	0.1354	0.2914	0.1387	0.2753
<b>39</b>	<b>2,4,6</b>	<b>0.1219</b>	<b>0.1949</b>	<b>0.1229</b>	<b>0.1894</b>

40	2,4,7	0.1279	0.3051	0.1318	0.2986
41	2,5,6	0.1360	0.2084	0.1372	0.2173
42	2,5,7	0.1514	0.2754	0.1525	0.2743
43	3,4,5	0.1346	0.2904	0.1399	0.2680
44	3,4,6	0.1224	0.2145	0.1238	0.2126
45	3,4,7	0.1282	0.2941	0.1336	0.2841
46	3,5,6	0.1408	0.2453	0.1421	0.2459
47	3,5,7	0.1574	0.2667	0.1582	0.2652
48	4,5,6	0.1288	0.2174	0.1321	0.2289
49	4,5,7	0.1354	0.2967	0.1419	0.2731
50	1,5,6	0.1438	0.2167	0.1463	0.2158
51	1,5,7	0.1589	0.2857	0.1591	0.2816
52	1,6,7	0.1455	0.2047	0.1478	0.2194
53	2,6,7	0.1412	0.1896	0.1451	0.2036
54	3,6,7	0.1465	0.2321	0.1506	0.2248
55	4,6,7	0.1266	0.2185	0.1328	0.2114
56	5,6,7	0.1496	0.2606	0.1529	0.2431
57	4,5,6,7	0.1253	0.2124	0.1315	0.2153
58	3,5,6,7	0.1396	0.2284	0.1430	0.2217
59	3,4,6,7	0.1207	0.2133	0.1267	0.2088
60	3,4,5,7	0.1278	0.2825	0.1353	0.2533
61	3,4,5,6	0.1219	0.2106	0.1267	0.2154
62	2,5,6,7	0.1347	0.2005	0.1383	0.2108
63	2,4,6,7	0.1197	0.2028	0.1243	0.2009
64	2,4,5,7	0.1271	0.2873	0.1327	0.2625
65	2,4,5,6	0.1210	0.1943	0.1245	0.2033
66	2,3,6,7	0.1344	0.1836	0.1377	0.1981
67	2,3,5,7	0.1446	0.2475	0.1454	0.2476
68	2,3,5,6	0.1317	0.1929	0.1336	0.2025
69	1,5,6,7	0.1407	0.2059	0.1432	0.2168
70	1,4,6,7	0.1227	0.2100	0.1272	0.2105
71	1,4,5,7	0.1307	0.2998	0.1379	0.2672
72	1,4,5,6	0.1254	0.1980	0.1311	0.2032
73	1,3,6,7	0.1349	0.2020	0.1376	0.2136
74	1,3,5,7	0.1461	0.2666	0.1475	0.2583
75	1,3,5,6	0.1344	0.2025	0.1371	0.2083
76	1,3,4,7	0.1247	0.2857	0.1290	0.2804
77	1,3,4,6	0.1198	0.2035	0.1222	0.2011
78	1,2,6,7	0.1315	0.1905	0.1341	0.2110
79	1,2,5,7	0.1423	0.2840	0.1438	0.2728
80	1,2,5,6	0.1310	0.1868	0.1339	0.2016
81	1,2,4,7	0.1242	0.2990	0.1275	0.2980
82	1,2,4,6	0.1191	0.1926	0.1212	0.1974
83	1,2,3,7	0.1384	0.2804	0.1391	0.2798

84	1,2,3,6	0.1284	0.1816	0.1292	0.1932
85	2,3,4,7	0.1237	0.2792	0.1287	0.2663
86	2,3,4,6	0.1182	0.1949	0.1206	0.1918
87	2,3,4,5	0.1282	0.2730	0.1329	0.2498
88	1,3,4,5	0.1299	0.2872	0.1368	0.2542
89	1,2,4,5	0.1307	0.2931	0.1356	0.2677
90	1,2,3,5	0.1416	0.2594	0.1428	0.2515
91	1,2,3,4	0.1251	0.2815	0.1276	0.2758
92	3,4,5,6,7	0.1205	0.2105	0.1278	0.2092
93	2,4,5,6,7	0.1194	0.1999	0.1251	0.2048
94	2,3,5,6,7	0.1312	0.1897	0.1347	0.2022
95	2,3,4,6,7	0.1173	0.1992	0.1225	0.1991
96	2,3,4,5,7	0.1236	0.2738	0.1301	0.2448
97	2,3,4,5,6	0.1181	0.1938	0.1229	0.1986
98	1,4,5,6,7	0.1226	0.2069	0.1292	0.2091
99	1,3,5,6,7	0.1335	0.1992	0.1368	0.2104
100	1,3,4,6,7	0.1186	0.2074	0.1234	0.2095
101	1,3,4,5,7	0.1247	0.2856	0.1324	0.2520
102	1,3,4,5,6	0.1198	0.2029	0.1261	0.2027
103	1,2,5,6,7	0.1299	0.1886	0.1335	0.2094
104	1,2,4,6,7	0.1177	0.2022	0.1215	0.2105
105	1,2,4,5,7	0.1242	0.2955	0.1305	0.2639
106	1,2,4,5,6	0.1190	0.1916	0.1244	0.2010
107	1,2,3,6,7	0.1278	0.1846	0.1304	0.2057
108	1,2,3,5,7	0.1374	0.2618	0.1388	0.2515
<b>109</b>	<b>1,2,3,5,6</b>	<b>0.1276</b>	<b>0.1816</b>	<b>0.1305</b>	<b>0.1976</b>
110	1,2,3,4,7	0.1214	0.2773	0.1252	0.2708
111	1,2,3,4,6	0.1166	0.1931	0.1192	0.1981
112	1,2,3,4,5	0.1250	0.2785	0.1306	0.2493
113	1,2,3,4,5,6	0.1166	0.1930	0.1224	0.1991
114	1,2,3,4,5,7	0.1214	0.2771	0.1280	0.2485
115	1,2,3,4,6,7	0.1159	0.1986	0.1203	0.2071
116	1,2,3,5,6,7	0.1272	0.1831	0.1308	0.2041
117	1,2,4,5,6,7	0.1177	0.2015	0.1240	0.2079
118	1,3,4,5,6,7	0.1187	0.2066	0.1260	0.2068
119	2,3,4,5,6,7	0.1173	0.1983	0.1239	0.2008
120	1,2,3,4,5,6,7	0.1159	0.1986	0.1227	0.2046

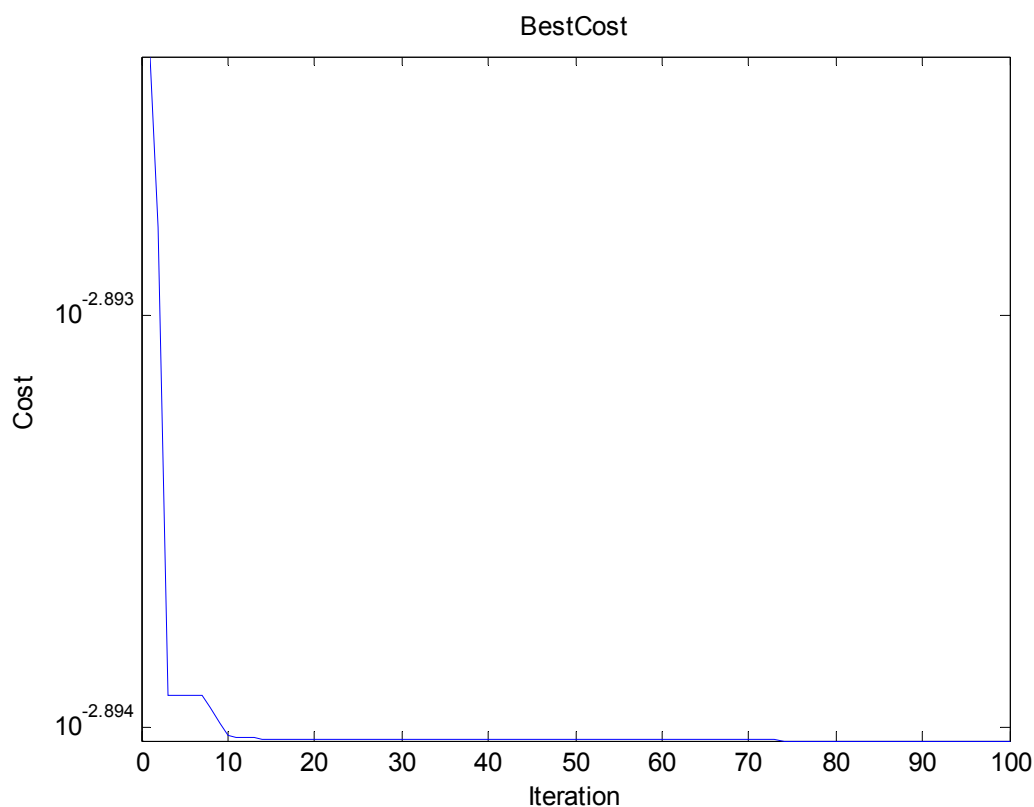
با توجه به جدول، ترکیب مربوط به ردیف ۱۰۹ (ترکیب شبکه های ۱، ۲، ۳، ۵ و ۶) برای الگوریتم

ژنتیک و ترکیب مربوط به ردیف ۳۹ (ترکیب شبکه های ۲، ۴ و ۶) برای میانگین گیری ساده، بهترین

جواب را از جهت کمترین خطای آزمون به همراه دارند. نتایج کامل تر در جدول ۴-۱۵ ارائه شده‌اند.

جدول ۴-۱۵: نتایج بهترین ترکیبات مربوط به هر دو روش میانگین گیری ساده و الگوریتم ژنتیک.

روش ترکیب خطی		نتایج
الگوریتم ژنتیک	میانگین گیری ساده	
۶ و ۵، ۳، ۲، ۱	۶ و ۴، ۲	ترکیب آنسامبلی
۰/۹۸۴۰	۰/۹۸۴۶	ضریب همبستگی داده های آموزش
۰/۱۲۷۶	۰/۱۲۲۹	MSE داده های آموزش(%)
%۴۱/۲۳	%۴۳/۳۹	کاهش در MSE داده های آموزش
۰/۹۸۳۱	۰/۹۸۲۷	ضریب همبستگی داده های آزمون
۰/۱۸۱۶	۰/۱۸۹۴	MSE داده های آزمون(%)
%۵۵/۱۵	%۵۳/۲۲	کاهش در MSE داده های آزمون
۰/۱۸۶۶	۰/۳۳	$W_1$
۰/۲۳۳۴	۰/۳۳	$W_2$
۰/۱۸۱۰	۰/۳۳	$W_3$
۰/۰۸۷۶	۰	$W_4$
۰/۳۱۱۴	۰	$W_5$
۰	۰	$W_6$
۰	۰	$W_7$
۱	۱	$\sum W$



شکل ۴-۱۳: روند کاهش تابع هزینه (MSE) در الگوریتم ژنتیک.

ساختار این جدول نیز مانند جدول (۴-۵) می باشد. میزان کاهش میانگین مربعات خطای داده‌های آموزش از طریق تقسیم حاصل تفریق میانگین مربعات خطای ترکیب و بهترین شبکه منفرد بر میانگین مربعات خطای بهترین شبکه منفرد محاسبه شده است. برای روش الگوریتم ژنتیک شبکه‌ی شماره‌ی ۶ جدول (۴-۱۳) بیشترین ضریب و شبکه‌ی شماره ۵ کمترین ضریب را به خود اختصاص داده‌اند. میزان کاهش خطای تخمین بهترین ترکیب برای داده‌های آزمون در روش الگوریتم ژنتیک بیشتر از روش میانگین گیری می‌باشد.

