

بسم الله الرحمن الرحيم



دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

پایان نامه کارشناسی ارشد

استفاده از الگوریتم یادگیری عمیق برای دسته‌بندی سیگنال‌های EEG

مینا سادات حسینی شמושکی

اساتید راهنما:

دکتر علی اکبر پویان

دکتر سعیده فردوسی

استاد مشاور:

دکتر هدی مشایخی

بهمن ماه ۱۳۹۴

دانشگاه صنعتی شاهرود

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

پایان نامه کارشناسی ارشد آقای / خانم

تحت عنوان:

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 :

تقدیم به پدر و مادر عزیز و مهربانم

که همواره در مسیر پرتلاطم زندگی چون کوهی استوار پشتیبان محکم و مطمئن برایم بودند و با ایثار و از خودگذشتگی و مهر و محبت بی دریغ سلامت

و امنیت و آرامش و آسایش را برای من فراهم آوردند.

تقدیم به خواهر و برادر عزیزم

که همراهان همیشگی و پشتوانه‌های زندگیم هستند.

بامشکر و قدرانی فراوان

از اساتید راهنمای فرهیخته ام جناب آقای دکتر پویان و سرکار خانم دکتر فردوسی

بخاطر سه صدر و رهنمودهای دلسوزانه که در تهیه این پایان نامه مرا مورد لطف قرار دادند و از هیچ کجی در این عرصه بر من دریغ ننمودند.

هم چنین سرکار خانم دکتر مشایخی که با قبول زحمت مشاوری مراد انجام این رساله یاری دادند.

در نهایت از زحمات پدر، مادر، خواهر، برادر و تمام کسانی که در انجام این تحقیق مرا از الطاف خود بی نصیب نگذاشته اند کمال تشکر را

دارم.

تعهد نامه

اینجانب **میناسوات حسینی شومشکی** دانشجوی دوره کارشناسی ارشد رشته **مهندسی کامپیوتر گرایش هوش مصنوعی** دانشکده **مهندسی کامپیوتر و**

فناوری اطلاعات دانشگاه صنعتی شاهرود نویسنده پایان نامه **استفاده از الگوریتم یادگیری عمیق برای دستبندی سیگنال های EEG** تحت

راهنمایی دکتر علی اکبر پویان و دکتر سعیده فردوسی متعهد می شوم .

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در بدست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

یادگیری عمیق یک شاخه نسبتاً جدید در یادگیری ماشین می‌باشد که در کاربردهای بسیاری نسبت به سایر الگوریتم‌های یادگیری نتایج بهتر و امیدوارکننده‌ای دارد. الگوریتم‌های مربوط به یادگیری عمیق به دلیل داشتن ساختار سلسه مراتبی، قابلیت بالایی در دسته‌بندی داده‌ها دارند. همچنین ساختارهای عمیق می‌توانند ارائه جامع‌تری از توابع در مقایسه با ساختارهای سطحی فراهم کنند. در حال حاضر یادگیری عمیق در زمینه‌های بسیاری از جمله بینایی ماشین، تشخیص صدا، پردازش زبان طبیعی و پردازش سیگنال کاربرد دارد. در این پایان‌نامه از یادگیری عمیق برای دسته‌بندی سیگنال‌های مغزی بهره گرفته می‌شود. حالت‌های احساسی در زمینه‌های مختلف زندگی روزمره انسان‌ها تأثیر می‌گذارند. توانایی کشف و تشخیص حالات احساسی فرد در تعاملات انسان و ماشین ضروری می‌باشد تا هرچه بیشتر این تعاملات به رابطه انسان با انسان نزدیک شوند. بدین منظور از الگوریتم یادگیری عمیق برای دسته‌بندی سیگنال EEG در کاربرد تشخیص احساسات مورد استفاده قرار گرفته است.

در روش پیشنهادی ابتدا تجزیه حالت تجربی (EMD) بر روی سیگنال EEG اعمال می‌شود. EMD سیگنال را به مجموعه‌ای از توابع حالت ذاتی (IMF) تجزیه می‌کند. انتخاب بهترین و مؤثرترین اجزای IMF توسط الگوریتم تجزیه مقدار منفرد (SVD) انجام می‌شود. به این منظور اجزای IMF متناظر با بیشترین مقادیر ویژه انتخاب شده، سپس بقیه اجزا حذف می‌شود. IMFها با استفاده از اجزا با بیشترین مقادیر ویژه بازسازی می‌شوند. بدست آوردن طیف IMF توسط تبدیل هیلبرت انجام می‌شود. در مرحله استخراج ویژگی از گشتاور طیفی که از طیف IMF استخراج می‌شود به عنوان ویژگی استفاده می‌شود. سپس تعداد ویژگی‌ها با روش مبتنی بر وابستگی آماری کاهش پیدا می‌کند. در نهایت از نوعی شبکه باور عمیق که واحد سازنده آن ماشین بولتزمن محدود می‌باشد به منظور دسته‌بندی سیگنال به احساسات مختلف بر اساس مدل‌های دوبعدی در فضای arousal و valence استفاده شده است.

در این تحقیق از پایگاه داده DEAP جهت ارزیابی عملکرد الگوریتم پیشنهادی استفاده شده است. روش پیشنهادی در مقایسه با روش‌های دیگر که از این پایگاه داده استفاده کرده‌اند، بهبود قابل توجهی داشته است. نتایج حاصل از دسته‌بندی هر کدام از معیارهای valence و arousal به سه دسته به ترتیب ۶۰/۵۲ و ۶۷/۳۶ درصد می‌باشد.

کلمات کلیدی: یادگیری عمیق، دسته‌بندی سیگنال EEG، تشخیص احساس، تجزیه حالت تجربی.

لیست مقالات مستخرج از پایان نامه

- Hosseini M, Pouyan A, Ferdosi S, Mashayekhi H. (1394), “EEG–Based Emotion Recognition using Deep Belief Network and Empirical Mode Decomposition”, Third national and first international conference on applied researches in Electrical, Mechanical and Mechatronic Engeeniring, Tehran, Iran.
- Hosseini M, Pouyan A, Ferdosi S, Mashayekhi H. (2016), “Emotion Recognition of EEG data using Deep Belief Network and Empirical Mode Decomposition”, submitted on journal of computational of Neuroscience.

فهرست مطالب

فصل ۱ مقدمه.....	۱
۱ - ۱ یادگیری عمیق.....	۲
۱ - ۲ انگیزه استفاده از یادگیری عمیق.....	۴
۱ - ۲ - ۱ انگیزه‌های بیولوژیکی.....	۴
۱ - ۲ - ۲ توان محاسباتی.....	۴
۱ - ۲ - ۳ توان پردازش حجم زیاد داده‌ها.....	۵
۱ - ۳ کاربردهای یادگیری عمیق در دنیای واقعی.....	۵
۱ - ۴ تعریف و هدف پایان‌نامه.....	۸
۱ - ۵ ساختار پایان‌نامه.....	۹
فصل ۲ یادگیری عمیق و احساسات.....	۱۱
۲ - ۱ مفاهیم یادگیری عمیق.....	۱۲
۲ - ۱ - ۱ شبکه عصبی مصنوعی.....	۱۲
۲ - ۱ - ۲ مدل‌های بر مبنای انرژی شبکه‌های عصبی.....	۱۴
۲ - ۱ - ۳ مدل‌های شرطی بر مبنای انرژی.....	۱۶
۲ - ۱ - ۴ نمونه‌برداری گیبز.....	۱۶
۲ - ۱ - ۵ ماشین بولتزمن محدود (RBM).....	۱۷
۲ - ۱ - ۶ همگرایی متقابل.....	۲۱
۲ - ۱ - ۷ شبکه باور عمیق.....	۲۴
۲ - ۱ - ۸ شبکه عصبی عمیق.....	۲۵
۲ - ۲ مقدمه‌ای بر احساسات.....	۲۸
۲ - ۲ - ۱ هوش احساسی و تعامل انسان و ماشین.....	۲۹
۲ - ۲ - ۲ محاسبه‌های احساسی.....	۳۰
۲ - ۲ - ۳ کاربردهای سیستم‌های بر پایه احساسات.....	۳۱
۲ - ۲ - ۴ تعریف احساسات.....	۳۴
۲ - ۲ - ۵ تعریف اجزا.....	۳۶
۲ - ۲ - ۶ مدل کردن احساسات.....	۳۶
۲ - ۲ - ۷ چالش‌های فنی.....	۴۰
فصل ۳ دانش پیش‌زمینه و کارهای گذشته.....	۴۵
۳ - ۱ مغز انسان.....	۴۶
۳ - ۱ - ۱ ساختار مغز و فعالیت‌های آن‌ها.....	۴۷

۴۹.....	۳ - ۱ - ۲ فیزیولوژی عصبی مغز انسان.....
۵۰.....	۳ - ۱ - ۳ الکتروانسفالوگرافی.....
۵۵.....	۳ - ۲ - ۲ مرور بر دسته‌بندی سیگنال EEG.....
۵۶.....	۳ - ۲ - ۱ تشخیص احساس با دسته‌بندی سیگنال EEG.....
۶۳.....	۳ - ۲ - ۲ الگوریتم یادگیری عمیق در دسته‌بندی احساسات.....
۶۵.....	فصل ۴ روش پیشنهادی.....
۶۶.....	۴ - ۱ معرفی پایگاه داده‌های موجود.....
۶۶.....	۴ - ۱ - ۱ IAPS.....
۶۸.....	۴ - ۱ - ۲ IADS.....
۶۹.....	۴ - ۱ - ۳ eINTERFACE 2005.....
۶۹.....	۴ - ۱ - ۴ MAHNOB-HCI.....
۷۰.....	۴ - ۱ - ۵ DEAP.....
۷۲.....	۴ - ۲ نحوه دسته‌بندی احساسات در روش پیشنهادی.....
۷۴.....	۴ - ۱ روش پیشنهادی.....
۷۶.....	۴ - ۲ استخراج ویژگی.....
۷۷.....	۴ - ۲ - ۱ تجزیه حالت تجربی (EMD).....
۸۲.....	۴ - ۲ - ۲ تجزیه مقدار منفرد (SVD):.....
۸۳.....	۴ - ۲ - ۳ گشتاور طیفی.....
۸۵.....	۴ - ۳ کاهش ویژگی.....
۸۶.....	۴ - ۳ - ۱ وابستگی آماری (SD) بین ویژگی‌ها و برچسب‌ها.....
۸۷.....	۴ - ۴ دسته‌بندی با شبکه عصبی عمیق.....
۹۱.....	فصل ۵ آزمایش‌ها و نتایج.....
۹۲.....	۵ - ۱ اعتبار سنجی متقابل.....
۹۳.....	۵ - ۲ نرخ صحت.....
۹۴.....	۵ - ۳ پیاده‌سازی روش پیشنهادی.....
۹۶.....	۵ - ۳ - ۱ تنظیم پارامترهای اولیه شبکه DNN.....
۹۸.....	۵ - ۴ نتایج روش پیشنهادی در دسته‌بندی ۳ کلاسه.....
۹۸.....	۵ - ۵ انتخاب کانال‌ها.....
۹۹.....	۵ - ۶ ارزیابی روش با اعمال روش‌های کاهش ویژگی مختلف.....
۱۰۲.....	۵ - ۷ ارزیابی روش با استفاده از ویژگی‌های مختلف.....
۱۰۶.....	۵ - ۸ ارزیابی روش با استفاده از دیگر دسته‌بندها.....
۱۰۹.....	۵ - ۹ مقایسه روش با کارهای دیگر انجام شده در DEAP.....
۱۱۱.....	فصل ۶ نتیجه‌گیری.....

۱۱۲.....	۶- ۱ کلیات پایان نامه.....
۱۱۳.....	۶- ۲ نتیجه گیری و پیشنهادها.....
۱۱۵.....	مراجع:.....

فهرست شکل‌ها

شکل ۱-۱. نمودار ویژگی‌های عمیق.....	۳
شکل ۱-۲. واحد محاسباتی، نورون.....	۱۳
شکل ۲-۲. ساختار سه لایه شبکه عصبی.....	۱۴
شکل ۳-۲. ماشین بولتزمن محدود که در آن h_i واحدهای مخفی و v_i واحدهای مشاهده‌پذیر هستند.....	۱۸
شکل ۴-۲. الگوریتم آموزش RBM.....	۲۳
شکل ۵-۲. شبکه باور عمیق (DBN).....	۲۵
شکل ۶-۲. الگوریتم یادگیری DBN.....	۲۸
شکل ۷-۲. نمونه‌ای از مدل ابعادی احساس.....	۳۹
شکل ۸-۲. طرح کلی یک مسئله یادگیری ماشین.....	۴۲
شکل ۱-۳. نواحی مختلف مغز [۳۵].....	۴۸
شکل ۲-۳. اولین داده‌های ضبط شده سیگنال EEG توسط هنس برگر [۳۸].....	۵۱
شکل ۳-۳. سیستم بین‌المللی قرارگیری الکتروود ۱۰-۲۰ [۴۰].....	۵۳
شکل ۱-۴. خود ارزیاب SAM برای حالات احساسی [۶۲].....	۷۲
شکل ۲-۴. خود ارزیاب SAM برای حالات احساسی [۶۲].....	۷۳
شکل ۳-۴. دسته‌های حالات احساسی.....	۷۳
شکل ۴-۴. روش پیشنهادی.....	۷۵
شکل ۵-۴. نمونه‌ای از نتیجه اعمال EMD روی سیگنال EEG ثبت شده در یک کانال.....	۸۰
شکل ۶-۴. DNN با سه لایه مخفی و دو دسته‌بند softmax.....	۸۸
شکل ۷-۴. خلاصه‌ای از فرآیند آموزش DNN.....	۹۰
شکل ۱-۵. تعیین تعداد بهینه ویژگی.....	۱۰۱
شکل ۲-۵. ارزیابی روش پیشنهادی با دسته‌بندهای مختلف.....	۱۰۸

فهرست جدول‌ها

جدول ۵-۱. دسته‌بندی سه کلاسه arousal و valence.....	۹۷
جدول ۵-۲. نتایج انتخاب کانال‌های متفاوت.....	۹۹
جدول ۵-۳. ارزیابی روش با اعمال روشهای کاهش ویژگی مختلف.....	۱۰۰
جدول ۵-۴. مقایسه ویژگی‌های مختلف برای دسته‌بندی.....	۱۰۵
جدول ۵-۱. نتایج مقایسه با سایر دسته‌بندها در دسته‌بندی valence.....	۱۰۶
جدول ۵-۶. نتایج مقایسه با سایر دسته‌بندها در دسته‌بندی arousal.....	۱۰۷
جدول ۵-۷. نتایج مقایسه روش پیشنهادی با سایر کارها.....	۱۰۹

فصل ۱

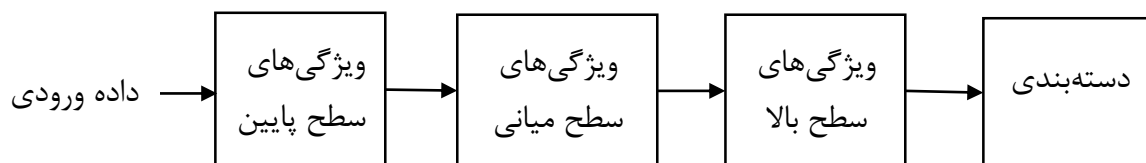
مقدمه

در سال‌های اخیر ما شاهد یک تجدید حیات در زمینه شبکه‌های عصبی بودیم که پایه و اساس آن نمایش داده‌ها از طریق چند لایه می‌باشد. این موضوع که به نام یادگیری عمیق شناخته می‌شود. با زمینه‌های مطالعاتی موجود مثل شبکه‌های عصبی، مدل‌های گرافیکی، مدل‌های مبتنی بر انرژی، استخراج ویژگی، بهینه‌سازی و تشخیص الگو مرتبط می‌باشد. از آنجایی که ساختار یادگیری عمیق شباهت بسیاری با نحوه درک مغز انسان دارد، با علم عصب‌شناسی نیز مرتبط می‌باشد. یادگیری عمیق در حال حاضر کاربردهای بسیاری در زمینه‌های بینایی ماشین، تشخیص صدا، پردازش زبان طبیعی، تشخیص دست‌نوشته و پردازش سیگنال دارد. یادگیری عمیق یکی از زمینه‌های پیشرو و امیدبخش در مسائل یادگیری ماشین به خصوص در حوزه شبکه‌های عصبی می‌باشد. در این فصل با یادگیری عمیق، انگیزه استفاده از آن و کاربردهای آن آشنا می‌شویم و در نهایت هدف و تعریف پایان‌نامه و ساختار آن بیان می‌شود.

۱ - ۱ یادگیری عمیق

دلایل بسیاری برای افزایش محبوبیت یادگیری عمیق وجود دارد که به پیشرفت‌های نرم‌افزار و سخت‌افزار مثل افزایش توانایی‌های پردازش واحد پردازش گرافیکی (GPU) و توسعه الگوریتم‌های پردازش اطلاعات جدید یادگیری ماشین مربوط می‌شود. یادگیری عمیق یک روش سلسله مراتبی برای ارائه و نمایش داده‌ها می‌باشد. ویژگی‌ها در لایه‌های مختلف و واحدهای مختلف هر لایه یکی پس از دیگری توزیع می‌شوند، که این امر سبب تمایز ساختارهای عمیق از دیگر ساختارها می‌شود [۱].

در شکل ۱-۱ می‌بینیم هرچه به لایه‌های بالایی نزدیک‌تر می‌شویم سطح انتزاع افزایش پیدا می‌کند. برای مثال در پردازش تصاویر، در پایین‌ترین سطح از پیکسل‌ها شروع می‌شود و در نهایت به شیء موردنظر می‌رسد. همین رویه در پردازش متن نیز وجود دارد که از کاراکترها به سطوح انتزاعی بالاتر مثل کلمات، گروهی از کلمات، عبارات، جملات و در نهایت به داستان کلی می‌رسد. در تشخیص صدا از نمونه‌های سیگنال، باندهای طیفی، صداها شروع می‌شود تا به آواها و کلمات برسد که برای انتزاع بیشتر به گروهی از کلمات و عبارات و جملات نیز می‌رسد [۲].



شکل ۱-۱. نمودار ویژگی‌های عمیق

یادگیری عمیق با این فرض شکل گرفته است که حقیقت^۱ یک موجودیت یکپارچه و تک‌جزئی نیست. حقیقت یک مفهوم توزیع شده است که در بین لایه‌ها و واحدها جریان دارد که در نهایت با هم ترکیب می‌شوند و در سطح ظاهری یک موجودیت کامل را شکل می‌دهند. یک شخص یا ماشین ممکن است که فرآیند یا بخشی از آن را بدون تلاش برای فهم ساختار و تعاملات بین اجزاء، متوجه نشود. گاهی اوقات این اجزای تنها هیچ معنی و مفهومی ندارند ولی در نهایت با ترکیب شدن یک موجودیت کامل را تشکیل می‌دهند.

^۱Truth

۱ - ۲ انگیزه استفاده از یادگیری عمیق

۱ - ۲ - ۱ انگیزه‌های بیولوژیکی

انگیزه اصلی استفاده از یادگیری عمیق شباهت عملکرد آن با مغز انسان است. قشر مغز انسان لایه‌ای و سلسله مراتبی است. بطور تخصصی تر قشر مغز تقریباً شش لایه دارد که لایه‌های پایینی به لایه‌هایی بالایی و بالعکس نگاشت می‌شوند [۳]. ارائه مفاهیم انتزاعی تر در لایه‌های بالایی ماهیت سلسله مراتبی از مشاهدات را بوجود می‌آورد.

یادگیری عمیق باور دارد که این ساختار سلسله‌مراتبی که توسط قشر مغز بکار گرفته می‌شود، دلیل اصلی قدرت بالای مغز می‌باشد و تلاش می‌کند این قدرت و توان بالا را با استفاده از ساختارهای سلسله مراتبی و لایه‌ای بدست آورد.

۱ - ۲ - ۲ توان محاسباتی

یک انگیزه مهم برای مطالعه یادگیری عمیق توان بالای محاسباتی آن می‌باشد. در برخی موارد، محاسباتی که با k لایه بدست می‌آیند کارآمدتر از محاسباتی هستند که با $k-1$ لایه بدست می‌آیند [۳]. لایه‌ها به تعداد طولانی‌ترین مراحل محاسباتی برای نتیجه گرفتن از ورودی اشاره دارند. در این مورد منظور از کارآمدتر بودن تعداد عناصری است که برای انجام محاسبات مورد نیاز است. در شبکه‌های عصبی عنصر محاسباتی نورون می‌باشد و در یک مدار منطقی المان‌های AND، OR، XOR و ... می‌باشند.

¹Neocortex

کارایی محاسباتی در مسائل یادگیری ماشین مهم می‌باشد. از آنجایی که پارامترهای عناصر محاسباتی در طول آموزش یاد گرفته می‌شوند و سایر مجموعه آموزشی معمولاً ثابت می‌باشد، عناصر محاسباتی اضافی به پارامترهای بیشتری برای هر نمونه آموزشی نیاز دارند که به عملکرد بدتری منجر می‌شود. بعلاوه با اضافه کردن پارامترهای اضافی همیشه خطر بیش‌برازش^۱ و یا عمومیت بخشیدن^۲ ضعیف وجود دارد.

۱ - ۲ - ۳ توان پردازش حجم زیاد داده‌ها

طبق تحقیقات شرکت IBM، در هر روز ۲,۵ اگزا بایت اطلاعات در جهان تولید می‌شود و شرکت IDC تخمین زده است که تا سال ۲۰۲۰ تقریباً ۴۰ زتا بایت اطلاعات در هر روز تولید خواهد شد. با این حجم زیاد اطلاعات کار پردازش و پیدا کردن اطلاعات نهان با الگوریتم‌های کم‌عمق متداول دشوار خواهد بود. سیستم‌های بر مبنای یادگیری عمیق امکان کشف و فهم داده‌های بدون ساختار را برای استخراج اطلاعات در جهت تحلیل داده‌ها را فراهم می‌کنند.

۱ - ۳ کاربردهای یادگیری عمیق در دنیای واقعی

یادگیری عمیق کاربردهای بسیاری در زمینه‌های مختلف دارد. در سال‌های اخیر تحقیقات و مطالعات زیادی در مراکز تحقیقاتی و شرکت‌های بزرگ برای کارهای تجاری و ساخت برنامه‌های کاربردی خلاقانه در این

^۱Overfitting

^۲Generalization

زمینه شروع شده و ادامه دارد. در این بخش به برخی از کاربردهای یادگیری عمیق در دنیای واقعی اشاره شده است.

- تبلیغات:

یکی از کاربردهای یادگیری عمیق در گستره وب بوجود آوردن تبلیغات بر پایه اطلاعات و مزایده‌های محصولات می‌باشد. برای مثال شرکت تبلیغاتی ادتئورنت^۱ کارهای مربوط به مزایده و تبلیغات محصولات در گستره وب را با بهره‌گیری از یادگیری عمیق انجام می‌دهد. این شرکت با استفاده از روش یادگیری عمیق ۲۰٪ افزایش در نرخ کلیک داشته است. استخراج کلمات کلیدی و آنالیز احساسات مبنای کار یادگیری عمیق برای دستیابی به این افزایش بود.

- موتورهای جست‌وجو

این کاربرد در قالب مثال بیان می‌شود. حالتی را در نظر بگیرید که فرد قصد تماشای فیلمی را دارد بدون اینکه اطلاعات و نظری در رابطه با انتخابش داشته باشد. در این مواقع شخص می‌تواند از موتورهای جست‌وجوی پیشنهاد کننده استفاده کند. نتفلیکس^۲ یک موتور جست‌جوی معروف و شناخته‌شده در این زمینه می‌باشد. دو سوم فیلم‌هایی که کاربران در نتفلیکس تماشا می‌کنند با استفاده از یادگیری عمیق پیشنهاد شده است.

- تشخیص چهره

یکی از کاربردهای مهم یادگیری عمیق در زمینه تشخیص چهره می‌باشد. در دنیای واقعی می‌توان به

^۱Adtheorent: www.adtheorent.com

^۲Netflix: www.netflix.com

استفاده از یادگیری عمیق در شبکه‌های اجتماعی مثل فیس‌بوک^۱ اشاره کرد.

مثالی دیگر در این زمینه نرم‌افزار فیس فرست^۲ می‌باشد. این نرم‌افزار چهره افراد را شناسایی می‌کند و می‌تواند برای افزایش امنیت در فروشگاه‌ها و فرودگاه‌ها استفاده شود. این نرم‌افزار با بکارگیری یادگیری عمیق چهره افراد را تشخیص می‌دهد.

- تشخیص صدا

تشخیص صدا با استفاده از یادگیری عمیق در زمینه‌های زیادی کاربرد دارد. یکی از معروف‌ترین کاربردهای واقعی آن در دستیاران هوشمند صوتی^۳ در اکثر گوشی‌های هوشمند می‌باشد. فرمان کاربر با استفاده از مکالمه صوتی صورت گرفته، انجام می‌شود.

- تشخیص الگو

از فرآیند تشخیص الگو در حجم زیاد داده در مواردی چون تشخیص ناهنجاری‌های موجود در داده‌ها برای جلوگیری از کلاهبرداری و کشف الگوهای ثابت برای تشخیص بیماری‌های خاص استفاده می‌شود.