#### ۴-۹-۲- واریانس و بایاس خطا

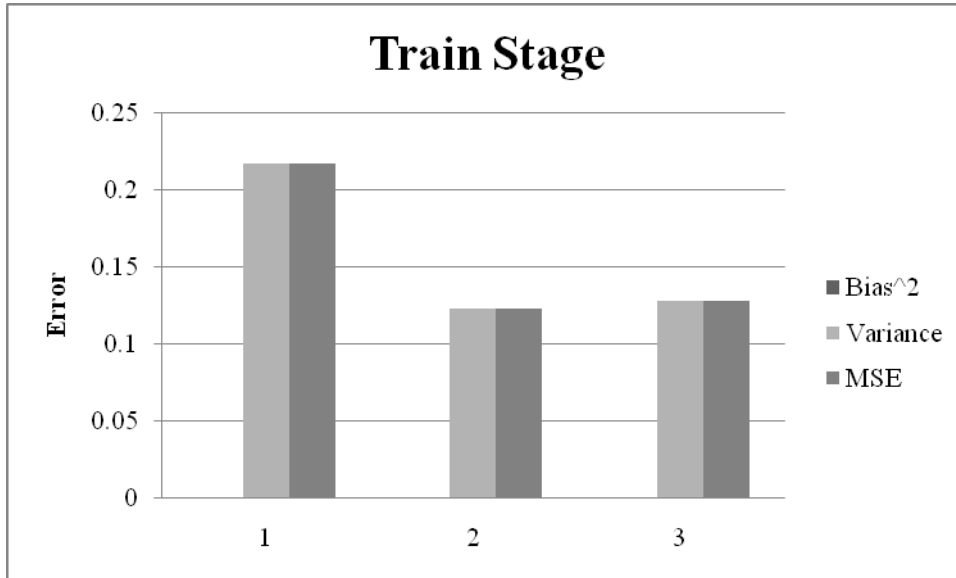
در جدول ۴-۱۶ واریانس، بایاس و میانگین مربعات خطای تخمین برای داده‌های آموزش و آزمون بهترین شبکه ی منفرد (۱) (شبکه ۱ جدول ۴-۱۳)، بهترین ترکیب با هر دو روش میانگین گیری ساده (۲) و الگوریتم ژنتیک (۳) ارائه شده است.

شکل ۴-۱۴ نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین برای داده‌های آموزش و آزمون حالات مربوط به جدول ۴-۱۶ را نشان می‌دهد. با توجه به شکل ۴-۱۴-الف برای داده‌های آموزش می‌بینیم که واریانس مؤلفه‌ی اصلی خطا است و بایاس تقریباً نقشی در خطا ندارد. برای داده‌های آزمون اگر چه نقش بایاس نسبت به داده‌های آموزش بیشتر است ولی باز هم می‌توان گفت که تقریباً نقشی در خطا ندارد و واریانس مؤلفه‌ی اصلی خطاست.

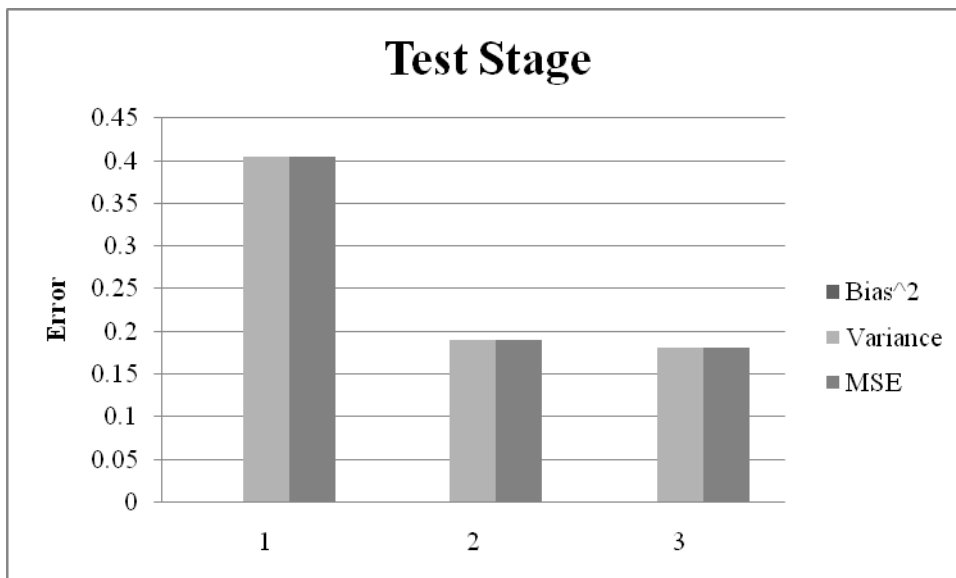
جدول ۴-۱۶: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه‌ی ۱ جدول ۴-۱۳ و بهترین ترکیب آن جدول.

ردیف	بایاس تخمین داده های آموزش	واریانس تخمین داده- های آموزش	میانگین مربعات خطای تخمین داده‌های آموزش	بایاس تخمین داده های آزمون	واریانس تخمین داده های آزمون	میانگین مربعات خطای تخمین داده های آزمون
۱	۰/۰۰۳۴	۰/۲۱۷۱	۰/۲۱۷۱	۰/۰۰۵۸	۰/۴۰۴۹	۰/۴۰۴۹
۲	۰/۰۰۱۲	۰/۱۲۲۹	۰/۱۲۲۹	۰/۰۰۷۳	۰/۱۸۹۳	۰/۱۸۹۴
۳	$-۵/۰۲ \times ۱۰^{-۵}$	۰/۱۲۷۶	۰/۱۲۷۶	۰/۰۰۳۳	۰/۱۸۱۶	۰/۱۸۱۶





شکل ۴-۱۴الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آموزش (۱: شبکه ۱ جدول ۴-۱۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).



شکل ۴-۱۴ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای داده‌های آزمون (۱: شبکه ۱ جدول ۴-۱۳، ۲: بهترین ترکیب با روش میانگین گیری ساده و ۳: بهترین ترکیب با روش الگوریتم ژنتیک).

## فصل پنجم

### نتایج و پیشنهادات

## ۵-۱- جمع‌بندی و نتیجه‌گیری

در این تحقیق از الگوریتم ژنتیک برای ترکیب آنسامبلی شبکه‌های جزء استفاده شد. این کار در سه بخش انجام شد:

۱- روش فرا آموزش

۲- روش مرتب سازی

۳- روش توقف سریع

در هر کدام از این موارد ۵۹ شبکه با ساختارهای مختلف ایجاد شدند و ۷ شبکه با کمترین میزان خطای آزمون برای ساختن ترکیبات آنسامبلی انتخاب شدند. تمام ترکیبات ممکن این ۷ شبکه با دو روش الگوریتم ژنتیک و میانگین گیری ساده ایجاد شد و میزان کاهش خطای MSE در هر دو بخش آموزش و آزمون برای بهترین ترکیبات در هر کدام از این دو روش محاسبه شد.

در روش فرا آموزش و برای داده‌های آموزش، الگوریتم ژنتیک خطای MSE را  $50/68\%$  کاهش داد و در روش میانگین گیری ساده خطای MSE به میزان  $49/32\%$  کاهش پیدا کرد. برای داده‌های آزمون نیز میزان کاهش خطای MSE برای هر دو روش الگوریتم ژنتیک و میانگین گیری ساده به ترتیب  $56/19\%$  و  $57/74\%$  به دست آمد.

در روش مرتب سازی و برای داده‌های آموزش، الگوریتم ژنتیک خطای MSE را  $25/57\%$  کاهش داد و در روش میانگین گیری ساده خطای MSE به میزان  $32/43\%$  کاهش پیدا کرد. برای داده‌های آزمون نیز میزان کاهش خطای MSE برای هر دو روش الگوریتم ژنتیک و میانگین گیری ساده به ترتیب  $44/69\%$  و  $46/62\%$  به دست آمد.

در روش توقف سریع و برای داده‌های آموزش، الگوریتم ژنتیک خطا را  $41/23\%$  کاهش داد و در روش میانگین گیری ساده خطا به میزان  $43/39\%$  کاهش پیدا کرد. برای داده‌های آزمون نیز میزان کاهش خطای MSE برای هر دو روش الگوریتم ژنتیک و میانگین گیری ساده به ترتیب  $55/15\%$  و  $53/22\%$  به دست آمد.

با توجه به این تحقیق می‌توان گفت ترکیب آنسامبلی شبکه‌های جزء تا حد قابل قبولی توانسته نتایج شبکه‌های منفرد را بهبود دهد. اینکه میزان کاهش خطای MSE و افزایش دقت چه میزان باشد به عوامل متعددی می‌تواند بستگی داشته باشد که به چند نمونه از آن‌ها اشاره می‌کنیم:

نوع مسأله‌ای که با آن روبرو هستیم بسیار مهم است. به این معنی که ماهیت کمیت‌های ورودی و خروجی چه می‌باشد. اینکه چه میزان داده در اختیار داشته باشیم و این داده‌ها حاوی نوفه هستند یا خیر نیز بر نتیجه‌ی نهایی تأثیر گذار است. همان‌طور که از نتایج مشخص است در بیشتر موارد میزان کاهش خطای MSE داده‌های آزمون برای روش میانگین گیری ساده بیشتر از میزان کاهش خطای MSE در روش الگوریتم ژنتیک است به گونه‌ای که در بهترین ترکیب روش فرا آموزش و مرتب‌سازی میزان این کاهش در روش میانگین گیری بیشتر از روش روش الگوریتم ژنتیک می‌باشد. اما در روش توقف سریع، در بهترین ترکیب الگوریتم ژنتیک، میزان کاهش خطا برای داده‌های آزمون از میزان کاهش خطای MSE داده‌های آزمون در روش میانگین گیری ساده بیشتر است. بنابراین شبکه‌های جزء و نوع روش آموزش به کار رفته نیز در نتیجه‌ی شبکه‌های آنسامبلی تأثیر بسزایی دارند به گونه‌ای که ممکن است در حالتی روش الگوریتم ژنتیک توان تعمیم بهتر و در حالت دیگر روش میانگین گیری ساده جواب بهتری را به همراه داشته باشد. باید توجه داشته باشیم که جواب بهتر روش میانگین گیری ساده در این مورد خاص می‌تواند به دلیل نویزی بودن داده‌ها نیز باشد.

نکته‌ی دیگر اینکه در ساخت ترکیبات آنسامبلی می‌توانیم ترکیبات بسیار زیادی را ایجاد کنیم. همان‌طور که در شبکه‌های منفرد نمی‌توانیم ادعا کنیم بهترین شبکه‌ی حاصل را به دست آورده‌ایم، در

ساخت ترکیبات هم نمی‌توانیم چنین ادعایی داشته باشیم. ترکیبات بسیار زیادی را می‌توان ساخت، از طرف دیگر ما بهترین شبکه‌های منفرد را برای ساخت این ترکیبات به کار بردیم و این در حالی است که شبکه‌های با خطای MSE بالاتر ممکن است در ترکیب بسیار خوب عمل کند و ضریب بزرگتری را به خود اختصاص دهد.