یک مثال کاربردی در دنیای واقعی سیستم پی‌پال^۴ می‌باشد. پی‌پال سیستم پرداخت آنلاین جهانی می‌باشد که انتقال الکترونیکی وجه را انجام می‌دهد. این سیستم از یادگیری عمیق استفاده می‌کند که از کلاهبرداری در خرید و انتقال وجه جلوگیری می‌کند.

- جست‌وجوی تصاویر یا برچسب گذاری تصاویر

^۱Facebook: www.facebook.com

^۲FaceFirst: www.facefirst.com

^۳voice-activated intelligent assistants

^۴PayPal: www.paypal.com

در فرآیند برچسب گذاری تصاویر، اشیا مختلف در تصاویر شامل ساختمان‌ها، طبیعت، مردم و غیره به کمک یادگیری عمیق تشخیص داده می‌شوند. مثال کاربردی در این زمینه موتورهای جست‌وجو همچون گوگل^۱ و شاتر استاک^۲ می‌باشند. استفاده از یادگیری عمیق در موتورهای جست‌وجو کار پیدا کردن تصاویر را به میزان زیادی تسهیل می‌کند.

۱-۴ تعریف و هدف پایان‌نامه

در این فصل با کاربردهای یادگیری عمیق در زمینه‌های مختلف آشنا شدیم. همانطور که ذکر شد یادگیری عمیق در زمینه‌های مختلف کارآمد و مؤثر می‌باشد. از این‌رو برای دسته‌بندی سیگنال EEG از یادگیری عمیق استفاده کردیم. در این پایان‌نامه سعی بر پیشنهاد الگوریتمی بر مبنای یادگیری عمیق جهت دسته‌بندی سیگنال EEG می‌باشد. در واقع قصد داریم از ساختارهای عمیق به جای ساختارهای کم‌عمق استفاده کنیم. تصمیم‌ها و فعالیت‌های انسان از احساساتش تأثیر می‌پذیرد، بنابراین شناسایی احساسات می‌تواند نقش مهمی در پیش‌بینی رفتار افراد داشته باشد. همچنین احساسات بخش بسیار مهمی از فرآیند ارتباطی انسان با انسان به شمار می‌رود. به همین دلیل در دسته‌بندی سیگنال EEG کاربرد تشخیص احساسات را در نظر گرفتیم. در واقع قصد داریم حالت‌های مختلف احساسی فرد را از سیگنال EEG ضبط شده تشخیص دهیم. بدین منظور سیگنال EEG به عنوان ورودی و ارزیابی شخص از احساسات خود به عنوان برچسب به سیستم داده می‌شود. ارزیابی شخص از حالت احساسی خود در قالب دو معیار valence و arousal بیان

^۱Google: www.google.com

^۲Shutterstock: www.shutterstock.com

می‌شود. از سیگنال EEG ویژگی استخراج می‌شود. دسته‌بندی ویژگی‌ها استخراج شده از سیگنال EEG با استفاده از یادگیری عمیق انجام می‌شود. دسته‌بندی برای مقادیر مختلف valence و arousal که بیانگر حالات احساسی متفاوت می‌باشند انجام می‌شود.

۵ - ۱ ساختار پایان‌نامه

فصل اول شامل مقدمه‌ای راجع به یادگیری عمیق و کاربردهای آن در دنیای واقعی می‌باشد. فصل دوم شامل مفاهیم اولیه یادگیری عمیق و احساسات می‌باشد. در این فصل پس از معرفی مفاهیم یادگیری عمیق، احساسات در تعاملات انسان و ماشین، مفهوم احساسات و روش‌های مدل کردن احساسات بیان شده است. در فصل سوم ابتدا توضیح مختصری در خصوص ساختار مغز انسان و EEG آمده است و با بررسی مطالعاتی که در گذشته با استفاده از سیگنال EEG در تشخیص احساسات انجام شده است، دنبال شده است. در فصل چهارم ابتدا پایگاه داده‌های پرکاربرد شامل سیگنال EEG در زمینه تشخیص احساس را معرفی کردیم. پس از آن، ابزارهای مورد استفاده در روش پیشنهادی را بطور مفصل مورد بررسی قرار دادیم و در نهایت روش پیشنهادی بیان شده است. در فصل پنجم نتایج آزمایش‌ها و ارزیابی روش پیشنهادی در مقایسه با دیگر روش‌های دسته‌بندی آورده شده است. فصل ششم نیز شامل نتیجه‌گیری و پیشنهادهایی برای کارهای آتی در این زمینه می‌باشد.

فصل ۲

یادگیری عمیق و احساسات

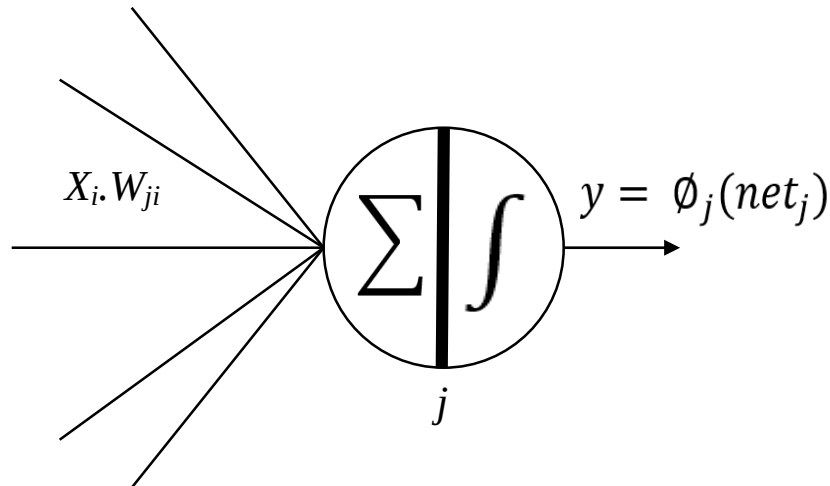
۲- ۱ مفاهیم یادگیری عمیق

در سال‌های اخیر یادگیری عمیق در بسیاری از مسائل تشخیص الگو مورد استفاده قرار گرفته است. این نوع یادگیری الگوهای انتزاعی موجود در داده‌ها را آشکار می‌کند. یادگیری عمیق به دلیل یافته‌های جدید در جنبه‌های بیولوژیکی مغز و پیشرفت و توسعه سخت‌افزارهایی که پردازش موازی را ممکن ساخته‌اند، مورد توجه قرار گرفته است. روش‌های یادگیری عمیق با الگوریتم‌های متداول بهینه‌سازی و آموزش برای کاربردهای فراوانی از جمله در تشخیص الگو مناسب و کارآمد هستند. یادگیری عمیق بر پایه شبکه‌های عصبی شکل گرفته‌اند در واقع شبکه‌های عصبی مسیر اصلی یادگیری عمیق هستند. به همین دلیل در ابتدا مروری بر شبکه‌های عصبی داریم سپس با یادگیری عمیق آشنا می‌شویم.

۲- ۱- ۱ شبکه عصبی مصنوعی

شبکه عصبی مصنوعی یک مدل ریاضی الهام گرفته از مغز انسان است که تلاش می‌کند توان محاسباتی سیستم عصبی انسان را مدل کند. شبکه عصبی مصنوعی از عناصری به نام نورون تشکیل شده است که در یک ساختار معین به هم مرتبط می‌باشند. هر نورون محاسبات جدا انجام می‌دهد و نتیجه را به نورون‌های بعدی می‌فرستد. در نهایت نتیجه با استفاده از کل سیستم بدست می‌آید. شبکه‌های عصبی یک مدل ریاضی برای هر تابع $f: X \rightarrow Y$ یا یک توزیع احتمالاتی روی X تعریف می‌کنند.

اساساً نورون‌ها دو نوع محاسبه انجام می‌دهند (شکل ۲-۱): ابتدا مجموع گرفتن روی همه‌ی مقادیر، فعال‌سازی نورون‌های ورودی وزندار با وزن‌های w_{ij} : $net_j = \sum_i x_i w_{ji}$ و سپس اعمال یک تابع بر روی این مجموع $y_j = \phi_j(net_j)$. y_j خروجی نورون j می‌باشد. تابع اعمال شده برای کاربردهای مختلف فرق می‌کند. مثل: خطی،^۱ منطقی،^۲ سیگموئید،^۳ softmax.



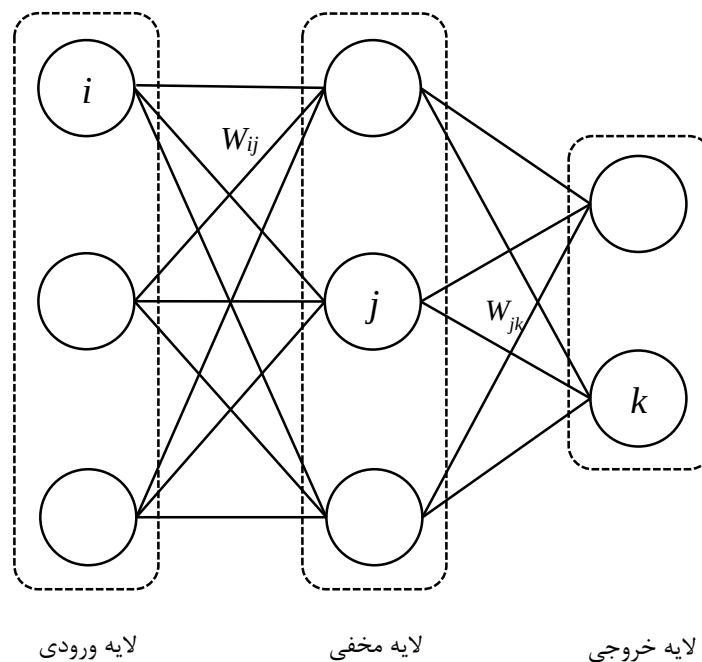
شکل ۲-۱. واحد محاسباتی، نورون

برای تشکیل یک شبکه عصبی مصنوعی در ساختارهای مختلف نورون‌ها به هم مرتبط هستند. یکی از متداول‌ترین ساختارها، ساختار سه لایه می‌باشد که در آن نورون‌ها در سه لایه قرار گرفته‌اند: لایه ورودی، لایه مخفی و لایه خروجی (شکل ۲-۲).

^۱Linear

^۲Logistic

^۳Sigmoidal



شکل ۲-۲. ساختار سه لایه شبکه عصبی

لایه ورودی معمولاً یک تابع فعال‌ساز خطی دارد در حالی که لایه‌های مخفی و خروجی یک تابع فعال‌سازی غیرخطی دارند. داده‌ها در لایه ورودی به شبکه وارد می‌شوند و از آنجا به لایه مخفی انتشار می‌یابند. لایه مخفی محاسبات را انجام می‌دهد و اطلاعات را به لایه خروجی انتشار می‌دهد. لایه خروجی نتیجه نهایی را محاسبه می‌کند. این نوع شبکه‌ها رو به جلو^۱ نامیده می‌شوند زیرا اطلاعات را از ورودی به خروجی بدون وجود حلقه انتشار می‌دهند.

۲ - ۱ - ۲ مدل‌های بر مبنای انرژی شبکه‌های عصبی

مدل‌های بر مبنای انرژی یک عدد را به عنوان انرژی $E(X.Y)$ به هر پیکربندی از متغیرهای دلخواه نسبت می‌دهند [۴]. فرآیند یادگیری در مدل‌های بر مبنای انرژی به پیدا کردن مقادیر بهینه پارامترها (مثل وزن‌ها

^۱Feed forward

و بایاس‌ها) منتهی می‌شود که کمینه انرژی را به مقادیر صحیح متغیرهای هدف و انرژی بیشتر را به مقادیر غیر صحیح متغیرهای هدف نسبت می‌دهد. در واقع در این مدل‌ها یادگیری تغییر تابع انرژی می‌باشد تا زمانی که تابع انرژی پیدا شود که در آن پیکربندی به متغیرهای مشاهده شده X انرژی کمتری نسبت به متغیرهای مشاهده نشده Y نسبت دهد. این تنها محدودیتی است که در آموزش مدل‌های بر مبنای انرژی وجود دارد که این مدل‌ها را برای آموزش بسیار انعطاف‌پذیر می‌سازد. آموزش می‌تواند از طریق جست‌وجو مجموعه‌ای از توابع انرژی که توسط وزن‌ها به عنوان پارامترهای سیستم شاخص‌گذاری شده‌اند مدل شوند. در استفاده از یک مدل بر مبنای انرژی، ورودی X داده می‌شود و مدل باید Y^* از مجموعه $\mathcal{Y}$ را بگونه‌ای انتخاب کند که $E(X.Y)$ کمترین مقدار انرژی را داشته باشد.

$$Y^* = \underset{Y \in \mathcal{Y}}{\operatorname{argmin}} E(X.Y) \quad (1-2)$$

مدل احتمالاتی بر مبنای انرژی یک توزیع احتمالاتی را از طریق یک تابع انرژی تعریف می‌کند:

$$p(x) = \frac{e^{-\text{Energy}(x)}}{Z} \quad (2-2)$$

ثابت نرمال‌سازی Z که تابع تقسیم نامیده می‌شود از مجموع همه حالت‌های انرژی روی تمام ورودی‌ها بدست می‌آید. در صورتی که x پیوسته باشد با انتگرال محاسبه می‌شود.

$$Z = \sum_x e^{-\text{Energy}(x)} \quad (3-2)$$

در موارد بسیاری تمام اجزای x_i از x قابل مشاهده نیستند که در این حالت از متغیر مخفی h استفاده می‌شود. بنابراین بخش قابل مشاهده x و بخش مخفی بعنوان h در نظر گرفته می‌شود.

$$p(x.h) = \frac{e^{-\text{Energy}(x.h)}}{Z} \quad (4-2)$$

اگر تنها x قابل مشاهده باشد

$$p(x) = \frac{e^{-Energy(x,h)}}{Z} \quad (5-2)$$

۲- ۱- ۳ مدل‌های شرطی بر مبنای انرژی

در حالت کلی محاسبه تابع تقسیم Z مشکل می‌باشد. به همین دلیل بجای در نظر گرفتن همه حالات پیکربندی

(x, h) تنها کافی است پیکربندی‌های h با داشتن x را در نظر بگیریم [۵].

$$p(x, h) = \frac{e^{-Energy(x,h)}}{\sum_h e^{-Energy(x,h)}} \quad (6-2)$$

در این مورد گرادیان احتمال شرطی با توجه به پارامترهای تابع انرژی می‌تواند محاسبه شود.

۲- ۱- ۴ نمونه‌برداری گیبز^۱

نمونه‌برداری گیبز یکی از الگوریتم‌های مورد استفاده برای آموزش ماشین بولتزمن محدود می‌باشد. نمونه‌برداری گیبز یک الگوریتم زنجیره مارکوف مونت کارلو^۲ می‌باشد. زمانی که نمونه‌گیری مستقیم مشکل می‌باشد، برای محاسبه کردن یک توالی از مشاهدات (نمونه‌ها) از توزیع احتمالاتی بکار می‌رود [۶].

^۱Gibbs sampling

^۲Markov Chain Monte Carlo (MCMC)

در نمونه‌برداری گیبز با داشتن توزیع چندمتغیره نمونه‌برداری با استفاده از توزیع شرطی نسبت به توزیع توأم ساده‌تر است. برای مثال فرض کنید که ما می‌خواهیم K نمونه از $X = \{x_1, x_2, \dots, x_n\}$ از توزیع توأم $p(x_1, x_2, \dots, x_n)$ آن بدست آوریم. i امین نمونه به صورت $x^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}\}$ مشخص می‌شود. مراحل نمونه‌برداری گیبز به صورت زیر می‌باشد.

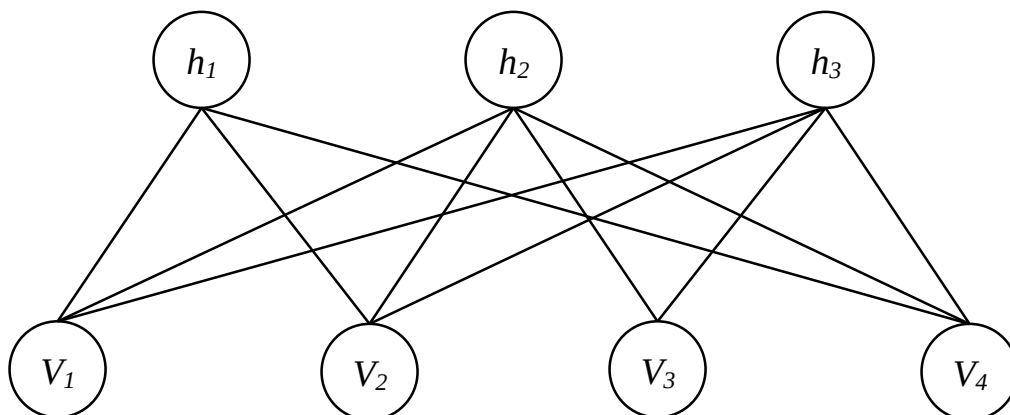
۱. به هر نمونه یک مقدار اولیه $X^{(0)}$ داده می‌شود.

۲. برای هر نمونه $i = \{1, \dots, k\}$ هر متغیر $x_j^{(i)}$ از توزیع شرطی $p(x_j | x_1^{(i)}, \dots, x_{j-1}^{(i)}, x_{j+1}^{(i-1)}, \dots, x_n^{(i-1)})$ نمونه گرفته می‌شود. به این صورت که برای نمونه‌گیری هر متغیر، توزیع شرطی آن متغیر روی همه متغیرهای دیگر محاسبه می‌شود. که در هر مرحله از آخرین مقادیر به‌روزرسانی شده متغیرها استفاده می‌شود.

۲-۱-۵ ماشین بولتزمن محدود^۱ (RBM)

ماشین بولتزمن محدود که نوعی مدل بر مبنای انرژی می‌باشد بعنوان یک مدل گرافیکی غیر جهت‌دار ارائه شده است [۷]. RBM از دو لایه تشکیل می‌شود: لایه مشاهده‌پذیر و لایه مخفی (شکل ۲-۳). در این مدل واحدهای لایه مخفی با اتصالات وزندار به لایه مخفی متصل شده‌اند با این محدودیت که واحدهای هر لایه به هم متصل نیستند. عموماً لایه مخفی تنها شامل واحدهای باینری می‌باشد اگرچه لایه مخفی می‌تواند مقادیر حقیقی یا باینری داشته باشد.

^۱Restricted Boltzman Machine



شکل ۲-۳. ماشین بولتزمن محدود که در آن h_i واحدهای مخفی و v_i واحدهای مشاهده‌پذیر هستند تابع انرژی برای نگاشت واحدهای مشاهده‌پذیر و مخفی RBM به یک فضای انرژی تعریف شده است و الگوریتم یادگیری شامل هدایت پارامترهای مدل (اتصالات وزندار و بایاس واحدها) است بگونه‌ای که این پیکربندی را به کمینه محلی انرژی نگاشت کند.

$$E(v, h) = -b'v - c'h - h'Wv \quad (۷-۲)$$

که در آن b و c به ترتیب بایاس واحدهای مشاهده‌پذیر و مخفی هستند. W ماتریس وزن‌های اتصالات، v و h

به ترتیب مقادیر فعال‌سازی واحدهای ورودی و واحدهای مخفی هستند. این یک شبکه عصبی تولیدی آماری می‌باشد که می‌تواند یک توزیع احتمالاتی را روی مجموعه‌ای از ورودی‌ها یاد بگیرد.

۲ - ۱ - ۵ فعال‌سازی واحدها

از آنجایی که RBM یک گراف دو قسمتی بدون اتصالات درون لایه‌ای می‌باشد، فعال‌سازی واحدهای مخفی مستقل از فعال‌سازی واحدهای مشاهده‌پذیر می‌باشد و بالعکس.

$$P(v|h) = \prod_i P(v_i|h) \quad (۸-۲)$$

$$P(h|v) = \prod_j P(h_j|v) \quad (۹-۲)$$

زمانی که بردار ورودی به لایه مشاهده‌پذیر داده می‌شود، فعال‌سازی واحدهای باینری مخفی بصورت زیر محاسبه می‌شود.

$$P(h|v) = \text{sigm}(c_i + W_i v) \quad (۱۰-۲)$$

واحدهای لایه مشاهده‌پذیر می‌توانند با داشتن مقادیر واحدهای لایه مخفی بازسازی شوند:

$$P(v|h) = \text{sigm}(b_j + W'_j h) \quad (۱۱-۲)$$

۲ - ۱ - ۵ - ۲ آموزش RBM

ماشین بولتزمن محدود برای بیشینه کردن احتمال مجموعه‌ای از بردارهای X آموزش داده می‌شود.

$$\underset{W, b, c}{\operatorname{argmax}} \prod_{x \in X} P(x) \quad (۱۲-۲)$$

یا با استفاده از لگاریتم داریم:

$$\underset{W, b, c}{\operatorname{argmin}} \sum_{n=1}^N \log P(x) \quad (۱۳-۲)$$

یادگیری بیشینه احتمال (ML) از پارامترهای θ با توزیع مستقل و یکسان $X = \{x_i\}_{n=1}^N$ می‌تواند توسط گرادیان صعودی انجام شود.

$$\theta^{(\tau+1)} = \theta^{(\tau)} + \delta \frac{\partial L(X; \theta)}{\partial \theta} \quad (۱۴-۲)$$

که در آن θ به مجموعه پارامترهای W ، b و c مربوط می‌باشند.

احتمال میانگین به صورت زیر محاسبه می‌شود: [۸]

$$L(X; \theta) = \frac{1}{N} \sum_{n=1}^N \log P(x; \theta) \quad (15-2)$$

$$= \langle \log P(x; \theta) \rangle_{data}$$

$$= - \langle E(x; \theta) \rangle_{data} - \log z(\theta)$$

که در آن N تعداد نمونه‌ها می‌باشد. یک مشکل متداول در محاسبات گرادیان به صورت زیر می‌باشد.

$$\frac{\partial L(X; \theta)}{\partial \theta} = - \langle \frac{\partial E(x; \theta)}{\partial \theta} \rangle_{data} + \langle \frac{\partial E(x; \theta)}{\partial \theta} \rangle_{model} \quad (16-2)$$

که در آن $\langle \rangle_{data}$ میانگین نمونه داده‌ها می‌باشد و $\langle \rangle_{model}$ با توجه به توزیع داده‌ها میانگین را حساب می‌کند. $E(x; \theta)$ تابع انرژی پارامترهای θ برای نمونه x می‌باشد. میانگین $\langle \rangle_{data}$ با آسانی با استفاده از داده نمونه X قابل محاسبه می‌باشد ولی $\langle \rangle_{model}$ با ثابت نرمال‌سازی $Z(W)$ (تابع تقسیم) سروکار دارد که به آسانی قابل محاسبه نیست. روش استاندارد این است که میانگین روی توزیع با استفاده از میانگین روی یک نمونه از $P(x; \theta)$

تخمین زده شود که با استفاده از اجرای زنجیره مارکوف به تعادل می‌رسد. برای مثال نمونه‌گیر گیبز با شروع از واحدهای مشاهده‌پذیر داده شده و نمونه‌گیری به ترتیب از واحد مخفی برای مدت زمان طولانی (بطور مطلوب بینهایت) انجام می‌شود. در [۸] نشان داده شده است که همگرایی تضمین شده می‌باشد. در این مورد توالی نمونه‌گیری با مراحل زیر انجام می‌شود:

$$\begin{aligned} x_1 &\sim \hat{P}(x) \\ h_1 &\sim P(h|x_1) \\ x_2 &\sim P(x|h_1) \\ x_1 &\sim P(h|x_2) \\ &\dots \\ x_{k+1} &\sim P(x|h_k) \end{aligned} \quad (17-2)$$

پس معادله زیر برای به‌روزرسانی وزن‌ها استفاده می‌شود.

$$\Delta w_{ij} = \delta(< v_i h_j >_{data} - < v_i h_j >_{model}) \quad (۱۸-۲)$$

که در آن $<>$ امید ریاضی می‌باشد.

اجرای یک زنجیره مارکوف (برای مثال نمونه‌برداری گیبز) آهسته می‌باشد زیرا به تعداد واحدهای موجود در شبکه زیر مرحله در زنجیره گیبز وجود دارد. به همین دلیل یک روش کارآمد برای آموزش RBM استفاده می‌شود که همگرایی متقابل [۸] نامیده می‌شود.

۲-۱-۶ همگرایی متقابل

از آنجایی که نمونه‌گیری گیبز بار محاسباتی زیادی دارد از همگرایی متقابل [۹] برای آموزش RBM استفاده می‌شود. این روش از گرادیان احتمال پیروی نمی‌کند ولی به جای آن گرادیان یک تابع دیگر را تخمین می‌زند. یادگیری بیشینه احتمالاتی همگرایی Kullback-Leibler را کمینه می‌کند.

$$KL(p_0||p_\infty) = \sum_x p_0(x) \log \frac{p_0(x)}{p(x; \theta)} \quad (۱۹-۲)$$

که در آن p_0 توزیع احتمالاتی داده می‌باشد و p_∞ توزیع احتمال مدل می‌باشد. یادگیری CD از گرادیان تفاضل دو همگرایی پیروی می‌کند.

$$CD_n = KL(p_0||p_\infty) - KL(p_n||p_\infty) \quad (۲۰-۲)$$

در به‌روزرسانی وزن‌ها این قضیه به جای مدل توزیع در توزیع احتمال بازسازی داده ورودی استفاده می‌شود.