## ۵-۲- پیشنهادات

محدودیت مطالعات شبکه‌های آنسامبلی در تخمین پارامترهای پتروفیزیکی باعث شده تا نقاط قوت و ضعف آن‌ها در شرایط گوناگون مبهم باشد. اگرچه در این مطالعه قابلیت آن‌ها برای هر دو روش الگوریتم ژنتیک و میانگین‌گیری ساده به خوبی ثابت شد ولی با استفاده از آن نمی‌توان به زوایای مختلف عملکرد آن‌ها پی برد.

پیشنهاد می‌شود تخمین اشباع از آب و سایر خصوصیات مخزنی از جمله تراوایی، تخلخل و ... و همچنین تخمین نگارهای چاه در میداین مختلف، با روش ترکیب آنسامبلی صورت بگیرد. در این جا از الگوریتم ژنتیک و میانگین‌گیری ساده برای این کار استفاده شد. می‌توانیم از روش تحلیلی هاشم نیز در کنار این دو روش استفاده کنیم و قابلیت این ۳ روش را در یک مورد خاص مورد بررسی قرار دهیم. همچنین می‌توانیم از روش‌های بهینه‌سازی دیگر مانند Simulated Annealing و ... نیز برای تعیین ضرایب ترکیب استفاده کنیم و قابلیت آن‌ها را مورد بررسی قرار دهیم.

در راستای بهبود نتایج می‌توانیم از ترکیب مدوله‌ای نیز استفاده کنیم. بدین منظور می‌توانیم دامنه‌ی تغییرات اشباع از آب را به چند زیر دامنه تقسیم کنیم و در هر زیر دامنه از شبکه یا ترکیب شبکه‌ها استفاده نمود. به طور کلی ترکیب مدوله‌ای در مواردی که دامنه‌ی تغییرات کمیت‌ها به خصوص برای خروجی وسیع است، می‌تواند دقت تخمین را بالا ببرد.

مواردی که به آنها اشاره شد گوشه‌ای از تحقیقات فراوانی است که می‌تواند در این موضوع صورت

بگیرد.

# فهرست منابع

## منابع فارسی

۱. افشارحرب، ع.، ۱۳۸۱، زمین‌شناسی نفت، انتشارات پیام نور.
۲. انصاری، ع.، ۱۳۷۶، چاه‌پیمایی در هیدروژئولوژی، (مؤلف: دومینیکو کاپلییر)، انتشارات دانشگاه یزد.
۳. باولی بهمنی، ع.، ۱۳۸۲، اصول و چگونگی استخراج پارامترهای پتروفیزیکی سنگ مخزن نفت با استفاده از نگارهای چاه‌پیمایی، سمینار کارشناسی ارشد مهندسی اکتشاف معدن، دانشگاه صنعتی شاهرود.
۴. تدینی، م.، نبی‌بیدهدی، م.، ۱۳۸۵، تعیین خواص پتروفیزیکی سنگ مخزن میدان پارسی با استفاده از شبکه عصبی، اولین کنگره مهندسی نفت ایران (صنایع بالادستی نفت).
۵. حبیبیان، ب.، نبی‌بیدهدی، م. و کاظم‌زاده، ع.، ۱۳۸۴، پیش‌بینی نفوذپذیری از روی داده‌های چاه‌نگاری با استفاده از شبکه‌های عصبی مصنوعی در یکی از مخازن کربناته جنوب ایران، مجله‌ی فیزیک زمین و فضا، جلد ۳۱، شماره‌ی ۱، ۸۶-۷۹.
۶. حسنی پاک، ع.ا. و شرف‌الدین، م.، ۱۳۸۰، تحلیل داده‌های اکتشافی، انتشارات دانشگاه تهران.
۷. ذاکری، م.، رحیمی، م.، کامکار روحانی، ا. و پویان، ع.ا.، ۱۳۸۶، تخمین تخلخل سنگ مخزن گازی کنگان در میدان پارس جنوبی با استفاده از ترکیب آنسامبلی شبکه‌های عصبی مصنوعی، مجموعه مقالات دومین کنگره‌ی ملی مهندسی نفت ایران، جلد دوم، ۴۳-۳۳.
۸. ذاکری، م.، ۱۳۸۷، تخمین تخلخل سنگ مخزن گازی کنگان در میدان پارس جنوبی با استفاده از شبکه‌های عصبی مصنوعی کمیته‌ای، پایان‌نامه کارشناسی ارشد ژئوفیزیک، دانشگاه صنعتی شاهرود.
۹. رضایی، م.ر.، ۱۳۸۰، زمین‌شناسی نفت، انتشارات علوی.



۱۰. رضایی، ع. و رنجبران، س.، ۱۳۸۶، آموزش کاربردی الگوریتم ژنتیک در نرم‌افزار MATLAB، چاپ اول، کتاب پدیده، تهران.
۱۱. علیرضا، م.، ۱۳۸۶، مقدمه‌ای بر الگوریتم‌های ژنتیک و کاربردهای آن، چاپ دوم، شرکت ناقوس اندیشه، تهران.
۱۲. کامکار روحانی، ا.، ذاکری، م.، پویان، ع.ا. و رحیمی، م.، ۱۳۸۷، تخمین نگار نوترون با استفاده از ترکیب آنسامبلی شبکه‌های عصبی مصنوعی، مجموعه مقالات سیزدهمین کنفرانس ژئوفیزیک ایران، ۲۹۴-۲۹۸.
۱۳. مرادزاده، ع. و قوامی ریابی، ر.، ۱۳۸۰، چاه‌پیمایی برای مهندسين، انتشارات دانشگاه صنعتی شاهرود.
۱۴. منهاج، م.ب.، ۱۳۷۹، مبانی شبکه‌های عصبی (جلد اول)، انتشارات دانشگاه صنعتی امیرکبیر.
۱۵. نصرآبادی، ا.، نبی‌بیدهندی، م.، جواهریان، ع.، حقیقی، م. و سپهوند، م.ر.، ۱۳۸۲، تخمین تخلخل، تراوایی و اشباع آب از روی نگارهای چاه در میدان گازی نار، یازدهمین کنفرانس ژئوفیزیک ایران، انجمن ژئوفیزیک ایران، تهران، ۳۰۱ - ۲۹۹.

#### منابع انگلیسی:

16. Al-Bulushi, N., Araujo, M. and Kraaijveld, N., 2007, Predicting Water Saturation Using Artificial Neural Networks (ANNS), 25th IASTED International Multi-Conference: Artificial Intelligence and Applications, 57-62, Innsbruck, Austria.
17. Bhatt, A. and Helle H. B., 2002, Committee Neural Networks for Porosity and Permeability Prediction from Well Logs, Geophysical Prospecting, 50, 645-660.
18. Bhatt, A., 2002, Reservoir Properties from Well Logs Using Neural Networks, PhD thesis, Department of Petroleum Engineering and

Applied Geophysics, Norwegian University of Science and Technology.

19. Chen, Ch. H. and Lin Z. Sh., 2006, A Committee Machine with Empirical Formulas for Permeability Prediction, *Computer and Geosciences*, 32, 485-496.
20. Cho, S. B., 1999, Pattern Recognition with Neural Networks Combined by Genetic Algorithm, *Fuzzy sets and systems*, 103, 339-347.
21. Demuth, H. and Beale M., 2008, *Neural network Toolbox User's Guide of MATLAB, Version 6*, Online in [www.mathworks.com](http://www.mathworks.com).
22. Elliis, D. V., Singer, J. M. 2007, *Well Logging for Earth Scientists*, Chapter 23, Springer.
23. Fung, Ch. Ch., Wong K. W. and Eren H., 1997, Modular Artificial Neural Network for Prediction of Petrophysical properties from well log data, *IEEE Transaction on Instrumentation and Measurement*, 46, 1295-1299.
24. Geological Report, Research Institute of Petroleum Industry.
25. Gopinath, P. and Reddy N. P., 2000, Toward Intelligent Web Monitoring: Performance of Committee Neural Networks vs. Single Neural Network, *Information Technology Applications in Biomedicine, Proceedings of IEEE EMBs International Conference*, 179-182.
26. Hashem, S., 1993, *Optimal Linear Combination of Neural Networks*, PhD Thesis, School of Industrial Engineering, Purdue University.
27. Haupt, R. L. and Haupt, E., 2004, *Practical Genetic Algorithms*, Second Edition, John Wiley and sons, New York.

28. Hearst, J. R., Nelson, P. H. and Paillet F. L., 2000, Well Logging for Physical Properties, A Hand Book for Geophysicists, Geologists and Engineers, John Wiley and sons Ltd.
29. Helle, H. B., Bhatt, A. and Ursin B., 2001, Porosity and Permeability Prediction from Wire Line Logs Using Artificial Neural Networks: A North Sea Case Study, Geophysical Prospecting, 49, 431-444.
30. Helle, H. B. and Bhatt, A., 2002, Fluid Saturation from Well Logs Using Committee Neural Networks, Petroleum Geoscience, 8, 109-118.
31. <http://www.obitko.com/tutorials/genetic-algorithms/>.
32. Huang, Z., Shimeld, J., Williamson, M. and Katsube J., 1996, Permeability Prediction with Artificial Neural Network Modeling in the Venture Gas Field Offshore Eastern Canada, Geophysics, 61, 422-436.
33. Huang, Z. and Williamson, M., 1997, Determination of Porosity and Permeability in Reservoir Intervals by Artificial Neural Networks Modeling Offshore Eastern Canada, Petroleum Geoscience, 3, 245-258.
34. Kamkar Rouhani, A. and Zakeri, M., 2009, Using a Genetic Algorithm in Constructing Ensemble Combinations of Neural Networks to Estimate Reservoir Rock Porosity, First International Petroleum Conference and Exhibition, P53, Shiraz, Iran.
35. Link, C. A., Himmer Ph. A., 2003, Oil Reservoir Porosity Prediction Using a Neural Network Ensemble Approach, In Geophysical Application of Artificial Neural Networks and Fuzzy Logic (Eds W.Sandham and M.Leggett), 197-213, Kluwer academic publishers.

36. Lucia, F. J., 2007, Carbonate Reservoir Characterization, an Integrated Approach, Second Edition, Springer.
37. Natick, M., 2004, Genetic Algorithm and Direct Search Toolbox for Use with MATLAB, version 1, Online in [www.mathworks.com](http://www.mathworks.com).
38. Olson, T. M., 1998, Porosity and Permeability Prediction in Low-Permeability Gas Reservoirs From Well Logs Using Neural Networks, Rocky Mountain Regional Meeting / Low Permeability Reservoirs Symposium and Exhibition, 563-572, Denver, Colorado.
39. Selly, R. C., 1998, Elements of Petroleum Geology, Second edition, Academic press.
40. Sharkey, A., Sharkey, N., Gerecke, U. and Chandroth G.O., 2000, The "Test and Select Approach to Ensemble Combination, In Multiple Classifier Systems (Eds J.Kittler and F.Roli), 30-44, Lecture Notes in Computer Science, Springer-Verlag, Inc.
41. Sivanandam, S. N. and Deepa S. N., 2008, Introduction to Genetic Algorithms, Springer-Berlin Heidelberg, New York.
42. Tiab, D. and Donaldson E. C., 2004, Petrophysics. Theory and Practice of Measuring Reservoir Rock and Fluid Properties, Gulf publishing co.