نتیجه تجربی شگفت‌آور این می‌باشد که حتی برای $n=1$ (CD₁)، تنها یک تکرار) مدل معمولاً نتایج خوبی می‌دهد.

$$\Delta w_{ij} = \epsilon \left(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recon} \right) \quad (2-21)$$

الگوریتم آموزش RBM با استفاده از CD در شکل ۲-۴. الگوریتم آموزش RBM آورده شده است.

الگوریتم آموزش RBM با استفاده از CD_1

δ : نرخ یادگیری CD W : ماتریس وزن‌های RBM

b : واحدهای مشاهده‌پذیر RBM c : واحدهای مخفی RBM

h_1 : اولین نمونه از لایه مخفی با انتشار x_1 x_1 : ورودی لایه مشاهده‌پذیر

h_2 : دومین نمونه از لایه مخفی با انتشار x_2 x_2 : دومین لایه بازسازی شده از h_1

گام اول: انجام عملیات زیر برای همه واحدهای مخفی i در لایه مخفی

- محاسبه فعال‌سازی واحد i با استفاده از معادله (۲-۱۰)

- نمونه‌گیری $h_{1i} \in \{0,1\}$ از $P(h_{1i}|x_1)$

گام دوم: انجام عملیات زیر برای همه واحدهای مشاهده‌پذیر j در لایه مشاهده‌پذیر

- محاسبه فعال‌سازی واحد j با استفاده از معادله (۲-۱۱)

- نمونه‌گیری $x_{2j} \in \{0,1\}$ از $P(x_{2j}|h_1)$

گام سوم: انجام عملیات زیر برای همه واحدهای مخفی i در لایه مخفی

- محاسبه فعال‌سازی واحد i با استفاده از معادله (۲-۱۰)

- نمونه‌گیری $h_{2i} \in \{0,1\}$ از $P(h_{2i}|x_2)$

گام چهارم:

$$w \leftarrow w + \delta(h_1 x'_1 - h_2 x'_2)$$

$$b \leftarrow b + \delta(x_1 - x_2)$$

$$c \leftarrow c + \delta(h_1 - h_2)$$

شکل ۲-۴. الگوریتم آموزش RBM

۲ - ۱ - ۷ شبکه باور عمیق

یک شبکه باور عمیق (DBN) یک ساختار عمیق می‌باشد که واحد سازنده آن RBM می‌باشد [۱۰]. RBM‌های انباشته شده روی هم با لایه‌های مشترک یک DBN برای پیش آموزش نظارتی را می‌سازند. در RBM‌های انباشته شده واحد مخفی RBM پایینی واحد ورودی RBM بالایی می‌باشد (شکل ۲-۵). در فاز آموزش، RBM‌های انباشته شده یکی پس از دیگری آموزش می‌بینند. در فاز تولیدی اتصالات رو به بالا (پیکان‌های نقطه‌چین در شکل ۲-۵) برداشته می‌شوند و تنها دو لایه بالایی یک RBM باقی می‌مانند. سپس مدل قادر به تولید بردار داده می‌باشد که تقریباً به داده اولیه نزدیک می‌باشد و از توزیع داده اولیه پیروی می‌کند.

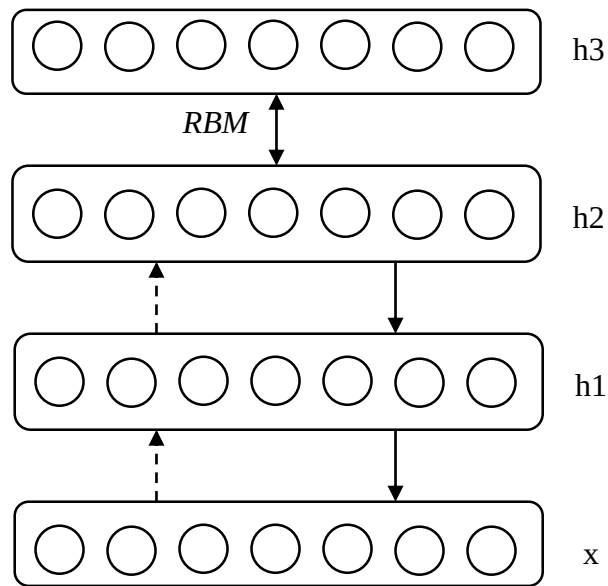
با توجه به سطح‌های چندگانه، احتمال توأم مدل تولیدی به صورت زیر تجزیه می‌شود:

$$P(x, h^1, h^2, \dots, h^l) = P(h^{l-1}, h^l) \left(\prod_{k=1}^{l-1} P(h^k | h^{k+1}) \right) P(x | h^l) \quad (22-2)$$

که در آن l تعداد لایه‌های مخفی می‌باشد. توزیع توأم دو لایه بالایی یک RBM می‌باشد.

$$P(h^{l-1}, h^l) \propto e^{b'h^{l-1} + c'h^l + h^{l'} W h^{l-1}} \quad (23-2)$$

آموزش ساختارهای عمیق با شروع از مقداردهی اولیه وزن‌ها و استفاده از تکنیک انتشار به عقب منجر به تعمیم ضعیف این ساختارها می‌شود. به همین دلیل از پیش‌آموزش غیر نظارتی استفاده می‌شود. الگوریتم‌های یادگیری غیر نظارتی اطلاعات برجسته راجع به توزیع آماری داده‌های ورودی را استخراج می‌کنند. پیش آموزش لایه اول با الگوریتم یادگیری غیر نظارتی این اطلاعات را استخراج می‌کند. ولی با توجه به ظرفیت محدود آن لایه، ویژگی‌های استخراج شده ویژگی‌های سطح پایین هستند. آموزش لایه دوم نیز منجر به استخراج ویژگی‌های برجسته از لایه اول می‌شود که ویژگی‌های سطح بالا هستند.



شکل ۲-۵. شبکه باور عمیق (DBN)

الگوریتم یادگیری DBN بصورت زیر می‌باشد (شکل ۲-۶):

زمانی که یک DBN بصورت لایه‌ای حریصانه آموزش داده می‌شود هر لایه به عنوان یک RBM در نظر گرفته می‌شود. $Q(h^k|h^{k-1})$ به تخمین $P(h^k|h^{k-1})$ در K مین RBM آموزش داده شده اشاره دارد و $P(h^k|h^{k-1})$ احتمالات RBM می‌باشد. طبق همین روال $Q(h^k|h^{k-1})$ تخمینی از $P(h^k|h^{k-1})$ می‌باشد.

۲ - ۱ - ۸ شبکه عصبی عمیق

آزمایش‌ها نشان می‌دهد که آموزش ساختارهای عمیق مشکل‌تر از ساختارهای سطحی می‌باشد. نتایج در [۱۱] نشان داد که آموزش با انتشار رو به عقب شبکه‌های عصبی چندلایه در کمینه محلی گیر می‌کند و

هر چه تعداد لایه‌ها بیشتر باشد بدست آوردن یک تعمیم خوب مشکل‌تر می‌شود. یکی از دلایل عمده تنظیم^۱ ضعیف لایه‌های پایین می‌باشد.

اما در [۱۰] توضیح داده شد که پیش‌آموزش با الگوریتم غیر نظارتی نتایج را بهبود می‌بخشد. با پیش‌آموزش لایه‌های پایین بهینه می‌شوند و خطای آموزش لایه‌های بالایی کم می‌شود و تعمیم بهتری انجام می‌شود. برای انجام فاز پیش‌آموزش یک شبکه باور عمیق بگونه‌ای که در بخش ۲ - ۱ - ۷ توضیح داده شد، ساخته می‌شود. یادگیری غیر نظارتی در یک DBN به مقداردهی اولیه وزن‌ها و بایاس واحدها منجر می‌شود که بعنوان مقداردهی اولیه برای آموزش نظارتی استفاده می‌شود.

بعد از آموزش غیر نظارتی، یک ساختار عمیق با پارامترهای DBN پیش‌آموزش (وزن‌ها و بایاس واحدها) با تعداد برابر لایه‌های مخفی و یک لایه خروجی از نوع دلخواه (در این تحقیق softmax) ساخته می‌شود، که در این صورت به عنوان یک شبکه عمیق رو به جلو یا شبکه عصبی عمیق^۲ (DNN) شناخته می‌شود. بعد از تبدیل یک DBN به DNN مرحله نهایی تنظیم یا همان آموزش نظارتی طبق بهینه‌سازی گرادیان نزولی با الگوریتم انتشار رو به عقب انجام می‌شود.

۲ - ۱ - ۸ - ۱ دسته‌بند نرم بیشینه^۳:

دسته‌بند نرم بیشینه احتمال مقادیر خروجی DNN را تخمین می‌زند. این دسته‌بند وزن‌ها و پارامترهای بایاس را با استفاده از ویژگی‌های آخرین لایه مخفی یاد می‌گیرد [۱۲]. خروجی فرضیه رگرسیون نرم بیشینه به صورت زیر محاسبه می‌شود.

^۱tune

^۲Deep neural network

^۳softmax

$$h_{\theta}(x) = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (22-2)$$

لایه نرم بیشینه نیاز دارد که وزن‌ها و پارامترها را با یادگیری نظارتی با کمینه کردن تابع هزینه زیر بدست آورد:

$$Cost = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k 1\{y_i = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} + \frac{\lambda}{2} \sum_{i=1}^k \sum_{j=1}^n \theta_{ij}^2 \quad (23-2)$$

که در آن m تعداد واحدهای لایه مخفی، n تعداد ورودی‌ها، k تعداد کلاس‌ها، y برچسب واقعی و θ وزن لایه‌های مخفی می‌باشد.

آموزش غیر نظارتی DBN ($\hat{P}, \delta, l, W, b, c$) با رویه‌ای لایه‌ای حریصانه (همگرایی متقابل)

$\hat{P}$: توزیع داده آموزشی ورودی شبکه

δ : نرخ یادگیری RBM

l : تعداد لایه‌ها برای آموزش

W^k : ماتریس وزن سطح k ، k از ۱ تا l

b^k : بایاس واحدهای مشاهده‌پذیر RBM سطح k ، k از ۱ تا l

c^k : بایاس واحدهای مخفی RBM سطح k ، k از ۱ تا l

برای $k = 1$ تا l انجام بده

- مقداردهی اولیه $c^k = 0$ ، $b^k = 0$ ، $W^k = 0$

- تا زمانی که شرط توقف برآورده شود مراحل زیر را انجام بده

- نمونه‌گیری $x = h^0$ از $\hat{P}$

- برای $i = 1$ تا $k - 1$ انجام بده

- نمونه‌گیری h_j^i از $Q(h_j^i | h^{i-1})$ ، برای همه عناصر j از h^i

- به‌روزرسانی RBM ($h^{k-1}, \delta, l, W^k, b^k, c^k$)

شکل ۲-۶. الگوریتم یادگیری DBN

۲ - ۲ مقدمه‌ای بر احساسات

حالت‌های مختلف احساسی بطور چشم‌گیری روی بخش‌های مختلف زندگی روزمره انسان‌ها تأثیر می‌گذارند مثل یادگیری، تصمیم‌گیری و ارتباط برقرار کردن با دیگران. با پیشرفت تکنولوژی زندگی روزمره انسان‌ها به ارتباط بین انسان با انسان و انسان با ماشین گره خورده است. انسان‌ها در رمزگشایی نشانه‌های احساسی که در روابطشان با یکدیگر بوجود می‌آید، بسیار ماهر هستند و قادرند پاسخ‌های مناسب و درخور شرایط

بیان کنند. این توانایی باعث می‌شود یک ارتباط مؤثر و راحت‌تر بین انسان‌ها برقرار شود. اما در روابط انسان با ماشین، از آنجایی که پاسخ ماشین‌ها از قبل تعریف شده و شرطی می‌باشد و شرایط فعلی انسان‌ها در نظر گرفته نمی‌شود، انسان‌ها معمولاً از پاسخ داده شده راضی نیستند و ارتباط مؤثر بین آن‌ها شکل نمی‌گیرد. این تحقیق تلاش می‌کند تا توانایی ماشین‌ها را در تشخیص صحیح حالت‌های احساسی انسان‌ها را بهبود بخشد. در این بخش پس از معرفی مفاهیم هوش احساسی و محاسبه‌های احساسی، کاربرد سیستم‌های بر پایه احساسات بیان می‌شود. پس از آن مفاهیم مهم و اساسی احساسات شامل: تعریف احساسات، تعریف اجزای احساسات و مدل‌های احساسی برای تعریف احساسات انسان‌ها مورد بحث و گفت‌وگو قرار می‌گیرد.