پیوست

پیوست الف:

انواع الگوریتم‌های آموزشی پس انتشار خطا و  
توابع مربوط به آنها در نرم‌افزار MATLAB

## انواع الگوریتم‌های آموزشی پس انتشار خطا و توابع مربوط به آن‌ها در نرم‌افزار

### MATLAB

در آموزش با الگوریتم پس انتشار خطا دو روند آموزش مختلف وجود دارد:

۱. آموزش الگو به الگو<sup>۱</sup>: در این روش به ازای هر بردار الگوی آموزشی ورودی به شبکه، مسیر

رفت و برگشت الگوریتم پس انتشار خطا طی شده و تغییر پارامترهای شبکه محاسبه و

تصحیح انجام می‌شود. این روند برای هر الگوی آموزشی تکرار می‌شود.

۲. آموزش گروهی<sup>۲</sup>: در این روش به ازای هر بردار الگوی آموزشی ورودی، مسیر رفت و برگشت

طی شده و تغییر پارامترهای شبکه محاسبه می‌گردد، اما تصحیح پارامترها انجام نمی‌شود. به

عبارت دیگر برای تمام الگوهای آموزشی تنها تغییرات پارامترها محاسبه می‌شوند. وقتی که

تمام الگوهای آموزشی به شبکه عرضه شد و تغییرات پارامترها محاسبه گردید، تغییرات هر

پارامتر مشخص به ازای تمام الگوهای آموزشی با هم جمع می‌شود تا تغییر آن پارامتر را

مشخص کند. سپس تصحیح انجام می‌شود.

بعضی از توابع آموزش مانند تابع تفریق گرادیان (Traingd) با هر دو روند یاد شده قابل استفاده

هستند ولی بعضی از آن‌ها مانند تابع لونبرگ – مارکوت (Trainlm) فقط با روند گروهی قابل استفاده

می‌باشند [۲۱].

حال به شرح انواع الگوریتم‌های آموزشی پس انتشار خطا می‌پردازیم:

همه‌ی الگوریتم‌های آموزشی پس انتشار خطا از گرادیان تابع عملکرد نسبت به بردار پارامترهای شبکه

استفاده می‌کنند تا پارامترها را به گونه‌ای تغییر دهند که تابع عملکرد شبکه را کمینه کند.

## الف-۱- الگوریتم پس انتشار خطای تفریق گرادیان<sup>۱</sup>

این الگوریتم در سال ۱۹۸۶ پیشنهاد شد. در این الگوریتم پارامترهای شبکه در جهت منفی گرادیان تغییر می‌کنند و تابع عملکرد به سرعت کاهش می‌یابد، به همین دلیل به آن الگوریتم بیشترین شیب نیز گفته می‌شود.

$$\vec{x}_{k+1} = \vec{x}_k - \alpha_k \vec{g}_k \quad (\text{الف-۱})$$

در این رابطه  $\vec{x}_k$  بردار پارامترهای شبکه،  $\vec{g}_k$  بردار گرادیان تابع عملکرد نسبت به پارامترهای شبکه و  $\alpha_k$  سرعت یادگیری<sup>۲</sup> است. تابع مربوط به این الگوریتم در نرم‌افزار MATLAB، Traingd است. در رابطه با این تابع هفت پارامتر آموزشی وجود دارد که می‌توانند توسط کاربر تعیین شوند و در غیر این صورت مقادیر پیش‌فرض خود را حفظ خواهند کرد. این هفت پارامتر عبارتند از: epochs, show, goal, time, min\_grad, max\_fail, lr و lr پارامتر سرعت یادگیری است. هر چه مقدار آن بزرگتر باشد، تغییرات پارامترهای شبکه در یک تصحیح، بزرگتر خواهد بود اما باید توجه داشت که چنانچه مقدار آن به طور نامناسبی بزرگ باشد، الگوریتم ناپایدار خواهد شد و اگر مقدار آن کوچک انتخاب شود، زمان زیادی برای همگرایی الگوریتم نیاز خواهد بود. پارامتر show مشخص می‌کند که بعد از هر چند دوره آموزش، وضعیت آموزش نشان داده شود. بقیه پارامترها مشخص می‌کنند که چه موقع آموزش متوقف شود. epochs تعداد دوره‌های تکرار را مشخص می‌کند. اگر تعداد تکرارها به آن برسد، یا اگر مقدار گرادیان کمتر از min\_grad شود و یا اگر مدت زمان آموزش بر حسب ثانیه بیشتر از Time شود، آموزش متوقف می‌شود. پارامتر max\_fail که در آموزش به روش توقف سریع به کار می‌رود تعداد تکرارهایی را مشخص می‌کند که چنانچه خطای داده‌های اعتبار سنجی افزایش یابد، آموزش متوقف می‌شود.

---

1. Gradient Descent Algorithm

2. Learning Rate



ضعف عمده‌ی این الگوریتم، سرعت کم همگرایی آن و نیز توقف یادگیری شبکه قبل از فراگیری کامل مسأله به دلیل به دام افتادن پارامترهای شبکه در نقاط کمینه‌ی محلی می‌باشد [۲۱].

## الف-۲- الگوریتم پس‌انتشار خطای تفریق گرادیان با تکانه<sup>۱</sup>

این الگوریتم نسبت به الگوریتم تفریق گرادیان اغلب همگرایی سریع‌تری را موجب می‌شود. تکانه به شبکه اجازه می‌دهد که تغییرات پارامترها را به گونه‌ای انجام دهد که علاوه بر گرادیان محلی، روند نهایی رویه‌ی خطا را نیز تأثیر دهد. بدون تکانه ممکن است شبکه در یک کمینه‌ی محلی به دام بیفتد ولی تکانه همانند یک فیلتر پایین‌گذر باعث می‌شود که شبکه ویژگی‌های کوچک رویه‌ی خطا را فراموش کند و از چنین کمینه‌هایی عبور کند. در این الگوریتم تغییرات پارامترهای شبکه مجموع دو جمله است. یک جمله کسری از حاصل ضرب گرادیان تابع عملکرد در سرعت یادگیری و جمله‌ی دیگر کسری از آخرین تغییرات پارامترهای شبکه است. میزان تأثیر آخرین تغییرات پارامترها به وسیله‌ی ضریب تکانه ( $mc$ ) تعیین می‌شود که می‌تواند عددی بین صفر و یک انتخاب شود. وقتی که این ثابت صفر انتخاب می‌شود تغییرات پارامترها تنها بر اساس گرادیان تابع عملکرد صورت می‌گیرد و وقتی که یک انتخاب می‌شود تغییرات پارامترها برابر آخرین تغییرات پارامترها است. در یک تکرار معین چنانچه تابع عملکرد جدید از تابع عملکرد تکرار قبل به نسبتی بیش از پارامتر  $max\_perf\_inc$  که معمولاً  $1/0.4$  تعیین می‌شود، فزونی گیرد، پارامترهای حاصل غیر مفید هستند و رها می‌شوند و  $mc$  صفر انتخاب و تغییرات پارامترها تنها بر اساس جمله‌ی گرادیان تعیین می‌شود. در MATLAB برای این الگوریتم تابع  $Traingdm$  پیش‌بینی شده است. با توجه به توضیحاتی که داده شد مشخص است که این تابع علاوه بر پارامترهایی که برای تابع  $Traingd$  ذکر شد دو پارامتر اضافی  $mc$  و  $max\_perf\_inc$  نیز دارد. این تابع از روند آموزش گروهی استفاده می‌کند [۲۱].

---

1. Gradient Descent with Momentum

### الف-۳- الگوریتم‌های پس‌انتشار خطا با سرعت یادگیری متغیر<sup>۱</sup>

در روش‌های تفریق‌گرادیان سرعت یادگیری در کل مرحله‌ی آموزش ثابت می‌ماند و کارایی الگوریتم به تعیین مناسب سرعت یادگیری وابستگی زیادی دارد. اگر سرعت یادگیری بسیار بزرگ انتخاب شود، ممکن است الگوریتم دچار نوسان و ناپایداری شود و چنان‌چه خیلی کوچک انتخاب شود، زمان زیادی برای همگرایی الگوریتم نیاز است. در عمل ممکن نیست که قبل از آموزش، سرعت یادگیری به طور بهینه تعیین شود و در واقع در طی آموزش با حرکت بر روی رویه‌ی خطا سرعت یادگیری بهینه تغییر می‌کند.

اگر اجازه دهیم که سرعت یادگیری در طی آموزش تغییر کند، کارایی الگوریتم بیشترین شیب بهبود خواهد یافت. سرعت یادگیری قابل تنظیم باعث می‌شود که تغییرات پارامترها تا آن‌جایی که ممکن است بزرگ باشد و در عین حال یادگیری پایدار بماند. در هر دوره با استفاده از سرعت یادگیری موجود، پارامترهای شبکه محاسبه و خروجی‌ها و خطاها تعیین می‌شوند. اگر خطای جدید نسبت به خطای تکرار قبل به نسبتی بیش از پارامتر `max_perf_inc` که معمولاً  $1/0.4$  انتخاب می‌شود فزونی گیرد، پارامترهای جدید غیر مفید هستند و رها می‌شوند. همچنین سرعت یادگیری با ضرب شدن در پارامتر `lr_dec` که معمولاً  $0.7$  انتخاب می‌شود، کاهش می‌یابد. در غیر این صورت پارامترهای جدید مفید هستند و نگه داشته می‌شوند. اگر خطای جدید از خطای قبلی کوچکتر باشد، سرعت آموزش با ضرب شدن در پارامتر `lr_inc` که معمولاً  $1/0.5$  انتخاب می‌شود، افزایش می‌یابد. این روند سرعت یادگیری را افزایش می‌دهد، البته فقط تا اندازه‌ای که شبکه بدون افزایش خطای زیادی آموزش ببیند. بنابراین یک سرعت یادگیری نزدیک به بهینه برای وضعیت محلی رویه‌ی خطا به دست می‌آید. وقتی که یک سرعت یادگیری بزرگتر بتواند به یادگیری پایدار منجر شود، سرعت یادگیری افزایش می‌یابد.

---

1. Variable Learning Rate

همچنین وقتی که سرعت یادگیری به اندازه‌ای بالاست که کاهش خطا را تضمین نکند، کاهش می‌یابد تا پایداری حفظ شود.

در نرم‌افزار MATLAB تابع Traingda به منظور آموزش به این روش طراحی شده است. در کاربرد این تابع علاوه بر پارامترهای آموزشی که در مورد تابع Traingd به آن‌ها اشاره شد، سه پارامتر اضافی دیگر نیز وجود دارد که عبارتند از max\_perf\_inc، lr\_dec و lr\_inc که در مورد آن‌ها توضیحات لازم بیان شد. علاوه بر تابع Traingda، تابع Traingdx نیز برای آموزش با سرعت یادگیری متغیر طراحی شده است. این تابع دو الگوریتم آموزش با سرعت یادگیری قابل تنظیم و آموزش با تکانه را با هم ترکیب می‌کند و واضح است که علاوه بر پارامترهای تابع Traingda پارامتر اضافی ضریب تکانه (mc) را نیز دارد [۲۱].