۲-۲-۱ هوش احساسی و تعامل انسان و ماشین

زمانی که واژه احساس در محیط‌های غیرعلمی مورد استفاده قرار می‌گیرد، معمولاً به احساسات عمومی انسان‌ها اشاره می‌کند. احساساتی مثل: عصبانیت، تعجب، ناراحتی، شادی، تنفر و ترس. اما بطور علمی احساس یک تجلی و ابراز شخصی است که با توجه به موقعیت اجتماعی، زمینه فرهنگی و تجربیات شخصی متفاوت است. مفهوم و تعریف بیان شده از احساس جهت استفاده در کاربردهای شناخت احساسات کمی و کاستی‌هایی دارد و فاقد مشخصه و ویژگی‌های کافی و ضروری می‌باشد.

راسل^۱ یک تفسیر جایگزین و کاربردی از احساس را ارائه داده است که طبق آن در هر زمانی، حالت احساسی هر شخص به عنوان یک نقطه بر روی مدل احساسات نمایش داده می‌شود [۱]. (جزئیات بیشتر در بخش

۲-۲-۴)

^۱Russel

هوش احساسی مفهومی است که جا دادن و دخیل کردن آن در تشکیل روابط ماشین و انسان، این روابط را قابل حس‌تر، انعطاف‌پذیرتر و کارآمدتر می‌کند [۱۵]. بدون نشانه‌های احساسی، روابط و تعامل ما با ماشین‌ها می‌تواند آزاردهنده باشد. برای مثال اینجا نمونه ساده یک کاربر را مطرح می‌کنیم، که نیاز دارد بعد از خرید یک محصول تماسی تلفنی با سرویس مشتری برقرار کند. این کاربر به سیستم پاسخگویی صوتی متقابل^۱ یا در اصطلاح عامل مجازی^۲ وصل می‌شود. در موارد بسیاری، هیچ راه دیگری وجود ندارد تا در صف تماس گیرنده‌ها قرار گیرد مگر اینکه از عامل مجازی عبور کند. اما بعد از چند بار تلاش کردن، عامل مجازی هنوز قادر به تشخیص این نیست که کاربر چه می‌گوید. درهم کوبیدن تلفن و فریاد زدن کاربر بر سر کامپیوتر از پاسخ‌های معمول چنین ناکامی‌هایی هستند. با توجه به استفاده روزافزون از ماشین‌ها، کم کردن فاصله در رابطه انسان با ماشین و بیشتر شبیه کردن این ارتباط به ارتباط بین انسان و انسان کار به‌جا و درستی می‌باشد. هوش احساسی واسط انسان و ماشین^۳ به‌موردی اشاره دارد که ماشین‌ها، مثل یک کامپیوتر شخصی، می‌توانند حالت‌های احساسی کاربرها را کشف کنند، تشخیص دهند و به آن‌ها پاسخ دهند. به این محاسبه‌ی انسان‌محور، محاسبه‌های احساسی^۴ (AC) نیز گفته می‌شود.

۲ - ۲ - ۲ محاسبه‌های احساسی

محاسبه احساسی یک موضوع تحقیق فعال و مهم در دو دهه‌ی گذشته بوده است و رشد چشمگیری در چند سال گذشته داشته است. این موضوع تلاش می‌کند تا اختلاف و شکاف ارتباطی بین انسان‌ها و ماشین‌ها

^۱Interactive Voice Response (IVR)

^۲Virtual agent

^۳Human Machine Interface (HMI)

^۴Affective computing

را کم کند. با توجه به پیشرفت واسط انسان ماشین، درک و پاسخ به حالت‌های احساسی کاربر یک نیاز اجتناب‌ناپذیر برای ماشین‌ها می‌باشد. احساسات بطور چشمگیری روی نتایج زندگی روزمره انسان‌ها مثل یادگیری و تصمیم‌گیری تأثیر می‌گذارند. در بیشینه‌ی محاسبات احساسی ۳ نوع کاربرد اصلی قرار دارد: آشکار ساختن احساسات، تقلید احساسات و تشخیص احساسات واقعی [۱۶]. با افزایش توانایی تشخیص احساسات، تعامل انسان و ماشین کارآمدتر و لذت‌بخش‌تر می‌شود. آشکار ساختن احساسات نه تنها قدم مهمی به سوی شناخت احساسات می‌باشد بلکه اولین قدم به سمت کاربردهای بر پایه احساسات نیز محسوب می‌شود. یک ماشین قبل از اینکه احساسات را بطور دقیق بشناسد نمی‌تواند به احساسات یک کاربر پاسخ بدهد.

۲ - ۲ - ۳ کاربردهای سیستم‌های بر پایه احساسات

۲ - ۲ - ۳ - ۱ روابط شبه انسانی

امروزه شمار زیادی از انسان‌ها بیشتر از اینکه با یکدیگر تعامل داشته باشند، با کامپیوترها تعامل دارند. علاوه بر این بیشتر مردم بواسطه‌ی کامپیوترها با هم در ارتباط هستند. مشکلی که تعامل با کامپیوتر یا ارتباط برقرار کردن بواسطه‌ی کامپیوتر دارد این است که کامپیوترها قادر نیستند عکس‌العمل‌های احساسی از خود نشان دهند. تبدیل کردن این ارتباطات به ارتباطات شبه‌انسانی می‌تواند همکاری بین انسان و کامپیوتر را بهبود ببخشد.

یکی از اشکال ارتباط احساسی محدود، ارتباط بواسطه ایمیل می‌باشد. تنها راه نشان دادن احساسات در یک ایمیل استفاده از شکل‌هایی مثل (-؛ ، (-: می‌باشد که به راحتی می‌تواند سوءتفاهم ایجاد کند و باعث کج‌فهمی می‌شود. برای توسعه دادن گستره‌ی یک ایمیل، ابزارهایی که توانایی تشخیص و بیان احساسات

را دارند می‌توانند استفاده شوند. برای مثال، ریتم تایپ و فشاری که به کلیدهای صفحه‌کلید وارد می‌شود را می‌توان اندازه‌گیری کرد و از آن برای تشخیص حالت احساسی فرستنده استفاده کرد.

یکی دیگر از روش‌های تبدیل کردن ارتباطات به ارتباطات شبه‌انسانی، تلفیق متن و سخن می‌باشد که معمولاً با لحن و تلفظ یکنواخت انجام می‌شود. افرادی که مجبور هستند از تلفیق صدا استفاده کنند می‌توانند احساسات خودشان را به سخنانشان اضافه کنند. در ورای همه‌ی اینها محققان به این نتیجه رسیده‌اند که خواندن یک داستان با صدای همراه احساس، نه تنها علاقه و لذت داستان را افزایش می‌دهد، بلکه سرعت فهم و درک داستان توسط کودکان را نیز افزایش می‌دهد [۱۷]. همچنین تلفیق متن و سخن کار را برای افراد نابینا نیز سهل‌تر می‌کند.

هم‌چنین کامپیوتر توانایی این را دارد که مکالمه‌های کوتاه بسازد، برای مثال یک مکالمه دوستانه صبح بخیر می‌تواند جایگزین مناسبی برای نام کاربری^۱ قدیمی شود. این امر می‌تواند تأثیر مثبت روی گرایش کاربر نسبت به کامپیوتر داشته باشد.

۲ - ۲ - ۳ - یادگیری

در طول فرآیند یادگیری، افراد یادگیرنده احساسات متفاوتی مثل رضایت، شادی، آشفتگی و ناراحتی را تجربه می‌کنند. حالت‌های احساسی و هیجانی یادگیرنده بطور چشمگیری روی نتیجه فرآیند یادگیری تأثیر می‌گذارد [۱۸]. اگر یک واسط کامپیوتر بتواند این تغییرات احساسی یادگیرنده را تشخیص دهد، می‌تواند به عنوان یک معلم در کلاس عمل کند و محتوای درسی را که ارائه شد تغییر دهد یا شیوه‌ی بیان آن را عوض کند. این عمل تأثیر مثبتی روی فرآیند یادگیری می‌گذارد و تجربه یادگیری را بطور کلی بهبود می‌بخشد.

¹Login name

۲-۲-۳ کاربردهای سلامتی و روانبخشی

کاربردهای حساس به احساس در پیش‌بینی رفتار و کنترل رفتار بسیار مفید هستند [۱۹]. کامپیوترها می‌توانند به افرادی که از اوتیسم رنج می‌برند کمک شایانی کنند، زیرا این افراد در مواجهه با شرایط احساسی مشکل دارند. اگر به این افراد در فهم و پاسخ به شرایط احساسی کمک شود، کمبود هوش هیجانی آن‌ها جبران خواهد شد. برای مثال کامپیوترها می‌توانند بازخوردهایی را به این افراد نشان دهند، مثلاً زمانی که در حال یادگیری مهارت‌هایی مثل همدردی هستند. انجام این عمل در محیط‌هایی که تعامل‌های احساسی مکرری نیاز است مثل محیط‌های بازی به نتایج بهتری منجر خواهد شد.

از جنبه‌ی سلامتی، ارزیابی مناسب از حالت احساسی بیمار، نشانه‌ای کلیدی از وضعیت سلامت فیزیکی و ذهنی بیمار محسوب می‌شود. فایده استفاده از احساسات در فرآیندهای بهبود دهنده‌ی بیماران مستندسازی شده است [۲۰]. در سناریو سلامتی از راه دور، مسئولین مراکز درمانی می‌توانند کیفیت خدمات خود را بطور چشمگیری افزایش دهند، در صورتی که مجهز به سیستمی باشند که بتواند بطور دقیق حالت‌های احساسی بیمار را ارزیابی کند. جمع‌آوری سیگنال‌های فیزیولوژیکی و نگاشت آن‌ها به حالت‌های احساسی می‌تواند اطلاعات احساسی بیمار را برای مراقبت از بیمار فراهم کند.

۲-۲-۴ بازی و سرگرمی

احساسات در کاربردهای بازی و سرگرمی در کامپیوترها نقش بسزایی را ایفا می‌کنند. محققان مفهوم دی‌جی^۳ احساسی را پیشنهاد دادند [۲۱]. آن‌ها یک کامپیوتر جامع طراحی کردند که می‌تواند اولویت‌ها و

^۱Recovery

^۲Tele-health

^۳DJ

ارجحیت‌های یک شخص را با توجه به نشانه‌های احساسی تشخیص داده و به آن‌ها پاسخ دهد. این اطلاعات به کامپیوتر کمک می‌کند تا شخص را در انتخاب آهنگ متناسب با احساسش یاری کند. سیگنال‌های فیزیولوژیکی، بازی‌ها را با حالت‌های احساسی کاربر سازگار می‌کنند. برای مثال یک رفتار آرام، با معرفی یک همراه و یار جدید پاداش داده می‌شود یا یک عمل شجاعانه زمانی که فرد در سطح برانگیختگی بالایی قرار دارد، می‌تواند منجر به امتیازات بیشتری شود.

۲ - ۲ - ۳ کمک و بازخورد

یکی از کاربردهایی که در نتیجه توسعه الگوریتم‌های کامپیوتری مبتنی بر احساسات بوجود آمده است، عکس‌العمل نشان دادن به اولویت‌های کاربر است. یک کامپیوتر می‌تواند برای شناخت اولویت‌های کاربر اطلاعاتی راجع به علاقه‌مندی‌ها و ناخوشایندی‌های شخص را توسط پاسخ‌های احساسی او جمع‌آوری کند. همچنین برای دستیابی به نتایج بهتر، محتوای جمع‌آوری شده می‌تواند با شرایط احساسی سازگار شود. برای مثال، اگر فردی تحت فشار و استرس باشد، فقط مهم‌ترین آیکن‌ها روی صفحه نمایش داده می‌شوند. علاوه بر کاربرد ذکر شده می‌توان از احساسات برای توسعه ابزارهایی استفاده کرد که به افراد کمک می‌کنند تا متوجه شوند که در شرایط مختلف در نظر دیگران چگونه به نظر می‌آیند. این آینه احساسی^۱ این امکان را فراهم می‌کند که انسان‌ها برای حضور مؤثر و مناسب در شرایط مختلف تمرین کنند. برای مثال برای یک سخنرانی مهم بازخوردهایی را در مورد اجرای خود دریافت کنند.

۲ - ۲ - ۴ تعریف احساسات

"هر کسی می‌داند احساسات چیست مگر اینکه از او بخواهیم تعریفی از احساسات ارائه دهد." [۲۲]

^۱Affective mirror

اگرچه احساسات در ارتباطات و رفتارهای انسانی نقش کلیدی دارند، هنوز هیچ تعریف جامع و قابل قبولی از احساسات وجود ندارد. در واقع، شیوه‌های متفاوت زیادی برای پیدا کردن یک تعریف مناسب وجود دارد که اکثر این تعاریف متناقض هستند. در [۲۳] دو دلیل بیان می‌شود که چرا رسیدن به یک تعریف دقیق برای احساس دشوار است: (۱)، این واژه کاربردهای متفاوتی در شرایط متفاوت دارد که همه‌ی آن‌ها تنها با واژه‌ی احساس به هم متصل شده‌اند. (۲)، تشخیص رخداد‌های احساسی از رخداد‌های غیر احساسی کار بسیار دشواری است، همانگونه یک برانگیختگی معین ممکن است بعنوان یک حالت احساسی یا حالت غیراحساسی تلقی شود.

در سال ۱۹۸۱، محققان ۹۲ تعریف مجزا راجع به احساس را آنالیز و دسته‌بندی کرده‌اند [۲۴]. آن‌ها در نهایت به این نتیجه رسیدند که تطابق کوچکی بین تعاریف وجود دارد و اکثر آن‌ها مبهم هستند. سرانجام آن‌ها براساس یافته‌هایشان تعریف زیر را ارائه دادند که بر جنبه‌های چندلایه‌ای احساس تأکید می‌کند:

"احساس یک مجموعه تعاملات پیچیده در میان عوامل ذهنی و واقعی می‌باشد که در میان سیستم‌های هورمونی و عصبی قرار گرفته است. احساس که می‌تواند:

- از تجربه‌های احساسی مثل حس برانگیختگی، خوشنودی یا ناخوشنودی ناشی شود.
- فرآیندهای شناختی مثل ادراک احساسی و ارزیابی‌ها را انجام دهد.
- سازگاری‌های فیزیولوژیکی گسترده با شرایط برانگیختگی را فعال کند.
- منتهی به رفتاری شود که اغلب نه همیشه، پرمعنی، هدف‌گرا و قابل تطبیق است." [۲۴]

این تعریف کاملاً کلی و عام تلاش می‌کند همه‌ی جنبه‌های مهم احساس را با هم ترکیب کند، اگرچه این جنبه‌ها اغلب با یکدیگر متناقض هستند. بنابراین این تعریف به جای یک نسخه واقعی از تعریف عملی یک حد وسط از همه‌ی نظرات و ایده‌ها ارائه می‌دهد.

۲-۲-۵ تعریف اجزا

احساس به عنوان یک موجودیت پیچیده تلقی می‌شود که از اجزای متفاوتی تشکیل شده است. تعداد و اسامی این اجزای احساسات در میان دانشمندان متفاوت است با توجه به [۲۵] سه جزء ضروری به شرح زیر می‌باشد:

- حسی که بطور ذهنی احساس می‌شود و به حالت‌هایی اشاره می‌کند که توسط خود فرد احساسات نامیده می‌شود.
- عکس‌العمل‌های فیزیولوژیکی و احساسی در مغز و سیستم عصبی که به محرک‌های احساسی نسبت داده می‌شوند.
- رفتار احساسی، رفتاری را توصیف می‌کند که با توجه به مشاهده‌های بیرونی به عنوان عکس‌العملی به یک محرک احساسی اتفاق می‌افتد.

حسی که بطور ذهنی درک می‌شود توسط توصیف‌های شخصی قابل شناسایی می‌باشد در حالی که عکس‌العمل‌های فیزیولوژیکی احساسی توسط سیگنال‌های بیوفیزیولوژیکی مثل EEG و EMG و رسانش پوستی اندازه‌گیری می‌شود. به علاوه رفتارهای احساسی نیز قابل شناسایی هستند برای مثال از ژست یا تشخیص چهره. این سه جزء احساسی بر روی یکدیگر تأثیر می‌گذارند. برای مثال خنده نه تنها بوسیله EMG قابل ارزیابی است بلکه از حالت چهره (رفتار احساسی) و گزارش‌های شخصی نیز قابل توصیف است.

۲-۲-۶ مدل کردن احساسات

در پیشینه‌ی روانشناسی دو مدل اصلی وجود دارد که برای نمایش حالت‌های احساسی واقعی انسان مورد استفاده قرار گرفته است، مدل گسسته و مدل ابعادی. تفاوت اصلی مدل‌ها در تعداد احساسات متفاوتی است

که هر مدل ارائه می‌دهد. مدل گسسته فقط شش احساس متفاوت را نشان می‌دهد که به عنوان احساسات مشترک همه مردم جهان است، در حالی که مدل ابعادی تعداد بیشتری از احساسات را نشان می‌دهد. انتخاب هر مدل بستگی به کاربردی دارد که کاربر قصد استفاده از آن را داشته باشد. مدل گسسته برای کاربردهای تشخیص احساس با استفاده از تصاویر چهره مورد استفاده قرار می‌گیرد. طرفداران این مدل بر این باور هستند که احساسات برای افراد با جنسیت و فرهنگ و قومیت متفاوت یکسان است [۲۶]. مدل ابعادی در روش‌های آنالیز احساسات مورد استفاده قرار می‌گیرد [۲۷]. قسمت بعدی تحقیق قواعد مهم این روش‌ها را فهرست می‌کند.

۲-۲-۶-۱ مدل گسسته

یکی از راه‌های مدل‌سازی احساسات، در نظر گرفتن آن‌ها به عنوان حس‌های جدا از هم و متنوع می‌باشد. مدل گسسته احساسات برای اولین بار در سال ۱۹۷۱ توسط اکمان^۱ [۲۶] معرفی شده است. اکمان رابطه بین احساسات انسانی و حالت‌های چهره نمونه‌هایی را مورد بررسی و آنالیز قرار داده است که پیش‌زمینه فرهنگی متفاوتی دارند. او در نهایت یک سری احساسات اولیه و عمومی مثل ترس، تنفر و عصبانیت را معرفی کرده است. شش احساس پایه و عمومی شامل: لذت، ناراحتی، تعجب، عصبانیت، عشق و ترس می‌باشند که به عنوان احساسات اصلی و الگو تلقی می‌شوند. این احساسات در میان نمونه‌هایی با جنسیت و زمینه‌ی فرهنگی متفاوت جامع و عمومی می‌باشند. با توجه به مدل پایه‌ای احساسات، در هر لحظه‌ای اگر شخصی احساسی را تجربه کند، باید یکی از احساسات شش‌گانه را انتخاب کند که به احساس واقعی او در آن لحظه نزدیک‌تر باشد. به عبارت دیگر مدل گسسته به عنوان یک مجموعه کامل برای توصیف احساسات در نظر گرفته می‌شود.

^۱Ekman

با وجود این هنوز توافقی بین دانشمندان در مورد اینکه چه تعداد احساس پایه وجود دارد و چرا این احساسات پایه هستند، وجود ندارد. نظریه پردازان متفاوتی، احساسات متفاوت را به عنوان حس‌های پایه در نظر می‌گیرند. ولی همه‌ی آن‌ها معتقد هستند که بعضی از حس‌ها پایه‌ای‌تر و اساسی‌تر از بقیه هستند. تعدادی از نظرات، که نظریه‌پردازان حس‌های پایه‌ای را پیشنهاد می‌دهند در [۲۸] موجود است. اگرچه مدل طبقه‌بندی احساس توسط دانشمندان زیادی حمایت می‌شود، هنوز محدودیت‌هایی وجود دارد که باید در نظر گرفته شود. اولاً، اطلاعات دقیقی در مورد تعداد حس‌های پایه‌ای وجود ندارد. ثانیاً تشخیص و تفکیک احساسات مبهم است.

۲-۲-۶-۲ مدل‌های ابعادی

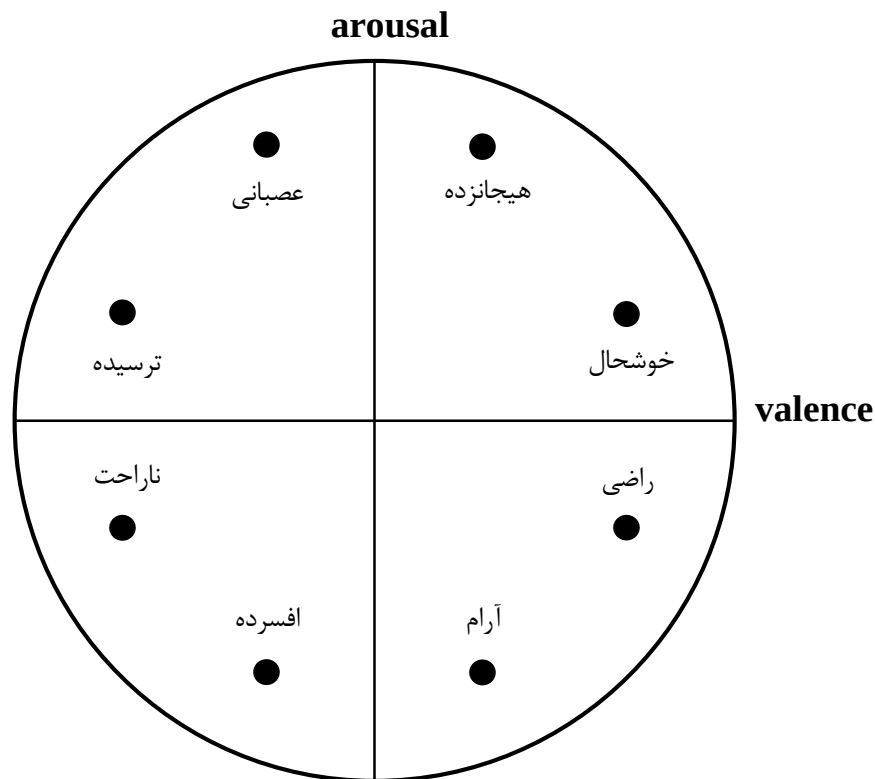
اسپنسر^۱ [۲۹] از اولین کسانی بود که احساسات را در بعدهایی در حوزه‌ی خودآگاهی معرفی کرد. در سال ۱۸۹۶، این نظریه با نام‌گذاری ۳ بعد به (۱) خوشایندی- ناخوشایندی (۲) آرامش-تنش (۳) آرام-هیجان‌زده توسعه یافت. با توجه به مطالعه انجام شده در [۳۰] یک مدل سه بعدی شامل arousal, valence, dominance ارائه شد:

- Valence کیفیت یک احساس را بیان می‌کند. (از خوشایند تا ناخوشایند)
- Arousal به سطح کمی فعال‌سازی اشاره می‌کند (از آرام تا هیجان‌زده)
- Dominance به درجه کنترلی که یک شخص در یک وضعیت خاص دارد مربوط می‌شود (از ضعیف تا قوی)

بین همه‌ی این‌ها valence و arousal پرکاربردترین ابعاد هستند. این واقعیت توسط نتایج مطالعات [۳۱] تأیید شده است که arousal و valence به علت تغییرات زیاد در اندازه‌گیری‌های احساسات بکار

^۱Spencer

می‌آیند، در حالی که dominance تغییرات کمی در احساسات مختلف دارد. یک مدل ابعادی احساس در شکل ۷-۲ نمایش داده شده است. با توجه به این مدل، احساسات با توجه با جایگاهشان در فضای دوبعدی مشخص می‌شوند. فضای دوبعدی دو محور valence (محور افقی) و arousal (محور عمودی) دارد. در مقایسه با مدل‌های احساسی گسسته، مدل‌های ابعادی یک مزیت اساسی دارند. در این مدل‌ها تجربه‌های احساسی بدون تأثیر محدودیت‌های مرزهای احساس‌های گسسته ارزیابی می‌شوند و همچنین همه حالات احساسی قابل تشخیص هستند. به علاوه زمانی که احساسات توسط جایگاه خود در فضای چند بعدی توصیف می‌شوند دیگر نیازی نیست که حالت‌های احساسی را نام‌گذاری کرد. اما نبود اسم نشان می‌دهد که جایگاه‌ها در فضای چند بعدی نسبت به مدل‌های گسسته دارای توصیف‌گری و تمایز پایین‌تری هستند. در این تحقیق از مدل ابعادی استفاده شده است.