#### الف-۴- الگوریتم پس‌انتشار خطای بهبودپذیر<sup>۱</sup>

معمولاً شبکه‌های چند لایه در لایه‌های میانی خود از توابع انتقال زیگموئیدی استفاده می‌کنند. این توابع یک دامنه‌ی ورودی گسترده و نا معین را به گستره‌ای مشخص و معین از خروجی‌ها می‌نگارند. شیب این توابع به ازای یک ورودی بزرگ به صفر نزدیک می‌شود. این موضوع باعث بروز مشکل به هنگام استفاده از روش بیشترین شیب می‌شود، زیرا گرادیان مقدار بسیار کوچکی پیدا می‌کند و حتی اگر پارامترهای شبکه بسیار دورتر از مقادیر بهینه‌شان باشند، تغییرات آن‌ها بسیار کوچک خواهد بود. هدف از الگوریتم پس‌انتشار بهبودپذیر آن است که اثرات مقدار مشتقات جزئی حذف شود. در این الگوریتم تنها از علامت مشتق به منظور تعیین جهت تغییر پارامترها استفاده می‌شود و مقدار مشتق تأثیری بر تنظیم پارامترها ندارد. اندازه‌ی تغییر پارامترها به وسیله‌ی مقدار قابل تنظیم جداگانه‌ای تعیین می‌شود. هرگاه که مشتق تابع عملکرد نسبت به یک پارامتر خاص در دو تکرار متوالی هم-

---

1. Resilient Backpropagation

علامت باشد، این مقدار قابل تنظیم برای آن پارامتر با استفاده از فاکتور  $delt\_inc$  افزایش می‌یابد. همچنین هر وقت که این مشتق نسبت به تکرار قبلی تغییر علامت دهد، مقدرا قابل تنظیم مقدار قابل تنظیم به وسیله‌ی فاکتور  $delt\_dec$  کاهش می‌یابد. اگر مشتق صفر باشد مقدار قابل تنظیم ثابت باقی می‌ماند. هرگاه که پارامتر مورد نظر نوسان کند، تغییرات آن کوچک می‌شود. اگر در چند تکرار متوالی تغییر پارامتر در یک جهت صورت گیرد، بزرگی تغییرات آن پارامتر افزایش می‌یابد. در نرم‌افزار MATLAB تابع  $Trainrp$  برای آموزش با این الگوریتم پیش‌بینی شده است. پارامترهای آموزشی این تابع عبارتند از: بیشینه تغییرات پارامترها در یک تصحیح می‌باشد. کارایی تابع  $Trainrp$  چندان متأثر از تنظیم پارامترهای آموزشی نیست و بیشتر این پارامترها می‌توانند مقدار پیش‌فرضشان باقی بمانند. به طور کلی این الگوریتم از الگوریتم‌هایی که پیش از این توضیح داده شد خیلی سریع‌تر همگرا می‌شود [۲۱].

## الف-۵- الگوریتم‌های گرادیان مزدوج<sup>۱</sup>

این روش‌ها در سال ۱۹۹۲ جهت آموزش ساختارهای عصبی ارائه شد. در الگوریتم پس‌انتشار خطای بیشترین شیب، پارامترهای شبکه در جهت منفی گرادیان تغییر می‌کنند. چنان‌چه قبلاً اشاره شد در این جهت تابع عملکرد سریعترین کاهش را خواهد داشت، اما این امر به این معنی نیست که سریع‌ترین همگرایی نیز در این جهت اتفاق می‌افتد. در روش گرادیان مزدوج تغییرات در جهت‌های مزدوج صورت می‌گیرد که به طور کلی نسبت به این روش بیشترین شیب ده‌ها برابر سریعتر همگرا می‌شود.

دو مسیر  $\bar{p}_1$  و  $\bar{p}_2$  را نسبت به ماتریس هسین<sup>۲</sup> مزدوج گویند هرگاه داشته باشیم:

$$\bar{p}_i^T \cdot H \cdot \bar{p}_j = 0 ; \forall i, j \quad , \quad i \neq j \quad (\text{الف-۲})$$

1. Conjugate Gradient Algorithm

2. Hessian Matrix

ماتریس هسین ماتریس مشتقات دوم تابع عملکرد است و با فرض این که تعداد پارامترهای شبکه  $N$  باشد یک ماتریس  $N \times N$  خواهد بود [۲۱].

در بیشتر الگوریتم‌های گرادیان مزدوج، در هر تکرار، اندازه تغییرات تنظیم می‌شود. جستجویی در جهت گرادیان مزدوج انجام می‌شود تا تغییرات به گونه‌ای انتخاب شود که تابع عملکرد را در امتداد آن مسیر کمینه کند. در جعبه ابزار MATLAB ۵ تابع جستجوی خطی مختلف وجود دارد. بعضی از این توابع جستجو یا برخی از توابع آموزشی روش گرادیان مزدوج، که در ادامه به آن‌ها می‌پردازیم، تناسب بیشتری دارند، اگرچه در نهایت انتخاب بهینه تابع جستجو به ماهیت مسأله بستگی دارد [۲۱]. در این جا به منظور اختصار، از شرح توابع جستجو صرف نظر می‌کنیم و فقط توابع آموزشی روش گرادیان مزدوج را شرح می‌دهیم.

### الف-۵-۱- الگوریتم فلچر - ریویس<sup>۱</sup>

همه‌ی الگوریتم‌های گرادیان مزدوج، در نخستین تکرار در جهت منفی گرادیان جستجو می‌کنند.

$$\vec{p}_0 = -\vec{g}_0 \quad (\text{الف-۳})$$

یک جستجوی خطی انجام می‌شود تا اندازه تغییرات بهینه مشخص شود و سپس تغییرات در جهت جاری اعمال می‌شود:

$$\vec{x}_{k+1} = \vec{x}_k + \alpha_k \vec{p}_k \quad (\text{الف-۴})$$

---

1. Fletcher - Reeves

جهت جستجوی جدید طوری انتخاب می‌شود که مزدوج جهت جستجوی قبلی باشد. روال کلی تعیین راستای جستجوی جدید آن است که جهت منفی گرادیان جاری را با جهت جستجوی قبلی ترکیب کنیم:

$$\vec{p}_k = -\vec{g}_k + \beta_k \vec{p}_{k-1} \quad (\text{الف-۵})$$

در روش‌های مختلف گرادیان مزدوج، ثابت  $\beta_k$  به روش‌های مختلفی محاسبه می‌شود. در روش فلچر - ریویس  $\beta_k$  به صورت زیر به دست می‌آید:

$$\beta_k = \frac{\vec{g}_k^T \vec{g}_k}{\vec{g}_{k-1}^T \vec{g}_{k-1}} \quad (\text{الف-۶})$$

در نرم‌افزار MATLAB تابع Traincgf برای آموزش با الگوریتم فلچر- ریویس تهیه شده است. پارامترهای آموزشی این تابع عبارتند از show, epochs, goal, time, min\_grad, max\_fail, srchFch, alpha, beta, delta, gama, low\_lim, up\_lim, maxstep, minstep و bmax. شش پارامتر اول قبلاً توضیح داده شده‌اند. پارامتر SrchFch نام تابع جستجوی خطی را مشخص می‌کند. بقیه پارامترها مربوط به تابع جستجو هستند. به طور کلی تابع Traincgf در مقایسه با Trainrp در تکرارهای کمتری همگرا می‌شود، اگر چه در هر تکرار محاسبات بیشتری نیاز است. معمولاً الگوریتم‌های گرادیان مزدوج از الگوریتم‌های پس‌انتشار خطای سرعت یادگیری متغیر بسیار سریع‌تر هستند و گاهی اوقات نیز نسبت به الگوریتم پس‌انتشار خطای بهبودپذیر سریع‌تر همگرا می‌شوند، اگر چه نتایج از مسأله‌ای به مسأله دیگر تفاوت دارد. این الگوریتم‌ها نسبت به الگوریتم‌های ساده‌تر فقط به میزان کمی حافظه بیشتر نیاز دارند و بنابراین انتخاب خوبی برای شبکه‌هایی با پارامترهای زیاد می‌باشند [۲۱].

## الف-۵-۲- الگوریتم پولاک-ریبیر<sup>۱</sup>

نوع دیگری از الگوریتم گرادیان مزدوج است که مانند روش فلچر-ریویس جهت جستجو در هر تکرار با رابطه زیر مشخص می‌شود:

$$\vec{p}_k = -\vec{g}_k + \beta_k \vec{p}_{k-1} \quad (\text{الف-۷})$$

در این روش ثابت  $\beta_k$  به صورت زیر تعیین می‌شود:

$$\beta_k = \frac{\Delta \vec{g}_{k-1}^T \vec{g}_k}{\vec{g}_{k-1}^T \vec{g}_{k-1}} \quad (\text{الف-۸})$$

این عبارت برابر است با ضرب داخلی تغییرات تکرار قبلی گرادیان با گرادیان جاری تقسیم بر مجذور اندازه بردار گرادیان تکرار قبلی. در نرم‌افزار MATLAB تابع Traincgp برای آموزش با این روش گرادیان مزدوج پیش‌بینی شده است. پارامترهای آموزشی این تابع مشابه تابع Traincgf است و عملکرد مشابهی دارند به طوری که پیش‌بینی آن که کدام یک از آن‌ها در یک مسأله خاص بهتر عمل می‌کند، مشکل است. حجم حافظه‌ای که Traincgp نیاز دارد تا حدودی از Traincgf بیشتر است [۲۱].

## الف-۵-۳- الگوریتم شروع پاول-بیل<sup>۲</sup>

در یک الگوریتم گرادیان مزدوج، به طور دوره‌ای جهت جستجو به همان جهت اولیه منفی گرادیان برگشت داده می‌شود. نقطه استاندارد برگشت وقتی است که تعداد تکرارها برابر با تعداد پارامترهای شبکه شود، ولی معیارهای برگشت دیگری وجود دارد که کارایی آموزش را بهبود می‌دهند. یکی از این

---

1. Polak - Ribiere

2. Powell - Beale

معیارها توسط پاول و بیل پیشنهاد شده است که بر اساس آن وقتی که نامساوی زیر برقرار شود، جهت جستجو به منفی گرادیان تغییر می‌یابد.

$$|\vec{g}_{k-1}^T \vec{g}_k| \geq 0.2 \|\vec{g}_k\|^2 \quad (\text{الف-۹})$$

در مطلب تابع Traincgb برای آموزش به این روش پیش‌بینی شده است که پارامترهای آموزشی آن مشابه تابع Traincgf است. برای برخی از مسائل، عملکرد تابع Traincgb تا اندازه‌ای از تابع Traincgp بهتر است، اگر چه قضاوت نهایی به مسأله مورد نظر بستگی دارد. حجم حافظه مورد نیاز تابع Traincgb تا اندازه‌ای از Traincgp بیشتر است [۲۱].

#### الف-۵-۴- الگوریتم گرادیان مزدوج مقیاسی<sup>۱</sup>

الگوریتم‌های گرادیان مزدوجی که تا این جا شرح داده شد، در هر تکرار، به یک جستجوی خطی نیاز دارند. این جستجو از نظر محاسباتی سنگین است زیرا لازم است که در هر جستجو پاسخ شبکه به همه ورودی‌های آموزشی چندین بار محاسبه شود. گرادیان مزدوج مقیاسی طراحی شده است تا از جستجوی خطی که وقت‌گیر است، اجتناب شود. توضیح این الگوریتم پیچیده در چند سطر مشکل است، ولی می‌توان گفت که این الگوریتم ترکیبی است از رهیافتی که در بخشی از روش لونبرگ-مارکوارت به کار می‌رود با رهیافت گرادیان مزدوج. در MATLAB تابع Trainscg برای آموزش به این روش پیش‌بینی شده است. این تابع نسبت به سایر توابع گرادیان مزدوج به تکرارهای بیشتری نیاز دارد تا همگرا شود، ولی تعداد محاسباتی که در هر تکرار نیاز است به میزان قابل توجهی کاهش می‌یابد، زیرا جستجوی خطی انجام نمی‌شود. حجم حافظه مورد نیاز این تابع مشابه تابع Traincgf است [۲۱].

---

1. Scaled Conjugate Gradient



## الف-۶- الگوریتم‌های شبه نیوتنی<sup>۱</sup>

### الف-۶-۱- الگوریتم بی‌اف‌جی‌اس<sup>۲</sup>

با هدف بهینه‌سازی سریع، روش نیوتن می‌تواند به عنوان جایگزینی برای روش‌های گرادیان مزدوج باشد. تغییر پارامترها در این الگوریتم مطابق رابطه زیر انجام می‌شود:

$$\vec{x}_{k+1} = \vec{x}_k - H_k^{-1} \vec{g}_k \quad (\text{الف-۱۰})$$

که در آن  $H_k$  ماتریس هسین تابع عملکرد به ازای مقادیر جاری پارامترهای شبکه است. اغلب روش نیوتن سریع‌تر از روش‌های گرادیان مزدوج همگرا می‌شود. متأسفانه محاسبه ماتریس هسین برای شبکه‌های عصبی پیش‌رو پیچیده و مشکل است. گروهی از الگوریتم‌ها وجود دارند که بر اساس روش نیوتن عمل می‌کنند اما به محاسبه مشتقات دوم نیازی ندارند و به آن‌ها روش‌های شبه‌نیوتنی یا متقاطع (Secant) گفته می‌شود. این روش‌ها در هر تکرار الگوریتم، تخمین ماتریس هسین مرحله قبل را تصحیح می‌کنند و از تخمین جدید برای محاسبه پارامترهای شبکه استفاده می‌کنند. روش شبه‌نیوتنی که در مطالعات منتشر شده موفق‌ترین نتایج را نشان داده است، روش تصحیح BFGS بوده است. در MATLAB تابع آموزشی Trainbfg به منظور آموزش شبکه‌ها با این الگوریتم پیش-بینی شده است. پارامترهای آموزشی این تابع همان پارامترهای تابع Traincgf است. این الگوریتم نسبت به روش‌های گرادیان مزدوج در هر تکرار به محاسبات و حافظه بیشتری نیاز دارد، زیرا ماتریس هسین که برای شبکه‌ای با  $N$  پارامتر یک ماتریس  $N \times N$  است، باید تخمین و ذخیره شود ولی به طور کلی در تکرارهای کمتری همگرا می‌شود. ممکن است برای شبکه‌های خیلی بزرگ که پارامترهای زیادی دارند، استفاده از روش‌های گرادیان مزدوج و یا پسانتشار خطای بهبودپذیر بهتر باشد ولی برای شبکه‌های کوچک‌تر Trainbfg یک تابع کارآمد است [۲۱].

---

1. Quasi – Newton Algorithm

2. BFGS Algorithm

## الف-۶-۲- الگوریتم متقاطع یک مرحله‌ای<sup>۱</sup>

چون الگوریتم BFGS نسبت به الگوریتم‌های گرادیان مزدوج در هر تکرار به محاسبات و حافظه بیشتری نیازمند است، بنابراین مطلوب است که تخمین به گونه‌ای انجام شود که این محاسبات و حجم حافظه کاهش یابد. روش متقاطع یک مرحله‌ای، روشی حد واسط الگوریتم‌های گرادیان مزدوج و الگوریتم‌های شبه نیوتنی است. در این الگوریتم فرض می‌شود که ماتریس هسین تکرار قبلی ماتریس واحد باشد. این الگوریتم نسبت به الگوریتم BFGS در هر دوره به محاسبات و حافظه کمتری نیاز دارد ولی در مقایسه با الگوریتم‌های گرادیان مزدوج اندکی به حافظه و محاسبات بیشتر نیازمند است. در MATLAB تابع آموزشی Trainoss برای آموزش با این الگوریتم پیش‌بینی شده است. پارامترهای آموزشی این تابع همان پارامترهای تابع Traincgf است [۲۱].

## الف-۷- الگوریتم لونبرگ-مارکوت<sup>۲</sup>

این روش در سال ۱۹۹۴ ارائه شد و تا به امروز به عنوان سریع‌ترین روش آموزش شبکه‌های عصبی برای شبکه‌های پیش‌رو با اندازه متوسط ( شبکه‌هایی با چند صد پارامتر) شناخته شده است. این روش ده‌ها برابر از روش گرادیان مزدوج و صدها برابر از روش بیشترین شیب سریع‌تر است. از ویژگی‌های برجسته این روش، کاهش شدید در نرون‌های لایه مخفی می‌باشد و از محدودیت‌های آن این است که تنها با روند یادگیری گروهی قابل استفاده است و هیچ تضمینی ندارد که با آموزش الگو به الگو همگرا شود [۲۱].

وقتی که تابع عملکرد به شکل مجموع مربعات (مثلاً تابع MSE) انتخاب شود، می‌توان ماتریس هسین را به صورت زیر تخمین زد:

---

1. One Step Secant Algorithm

2. Levenberg - Marquardt

$$H = J' J \quad (\text{الف-۱۱})$$

همچنین گرادیان به صورت زیر محاسبه می‌شود:

$$\vec{g} = J' \vec{e} \quad (\text{الف-۱۲})$$

$J$  ماتریس ژاکوبین است که مشتقات اول خطاهای تخمین شبکه نسبت به پارامترهای شبکه می‌باشد و  $\vec{e}$  یک بردار از خطاهای شبکه است. ماتریس ژاکوبین با استفاده از یک روش پسانتشار خطا با پیچیدگی بسیار کمتر نسبت به ماتریس هسین قابل محاسبه است. در الگوریتم لونبرگ-مارکوارت تغییر پارامترها با استفاده از رابطه زیر انجام می‌شود:

$$\vec{x}_{k+1} = \vec{x}_k - [J' J + \mu I]^{-1} J' \vec{e} \quad (\text{الف-۱۳})$$

وقتی که مقدار کمیت اسکالر  $\mu$  صفر در نظر گرفته شود، الگوریتم لونبرگ-مارکوارت به الگوریتم نیوتنی تبدیل می‌شود که از تخمین ماتریس هسین استفاده کرده است، ولی وقتی که مقدار این کمیت بزرگ است، به روش بیشترین شیب با سرعت یادگیری کوچک تبدیل می‌شود. در نزدیکی یک کمینه خطا، روش نیوتن سریع‌تر و دقیق‌تر است، بنابراین در چنین موقعیتی هدف آن است که تا آن جا که ممکن است به سرعت به روش نیوتنی نزدیک شویم. بنابراین بعد از هر مرحله موفقیت‌آمیز کاهش تابع عملکرد،  $\mu$  کاهش می‌یابد و فقط بعد از هر افزایش تابع عملکرد،  $\mu$  افزایش می‌یابد. با این روش تابع عملکرد بعد از هر تکرار الگوریتم، همیشه کاهش می‌یابد. در MATLAB برای آموزش با الگوریتم لونبرگ - مارکوارت تابع Trainlm پیش‌بینی شده است. پارامترهای این تابع عبارتند از: epochs, show, goal, time, min\_grad, max\_fail, mu, mu\_dec, mu\_inc, mu\_max, mem\_reduc. شش پارامتر اول قبلاً توضیح داده شده‌اند. mu مقدار اولیه  $\mu$  است. هر گاه تابع عملکرد کاهش یابد، mu در پارامتر mu\_dec ضرب می‌شود و هر گاه افزایش یابد، mu در mu\_inc

ضرب می‌شود. اگر  $\mu$  از  $\mu_{max}$  بزرگ‌تر شود، الگوریتم متوقف می‌شود. پارامتر  $mem\_reduc$  برای کنترل میزان حافظه‌ای که الگوریتم مورد استفاده قرار می‌دهد، به کار می‌رود.

باید توجه داشت که اگر در آموزش شبکه‌ها با روش توقف سریع از یک الگوریتم سریع استفاده می‌شود، بایستی پارامترهای آموزش به گونه‌ای تعیین شوند که الگوریتم نسبتاً آهسته همگرا شود، مثلاً در صورت استفاده از الگوریتم لونبرگ-مارکوارت،  $\mu$  نسبتاً بزرگ، مثلاً ۱ و  $\mu_{inc}$  و  $\mu_{dec}$  نیز نزدیک به ۱، مثلاً  $1/5$  و  $0/8$  انتخاب شوند [۲۱].

## الف-۸- تابع مرتب‌سازی بیزین (Trainbr)

در آموزش شبکه‌ها به روش مرتب‌سازی، مطلوب است که پارامترهای مرتب‌سازی بهینه به طور اتوماتیک تعیین شود. بنیان بیزین مک‌کی<sup>۱</sup> رهیافتی برای این منظور است که در آن، پارامترهای شبکه متغیرهایی تصادفی با توزیع‌های معین در نظر گرفته می‌شوند. پارامترهای مرتب‌سازی با واریانس‌های نامعلوم وابسته به این توزیع‌ها مربوط هستند. این پارامترها با استفاده از روش‌های آماری تعیین می‌شوند. شرح مفصل این روش، از سطح این نوشتار خارج است.

در MATLAB تابع آموزشی Trainbr برای آموزش با این روش پیش‌بینی شده است. یکی از ویژگی‌های این روش آن است که تعداد پارامترهایی از شبکه را که به طور مؤثر توسط شبکه استفاده می‌شوند، مشخص می‌کند. با فرض اطمینان از همگرایی، صرف‌نظر از این که تعداد کل پارامترهای شبکه چقدر است، تعداد مؤثر پارامترها بایستی تقریباً ثابت باشد.

---

1. Bayesian Framework of David Mackay

به هنگام استفاده از تابع Trainbr، باید اجازه دهیم تا تعداد مؤثر پارامترها همگرا شود. همچنین اگر مجموع مربعات خطا (SSE) یا مجموع مربعات پارامترهای شبکه (SSW) در چندین تکرار به مقداری نسبتاً ثابت برسد، می‌توان گفت که الگوریتم همگرا شده است [۲۱].

**پیوست ب:**

**روش های مختلف بهینه سازی**

## ب-۱- روش شبیه سازی تبرید تدریجی<sup>۱</sup>

SA مخفف Simulated Annealing به معنای شبیه سازی گداخت یا شبیه سازی حرارتی می باشد که برای آن از عبارات شبیه سازی بازپخت فلزات، شبیه سازی آب دادن فولاد و الگوریتم تبرید نیز استفاده شده است. برخی مسائل بهینه سازی صنعتی در ابعاد واقعی غالباً پیچیده و بزرگ می باشند. بنابراین روش های حل سنتی و استاندارد، کارایی لازم را نداشته و عموماً مستلزم صرف زمان های محاسباتی طولانی هستند. خوشبختانه، با پیشرفت فن آوری کامپیوتر و ارتقا قابلیت های محاسباتی، امروزه استفاده از روش های ابتکاری و جستجوگرهای هوشمند کاملاً متداول گردیده است. یکی از این روش ها SA است. SA شباهت دارد با حرارت دادن جامدات. این ایده ابتدا توسط شخصی که در صنعت نشر فعالیت داشت به نام متروپولیس<sup>۲</sup> در سال ۱۹۵۳ بیان شد. وی تشبیه کرد کاغذ را به ماده ای که از سرد کردن مواد بعد از حرارت دادن آنها بدست می آید. اگر یک جامد را حرارت دهیم و دمای آن را به نقطه ذوب برسانیم سپس آن را سرد کنیم جزئیات ساختمانی آن به روش و نحوه سرد کردن آن وابسته می شود. اگر آن جامد را به آرامی سرد کنیم کریستال های بزرگی خواهیم داشت که می توانند آن طور که ما می خواهیم فرم بگیرند ولی اگر سریع سرد کنیم آنچه که می خواهیم بدست نمی آید.

الگوریتم متروپولیس شبیه سازی شده بود از فرآیند سرد شدن مواد به وسیله کاهش آهسته دمای سیستم (ماده) تا زمانی که به یک حالت ثابت منجمد تبدیل شود. این روش با ایجاد و ارزیابی جواب های متوالی به صورت گام به گام به سمت جواب بهینه حرکت می کند. برای حرکت، یک همسایگی جدید به صورت تصادفی ایجاد و ارزیابی می شود. در این روش به بررسی نقاط نزدیک نقطه داده شده در فضای جستجو می پردازیم. در صورتی که نقطه جدید، نقطه بهتری باشد (تابع هزینه را کاهش دهد) به عنوان نقطه جدید در فضای جستجو انتخاب می شود و اگر بدتر باشد (تابع هزینه را

<sup>۱</sup> Simulated Annealing

<sup>۲</sup> metropolis

افزایش دهد) براساس یک تابع احتمالی باز هم انتخاب می‌شود. به عبارت ساده‌تر، برای کمینه سازی تابع هزینه، جستجو همیشه در جهت کمتر شدن مقدار تابع هزینه صورت می‌گیرد، اما این امکان وجود دارد که گاه حرکت در جهت افزایش تابع هزینه باشد.

روش بهینه سازی SA به این ترتیب است که با شروع از یک جواب اولیه تصادفی برای متغیرهای تصمیم‌گیری، جواب جدید در مجاورت جواب قبلی با استفاده از یک ساختار همسایگی مناسب به طور تصادفی تولید می‌شود. بنابراین یکی از مسائل مهم در SA روش تولید همسایگی است. برای پیاده سازی الگوریتم شبیه سازی حرارتی به سه عامل اساسی به شرح زیر نیازمندیم :

۱. نقطه شروع:

نقطه‌ای در فضای جستجو است که جستجو را از آنجا آغاز می‌کنیم. این نقطه معمولاً به صورت تصادفی انتخاب می‌شود .

۲. مولد حرکت:

این مولد وظیفه تولید حالات بعدی را بعهده دارد و با توجه به محاسبه هزینه نقطه فعلی و هزینه نقطه بعدی، وضعیت حرکت الگوریتم را مشخص می‌کند .

۳. برنامه سرد کردن<sup>۱</sup>:

پارامترهایی که نحوه سرد کردن الگوریتم را مشخص می‌کنند. بدین ترتیب که دما چند وقت به چند وقت و به چه میزان کاهش یابد و دماهای شروع و پایان چقدر باشند. در سال ۱۹۸۲ کرک پاتریک<sup>۲</sup> ایده متروپولیس را برای حل مسائل به کار برد. در سال ۱۹۸۳ کرک پاتریک و تعدادی از همکارانش از SA برای حل مسئله فروشنده دوره‌گرد یا TSP استفاده کردند.

---

<sup>1</sup> Cooling schedule

<sup>2</sup> Kirk patrick



جهت بررسی شرط توقف از دو معیار استفاده می‌شود:

یک معیار رسیدن به درجه حرارت نهایی است. معیار دیگر بر مبنای نسبت میزان پراکندگی جواب‌های پذیرفته شده در درجه حرارت فعلی به تفاوت متوسط مقادیر تابع هدف جهت جواب‌های پذیرفته شده در درجه حرارت اولیه و درجه حرارت فعلی است. در صورتی که این نسبت کم باشد یعنی سیستم به حالت انجماد رسیده و متوقف می‌شود در غیر اینصورت با کاهش درجه حرارت، الگوریتم پیشنهادی ادامه پیدا می‌کند [۲۷].

## ب-۲- الگوریتم کلونی مورچه‌ها<sup>۱</sup>

الگوریتم کلونی مورچه‌ها برای اولین بار توسط دوریگو (Dorigo) و همکارانش به عنوان یک راه حل چند عامله (Multi Agent) برای مسائل مشکل بهینه‌سازی مثل فروشنده دوره‌گرد (TSP: Traveling Sales Person) ارائه شد.

عامل هوشمند (Intelligent Agent) موجودی است که از طریق حسگرها قادر به درک پیرامون خود بوده و از طریق تاثیر گذارنده‌ها می‌تواند روی محیط تاثیر بگذارد. الگوریتم کلونی مورچه الهام گرفته شده از مطالعات و مشاهدات روی کلونی مورچه‌هاست. این مطالعات نشان داده که مورچه‌ها حشراتی اجتماعی هستند که در کلونی‌ها زندگی می‌کنند و رفتار آنها بیشتر در جهت بقا کلونی است تا در جهت بقا یک جزء از آن. یکی از مهمترین و جالبترین رفتار مورچه‌ها، رفتار آنها برای یافتن غذا است و بویژه چگونگی پیدا کردن کوتاه‌ترین مسیر میان منابع غذایی و آشیانه. این نوع رفتار مورچه‌ها دارای نوعی هوشمندی توده‌ای است که اخیرا مورد توجه دانشمندان قرار گرفته است. باید تفاوت هوشمندی توده‌ای (کلونی) و هوشمندی اجتماعی را روشن کنیم.

---

<sup>1</sup> Ant Colony Optimization

در هوشمندی اجتماعی عناصر میزانی از هوشمندی را دارا هستند. بعنوان مثال در فرآیند ساخت ساختمان توسط انسان، زمانی که به یک کارگر گفته میشود تا یک توده آجر را جابجا کند، آنقدر هوشمند هست تا بداند برای اینکار باید از فرغون استفاده کند نه مثلاً بیل!!! نکته دیگر تفاوت سطح هوشمندی افراد این جامعه است. مثلاً هوشمندی لازم برای فرد معمار با یک کارگر ساده متفاوت است.

در هوشمندی توده‌ای عناصر رفتاری تصادفی دارند و بین آن‌ها هیچ نوع ارتباط مستقیمی وجود ندارد و آنها تنها بصورت غیر مستقیم و با استفاده از نشانه‌ها با یکدیگر در تماس هستند. مثالی در این مورد رفتار موربانه‌ها در لانه سازیست.

همانطور که می‌دانیم مسئله یافتن کوتاه‌ترین مسیر، یک مسئله بهینه سازی می‌باشد که گاه حل آن بسیار دشوار و نیز زمانبر است. بعنوان مثال می‌توان به فروشنده دوره گرد اشاره کرد. در این مسئله، فروشنده دوره گرد باید از یک شهر شروع کرده، به شهرهای دیگر برود و سپس به شهر مبدا بازگردد بطوریکه از هر شهر فقط یکبار عبور کند و کوتاهترین مسیر را نیز طی کرده باشد. اگر تعداد این شهرها  $n$  باشد در حالت کلی این مسئله از مرتبه  $(n-1)!$  است که برای ۲۱ شهر زمان واقعا زیادی لازم است [۲۷].

از کاربردهای ACO می‌توان به بهینه کردن هر مسئله‌ای که نیاز به یافتن کوتاه‌ترین مسیر دارد، اشاره نمود:

۱. مسیر یابی داخل شهری و بین شهری

۲. مسیر یابی بین پست‌های شبکه‌های توزیع برق ولتاژ بالا

۳. مسیر یابی شبکه‌های کامپیوتری

### ب-۳- الگوریتم جستجوی تابو<sup>۱</sup>

منشا نام این روش را می‌توان از این نظریه یافت، که یک جستجوی محلی می‌تواند با یک بهینگی محلی منتهی شود، با این اطمینان که به طور دوره‌ای به یک بهینگی محلی یکسان بازنگردد، که نیازمند این است که از بازگشت جستجو به راه‌حل‌هایی که تاکنون بررسی شده‌اند جلوگیری شود، به عبارت دیگر، برخی راه‌حل‌ها باید دارای یک وضعیت ممنوعه باشند.

جستجوی تابو، در مقایسه با شبیه‌سازی حرارتی و الگوریتم ژنتیک، فضای راه‌حل را خیلی بیشتر جستجو می‌کند (یعنی حریص‌تر از آن‌هاست). الگوریتم‌های جستجوی تابو با یک پیکربندی (یا مجموعه‌ای از پیکربندی‌ها در زمانی که جستجو به شکل هم‌روند انجام می‌شوند) مقدار دهی اولیه می‌گردد، که پیکربندی جاری نامیده می‌شود. در هر دور تکرار الگوریتم، یک ساختار همسایگی برای پیکربندی جاری تعریف می‌شود؛ سپس یک حرکت انجام می‌شود تا به سوی بهترین پیکربندی در این همسایگی حرکت کند (یعنی در یک مسئله‌ی کمینه‌سازی، الگوریتم راهش را به سوی پیکربندی‌ای جهت می‌دهد که گویای کم‌ترین هزینه است). در حالت عادی، تنها همسایگان با امیدبخشی بیشتر مد نظر قرار می‌گیرند، در غیر این صورت ممکن است نتوان مسئله را به راه درستش هدایت کرد (مسئله‌ی رام نشدنی). بر خلاف انواع الگوریتم‌های حساس به تغییر (الگوریتم‌های گرادیانی)، که برای جستجوی محلی استفاده می‌شوند، در جستجوی تابو همسایگی به شکل پویا (دینامیک) به‌روزرسانی می‌گردد. تفاوت دیگر این است که انتقال به پیکربندی‌های با هزینه‌ی بالاتر (حالت نامناسب برای مسئله) مجاز است (این ویژگی روش را قادر می‌سازد تا از نقطه‌ی کمینگی محلی رهایی یابد). یک ویژگی ضروری الگوریتم‌های جستجوی تابو خارج کردن مستقیم گزینه‌های جستجویی است که به‌طور موقت در دسته‌ی مسیرهای ممنوع (تابو) قرار گرفته‌اند. نتیجه این‌که، در این الگوریتم‌ها استفاده از حافظه به گونه‌ای بسیار شدید صورت می‌گیرد که یکی از محدودیت‌های تابو است.

---

<sup>۱</sup> Tabu Search

ویژگی‌های جستجوی تابو این‌گونه خلاصه شده‌اند:

۱. حافظه‌ی سازگار شونده (شکل پویای حافظه)
  ۲. به‌گزینی (که دارای استراتژی فراموشی است)
  ۳. ساده‌سازی و تجزیه (در طول حافظه‌ی واضح و با دسترسی مستقیم)
  ۴. تنظیم زمان (یعنی هم تاخیر و هم تعداد تکرار وقایع و تفاوت میان کوتاه‌مدت و بلندمدت)
  ۵. کیفیت و فشردگی (یعنی قدرت کشش نسبی انتخاب‌های موجود و بزرگی تغییرات در ساختار یا روابط بازدارنده)
  ۶. سابقه (شامل سابقه‌ی ناحیه‌ای، ساختاری و وابستگی‌های متقابل ترتیبی)
  ۷. جستجوی قابل پیش‌بینی
  ۸. محدودیت‌ها و وسیله‌های تحمیل شده با توجه به موقعیت (یا، شرایط ممنوعه و سطوح پیش‌ران)
  ۹. تمرکز فراوان بر ناحیه‌ها و راه‌حل‌های خوب (فرایند تشدید)
  ۱۰. مشخص کردن و کشف ناحیه‌های امیدبخش تازه (فرایند تنوع)
- می‌توان الگوریتم‌های جستجوی تابوی گوناگونی را با ترکیب کارکردهای بالا برای حل مسایل مشخص تشکیل داد. البته، روش پیاده‌سازی که مورد استفاده قرار می‌گیرد به ویژگی‌های مسئله و درجه‌ی پیچیدگی مورد نیاز در یک کاربرد خاص بستگی دارد. اگرچه کارکردهای گفته شده در بالا می‌توانند تعمیم یا تغییر داده شوند، اما در بسیاری از مسایل با زیر مجموعه‌ای از این کارکردها موفقیت حاصل شده است [۲۷].

## ب-۴- الگوریتم جستجوی هارومونی<sup>۱</sup>

گیم<sup>۲</sup> (۲۰۰۲) و همکاران یک الگوریتم بهینه سازی به نام جستجوی هارومونی ابداع کرده‌اند. این الگوریتم از الگوریتم‌های فوق خلاقانه (Meta-heuristic) محسوب می‌شود. الگوریتم‌های فوق خلاقانه بر مبنای اطلاعات به دست آمده از پدیده‌های طبیعی بنا شده‌اند.

الگوریتم جستجوی هارومونی از فرآیند موسیقی برای جستجوی حالت متوازن مانند فی‌البداهه سازی جاز برگرفته شده است. فی‌البداهه سازی جاز، بهترین حالت توازن را مطابق با ظرافت طبع جستجو می‌کند، درست همان‌طور که الگوریتم بهینه سازی، بهترین حالت را با ارزیابی تابع هدف جستجو می‌کند. مراحل این الگوریتم به شرح زیر است [۲۷]:

۱. تنظیم پارامترها: در این مرحله پارامترهایی از قبیل نرخ در نظرگیری حافظه هارومونی، آهنگ تنظیم گام، تعداد کل تکرارهای الگوریتم و یا شرط توقف.
۲. تولید تصادفی حافظه هارومونی: در این مرحله حافظه هارومونی با بردارهای جواب تصادفی پر می‌شود. بردارهای جواب، ترکیبی از حالت‌های مختلف طراحی می‌باشند.
۳. تولید یک هارومونی جدید: یک بردار هارومونی جدید بر اساس سه قاعده تولید می‌شود: ۱- بررسی حافظه ۲- تعدیل درجه ۳- انتخاب تصادفی.
۴. به روز کردن حافظه هارومونی: اگر بردار هارومونی جدید بهتر از بدترین بردار هارومونی باشد هارومونی جدید انتخاب و بدترین حذف می‌شود.
۵. چک کردن شرط توقف.

---

<sup>۱</sup> Harmonic Search

<sup>۲</sup> Geem

## ب-۵- الگوریتم اتوماتیک سلولی<sup>۱</sup>

الگوریتم اتوماتیک سلولی (CA) در اواخر دهه ۱۹۴۰ توسط ون نیومن مطرح و پس از او توسط ریاضی دانی به نام اولام به عنوان مدلی برای بررسی رفتار سیستم‌های پیچیده پیشنهاد شد. CA در حقیقت سیستم‌های دینامیکی گسسته ای هستند که رفتارشان کاملاً بر اساس رفتار محلی استوار است. در CA فضا به صورت یک شبکه تعریف می‌گردد که به هر خانه آن یک سلول گفته می‌شود. زمان به صورت گسسته پیش می‌رود و قوانین آن به صورت سراسری است که از طریق آن در هر مرحله، هر سلول وضعیت جدید خود را با در نظر گرفتن همسایه‌های مجاور خود به دست می‌آورد. قوانین CA نحوه تأثیر پذیرفتن سلول از سلول‌های همسایه خود را مشخص می‌کند. یک سلول را همسایه سلول دیگر گوییم هرگاه بتواند آن را در یک مرحله و بر اساس قانون حاکم تحت تأثیر قرار دهد. در بدست آوردن وضعیت کنونی سلول علاوه بر وضعیت قبلی سلول‌های همسایه، می‌توان وضعیت قبلی خود سلول را نیز دخالت داد.

در مدل‌سازی سیستم‌های فیزیکی و بیولوژیکی، گاهی لازم است که قوانین را به صورت احتمالی در نظر بگیریم. رفتار احتمالی را می‌توان به عنوان نویز در سیستم تعبیر نمود. یکی از اشکالات CA تعیین فرم قطعی قوانین مورد نیاز برای یک کاربرد خاص است. راه‌حل‌های متفاوتی در برخورد با این مشکلات به نظر می‌رسد. یکی از این راه‌حل‌ها احتمالاتی کردن قوانین است. به این ترتیب، تمام قانون‌های امکان‌پذیر را در نظر می‌گیریم و برای فعال شدن آن‌ها احتمالی اختصاص می‌دهیم.

CA یادگیر ماشینی است که می‌تواند تعدادی متناهی عمل را انجام دهد. هر عمل انتخاب شده توسط یک محیط احتمالی ارزیابی می‌شود و نتیجه ارزیابی در قالب پاسخ مثبت یا منفی به CA داده می‌شود و CA از این پاسخ در انتخاب عمل بعدی تأثیر می‌گیرد. هدف نهایی این است که CA یاد بگیرد تا از

---

<sup>۱</sup> Automated cellular

بین اعمال خود بهترین عمل را انتخاب کند. بهترین عمل، عملی است که احتمال دریافت پاداش از محیط را به حداکثر برساند [۲۷].

## ب-۶- الگوریتم پرندگان<sup>۱</sup>

الگوریتم PSO یک الگوریتم جستجوی اجتماعی است که از روی رفتار اجتماعی دسته‌های پرندگان مدل شده است. در ابتدا این الگوریتم به منظور کشف الگوهای حاکم بر پرواز همزمان پرندگان و تغییر ناگهانی مسیر آن‌ها و تغییر شکل بهینه‌ی دسته به کار گرفته شد. در PSO ذرات در فضای جستجو جاری می‌شوند. تغییر مکان این ذرات در فضای جستجو تحت تأثیر تجربه و دانش خودشان و همسایگان آن‌ها است بنابراین موقعیت دیگر ذرات روی چگونگی جستجوی یک ذره اثر می‌گذارد. نتیجه مدل سازی این رفتار اجتماعی فرآیند جستجویی است که ذرات به سمت نواحی موفق میل می‌کنند. ذرات از یکدیگر می‌آموزند و بر مبنای دانش بدست آمده به سمت بهترین همسایگان خود می‌روند. اساس کار PSO بر این اصل استوار است که در هر لحظه هر particle مکان خود را در فضای جستجو با توجه به بهترین مکانی که تاکنون در آن قرار گرفته است و بهترین مکانی که در کل همسایگی‌اش وجود دارد، تنظیم می‌کند. از آنجا که دانش Swarm intelligence بسیار جدید است در حال حاضر کاربردهای کمی از آن شناخته شده است. ولی صاحب‌نظران معتقدند با این رشد روزافزون، Swarm intelligence می‌تواند نقش مهمی را در علوم مختلف مهندسی ایفا کند.

هر ذره برای اعمال تغییری مناسب در مکان و سرعت خود اطلاعات زیر را دارا می‌باشد:

"بهترین عمومی" که برای همه شناخته شده است و هنگامی که هر ذره بهترین مکان جدیدی را شناسایی کند، فوراً برای بقیه ذرات اطلاعات مربوطه را به روز رسانی می‌کند.

<sup>۱</sup> Particle Swarm Optimization(PSO)

بهترین همسایگی "که ذره از طریق ارتباط با زیر مجموعه های گروه ، آنرا بدست می آورد.

بهترین محلی "که بهترین راه حلی است که ذره تا کنون تجربه کرده است.

همه ذرات شروع به تاثیر پذیری از "بهترین عمومی" می کنند تا سرانجام به آن نزدیک شوند . ذرات در فضای جستجو در نزدیکی "بهترین عمومی" سیر می کنند و بقیه فضا را کاوش نمی کنند ، به این پدیده "همگرایی" گفته می شود. اگر ضریب اینرسی سرعت را کوچک انتخاب کنیم، تمام ذرات می توانند سرعتشان را کاهش دهند تا اینکه در "بهترین عمومی" به سرعت صفر نزدیکتر شوند. یک راه خروج از وضعیت همگرایی اولیه (نامطلوب) این است که دوباره به موقعیت ذرات (پس از رخ دادن همگرایی) مقدار اولیه بدهیم [۲۷].





**Shahrood University of Technology**  
**School of mining, Petroleum and Geophysics**

**Optimization of the Results of Committee Neural Network (Committee  
Machine = CM) in Prediction of Water Saturation Using Genetic  
Algorithm Method**

A thesis submitted in fulfillment of requirements for the degree of Master  
of Science in Geophysics (major Geoelectrics)

**Maki loveimi**

Supervisors:

**Dr. A. Kamkar-Rouhani**

**Dr. N. Keshavarz-Farajkhah**

**July 2010**

## **Abstract**

Water saturation is one of the most important and applied parameters in study of hydrocarbon reservoirs, which defined as percentage of empty spaces filled by water. In this study, we've used Committee neural networks to estimate the water saturation parameter in the study.

To this end we got benefit the data obtained from 5 wells locating in a southwestern field of Iran. The values of numerical sonic log, density, Gamma ray, resistivity, long normal resistivity, and effective porosity (which obtained by various logs and cores) are considered as the input, whereas water saturation is considered as the output. This method is more drastic and precise than using a single network method. The elements of the Committee neural networks are multilayer perceptron networks. The best structure of them was selected by trial and error method regarding the least test MSE. 59 networks with various structures have been trained and tested. The training for these 59 networks has been done by over-training, regularization and early stopping. Among the 59 networks, 7 networks with least MSE were selected for construction of ensemble combinations and 120 possible ensemble i.e. 21 two-network combination, 35 three-network combination, 35 four-network combination, 21 five-network combination, 7 six-network combination, 1 seven-network combination was constructed. Weight coefficients of the linear combinations have been obtained by 2 methods: Genetic Algorithm and Simple Averaging. Then the results of the combinations have been compared with respect to each other.

For over-training method, the single layer network with the structure of (6-10-1) has been known as the best. When the weight coefficients of linear combination of networks (6-10-1), (6-6-5-1), (6-13-1), (6-11-1) and (6-4-9-1) is obtained by using the genetic Algorithm, the most decrease in test MSE is observed in combination of these networks.

In regularization method, a network with the structure (6-10-2-1) has the least MSE in generalization stage. When the weight coefficients of linear combination of networks (6-10-2-1), (6-6-5-1), (6-5-6-1), (6-4-10-1) and (6-6-6-1) is obtained by using the genetic Algorithm, the most decrease in test MSE is observed in combination of these networks.

In early stopping method, the network with structure (6-8-8-1) When the weight coefficients of linear combination of networks (6-8-8-1), (6-8-9-1), (6-15-1), (6-6-9-1) and (6-7-10-1) is got by using the genetic Algorithm, the most decrease in test MSE is observed in combination of these networks.

Keywords: Water saturation, Artificial Neural Networks, Optimization, Ensemble Combinations, Genetic Algorithm.