شکل ۷-۲. نمونه‌ای از مدل ابعادی احساس

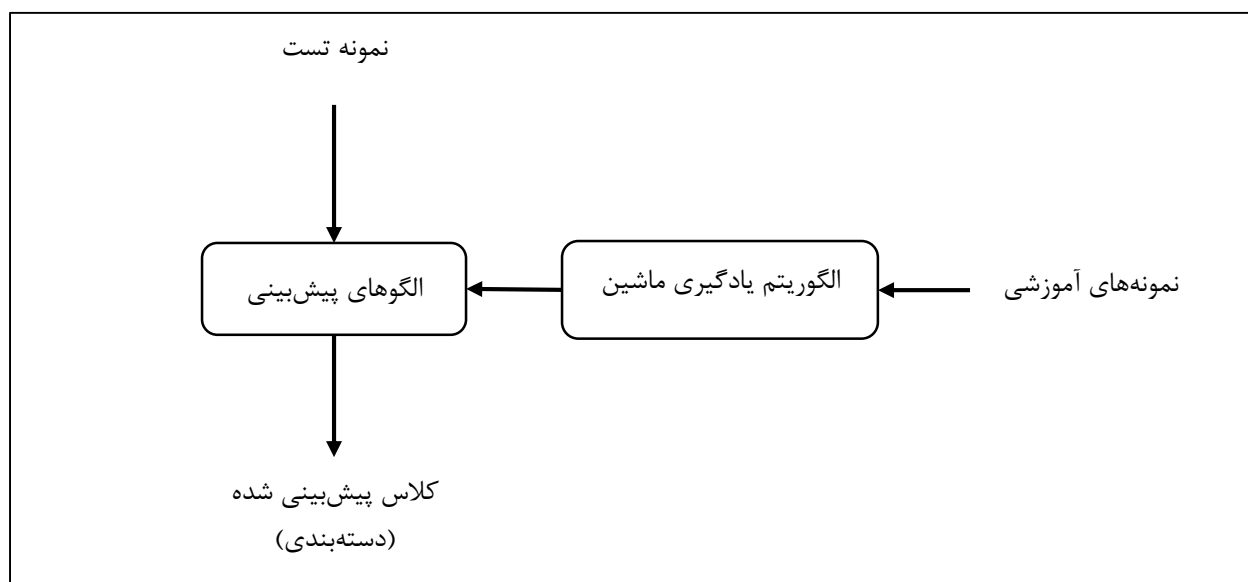
۲ - ۲ - ۷ چالش‌های فنی

اولین و مهم‌ترین قدم در طراحی سیستم‌های ارزیابی احساسات، تعریف احساساتی است که برای سیستم قابل شناسایی باشند. اما در حال حاضر یک مفهوم قابل تعریف و تفسیر برای احساس وجود ندارد و همچنان یک موضوع تحقیق در علم روانشناسی می‌باشد. در کنار توصیف‌های متعدد از احساسات، تفسیر، درک و بیان هر احساس به علت شرایط اجتماعی متفاوت بطور چشمگیری از فردی به فرد دیگر متغیر است. برای مثال لبخند یک شخص در یک دورهمی اجتماعی می‌تواند نتیجه خوشحالی واقعی فرد یا یک ژست مؤدبانه یا به عنوان یک ماسک برای پوشاندن احساسات واقعی باشد. شناسایی احساسات واقعی در محیط‌های طبیعی مشابه شرایط ذکر شده در جمله قبل، یک پروژه تحقیقاتی چالش‌برانگیز است. مدل‌های متعددی برای درک احساسات پیشنهاد شده است و هم‌چنین روش‌های متعددی برای تعمیم دادن تعریف احساسات با توجه به شرایط مختلف فرهنگی و تحصیلات افراد پیشنهاد شده است. امکان‌پذیر بودن این روش‌ها و مدل‌ها هنوز موضوع بحث و گفت‌وگو می‌باشند اما بیشتر تحقیقات تجربی، اثرگذاری مدل‌هایی مثل مدل ابعادی یا مدل احساسات گسسته را نشان می‌دهند.

در حوزه‌ی تشخیص احساسات، در بیست سال گذشته روش‌های متعددی بوجود آمده‌اند که از صدا، تصویر (حالت چهره)، آنالیز حرکات بدن و سیگنال‌های فیزیولوژیکی استفاده می‌کنند. البته در مقایسه با الگوهای شناختی دیگر و مسئله‌های یادگیری ماشین نتایج ناامیدکننده هستند. در حالی که نرخ صحت در مسائل تشخیص احساس بین ۶۰ درصد تا ۸۰ درصد بسیار متداول هستند [۲۷]، در بیشتر مسئله‌های شناختی دیگر درصدهای بالای ۹۰ و ۹۵ گزارش شده است. دلیل نرخ صحت پایین‌تر این الگو از الگوهای دسته‌بندی دیگر، پیچیدگی و نگاشت چندگانه بین نشانه‌ی دریافت شده و حالت احساسی واقعی می‌باشد.

هنوز مسائل چالش‌برانگیز زیادی وجود دارند که در کنار بهبود دقت شناسایی احساسات باید حل و فصل

شوند. تعریف برچسب درست برای سیستم شناسایی احساسات یک مشکل چالش برانگیز دیگر است. برچسب درست به برچسب‌های شناخته شده برای سیگنال‌های ورودی (نمونه‌ها) اشاره می‌کند به معنای اینکه هر حالت احساسی به کدام نمونه تعلق دارد که در طراحی سیستم‌های یادگیری ماشین مورد استفاده قرار می‌گیرند (شکل ۲-۸). برای یک مسئله یادگیری ماشین با نظارت (برچسب‌ها برای سیستم شناخته شده هستند)، برچسب‌گذاری صحیح نمونه‌های ورودی یک مسئله مهم و ضروری می‌باشد [۳۳]. یک الگوریتم یادگیری ماشین تلاش می‌کند الگویی را پیدا کند که با نمونه‌های مشاهده شده از هر کلاس سازگار باشد و سپس از این الگوی تعیین شده و قوانین تصمیم‌گیری برای پیش‌بینی کلاسی که نمونه‌های جدید به آن تعلق دارند استفاده می‌کند. اگر داده‌ای اشتباهی برچسب‌گذاری شده باشد، الگوی مخصوص هر کلاس در بین مشاهدات مبهم و گنگ می‌شود و در نتیجه اشتباهات یادگیری بیشتری رخ می‌دهد و فرآیند یادگیری به شکست می‌انجامد. بیشتر تحقیقات اخیر بر روی شناسایی احساسات بر روی پایگاه داده‌هایی طراحی شده‌اند که روش‌های استنباط و بیرون کشیدن احساسات متفاوتی دارند، به‌عنوان مثال روش‌های مبتنی بر حالت (چهره) بازیگران حرفه‌ای، مشاهده تصاویر یا کلیپ‌های صوتی که احساس موردنظر را برمی‌انگیزاند. اما تعداد کمی از کارها نشان داده‌اند که به خوبی می‌توانند احساسات موردنظر را به نمایش بگذارند. راه‌حلی وجود ندارد که با آن بتوان داده‌هایی را که درست طبقه‌بندی نشده‌اند را تعیین کرد. این واقعیت چالش بزرگی را برای ساخت یک مدل داده (اصولاً آماری) برای مسائل شناسایی را بوجود آورده است [۳۳].



شکل ۲-۸. طرح کلی یک مسئله یادگیری ماشین

در حال حاضر بیشتر نتایج تحقیقات از محیط‌های کنترل شده بدست می‌آیند. در زندگی واقعی، جمع‌آوری داده‌ها با مشکلاتی مثل در دسترس بودن تصاویر چهره یا سیگنال‌های شنیداری در محیط‌های شلوغ و پرسروصدا مواجه است. اخیراً نتایج تحقیقات نشان داده است که سیگنال‌های فیزیولوژیکی مثل سیگنال‌های مغزی (EEG)، سیگنال‌های قلبی (ECG)^۱، پاسخ برقی پوست (GSR)^۲، پالس حجم خون (BVP)^۳ ابزارهایی برای تشخیص حالت‌های احساسی می‌باشند. روش مطالعه به صورتی است که تغییرات ایجاد شده در این سیگنال‌ها در نتیجه‌ی بروز یک احساس خاص با استفاده از تکنیک‌های تشخیص الگو مناسب مورد بررسی و مطالعه قرار می‌گیرد. نتایج تعدادی از روش‌های مرتبط در تحقیق [۳۴] آورده شده است. لازم به ذکر است که تغییرات مشاهده شده در این سیگنال‌ها می‌تواند در نتیجه ایجاد عواملی که به احساسات مربوط نیستند، نیز تولید شود. برای مثال با افزایش فعالیت‌های فیزیکی ضربان قلب افزایش پیدا می‌کنند.

^۱Electrocardiogram

^۲Galvanic Skin Response

^۳Blood Volume Pulse

دمای بالای پیرامون می‌تواند رسانایی گرمایی پوست را افزایش دهد، در حالی که هیچ ترس یا هیجانی وجود ندارد. طراحی سیستمی برای نظارت‌های عملی و طولانی چندروزه به منظور جمع‌آوری سیگنال‌های متعددی که شبکه‌های حسگر نیاز دارند، غیرعملی و نشدنی می‌باشد. برای مثال برای جمع‌آوری سیگنال‌های جانبی مثل (GSR)، فرد باید یکی از انگشتان خود را ثابت نگه دارد تا بتوان سیگنال دریافت نمود. این امر با استفاده‌ی همزمان دست‌ها تداخل دارد و غیرکاربردی می‌باشد. بنابراین روش‌های جایگزینی برای ثبت مداوم سیگنال‌های حسی مورد نیاز هستند که کمتر حالت مداخله‌گرانه داشته باشند. مسئله دیگر زمان پاسخگویی می‌باشد. منظور از زمان پاسخ‌گویی زمانی است که یک نشانه احساسی نیاز دارد تا تأثیر خود را بر روی سیگنال فیزیولوژیکی انتخاب شده اعمال نماید. برای مثال، وقتی حالت‌های احساسی تغییر می‌کنند سیگنالی که در سیستم عصبی مرکزی^۱ (CNS) بوجود می‌آید در حدود چند میلی‌ثانیه تغییر می‌کند، در حالی که این فرآیند زمان بیشتری را نیاز دارد تا در یک سیگنال جانبی مثل GSR اثر خود را اعمال نماید. سیگنال‌هایی که از سیستم عصبی مرکزی (همان EEG) بوجود می‌آیند، ابزار جایگزین و بالقوه‌ای برای آنالیزهای تشخیص احساسات می‌باشند زیرا این سیگنال‌ها کمتر تحت تأثیر فعالیت‌های فیزیکی قرار گرفته‌اند. در این تحقیق از سیگنال‌های EEG برای تشخیص احساسات استفاده کردیم.

^۱Central nerve system

فصل ۳

دانش پیش‌زمینه و کارهای

گذشته

این فصل در ابتدا دانش پیش‌زمینه‌ای مربوط به سیگنال EEG را با هدف معرفی برخی از اصطلاحات و اطلاعات مربوط به این تحقیق ارائه می‌دهد. به دلیل اینکه سیگنال‌های EEG بیانگر فعالیت الکتریکی نورون‌ها می‌باشند، آناتومی و ساختار عصبی مغز انسان را توضیح می‌دهیم. این بخش شامل فعالیت‌های قسمت‌های مختلف مغز انسان و مروری بر EEG و ماهیت آن می‌شود. پس از آن به مقوله دسته‌بندی سیگنال EEG پرداخته و کارهای مرتبط انجام شده در زمینه دسته‌بندی احساسات با استفاده از سیگنال EEG بیان می‌شود.

۳-۱ مغز انسان

یکی از ارگان‌های مهم بدن، مغز انسان است. مغز عضو مرکزی سیستم عصبی است که فعالیت ارگان‌های مختلف بدن را کنترل می‌کند. در این قسمت، ابتدا ساختار آناتومی مغز و فعالیت‌های آن‌ها توضیح می‌دهیم. سپس بر روی تولید فعالیت‌های الکتریکی که از روی پوست سر توسط سیگنال EEG قابل ضبط می‌باشد تمرکز می‌کنیم.

۳ - ۱ - ۱ ساختار مغز و فعالیت‌های آن‌ها

از نظر آناتومی مغز به سه قسمت اصلی تقسیم می‌شود: مخ^۱، مخچه^۲، ساقه‌ی مغز^۳ [۳۵] (شکل ۱-۳) در ادامه توضیح مختصری در رابطه با سه قسمت معرفی شده داده می‌شود.

I. مخ

بزرگترین و مهم‌ترین قسمت مغز انسان است و بطور عمومی با فعالیت‌های مغز مربوط به افکار، احساسات و اعمال حرکتی مرتبط است. بیرونی‌ترین لایه مخ از بافت‌های عصبی تشکیل شده است که به عنوان قشر مخ^۴ شناخته می‌شود. مخ از دو نیمکره چپ و راست تشکیل شده است و هر نیمکره به چهار بخش یا لوب تقسیم می‌شود: پیشانی^۵، آهیانه^۶، پس سری^۷، گیجگاهی^۸ [۳۶].

- لوب پیشانی با شخصیت، احساسات، حل مسئله، شبیه‌سازی ذهنی، استدلال، برنامه‌ریزی، گفتار و حرکت مرتبط است.
- لوب آهیانه مسئول شناخت، جهت‌یابی و حرکت است.
- لوب پس سری مسئول پردازش بصری است.

^۱Cerebrum

^۲Cerebellum

^۳Brainstem

^۴Cerebral cortex

^۵Frontal

^۶Parietal

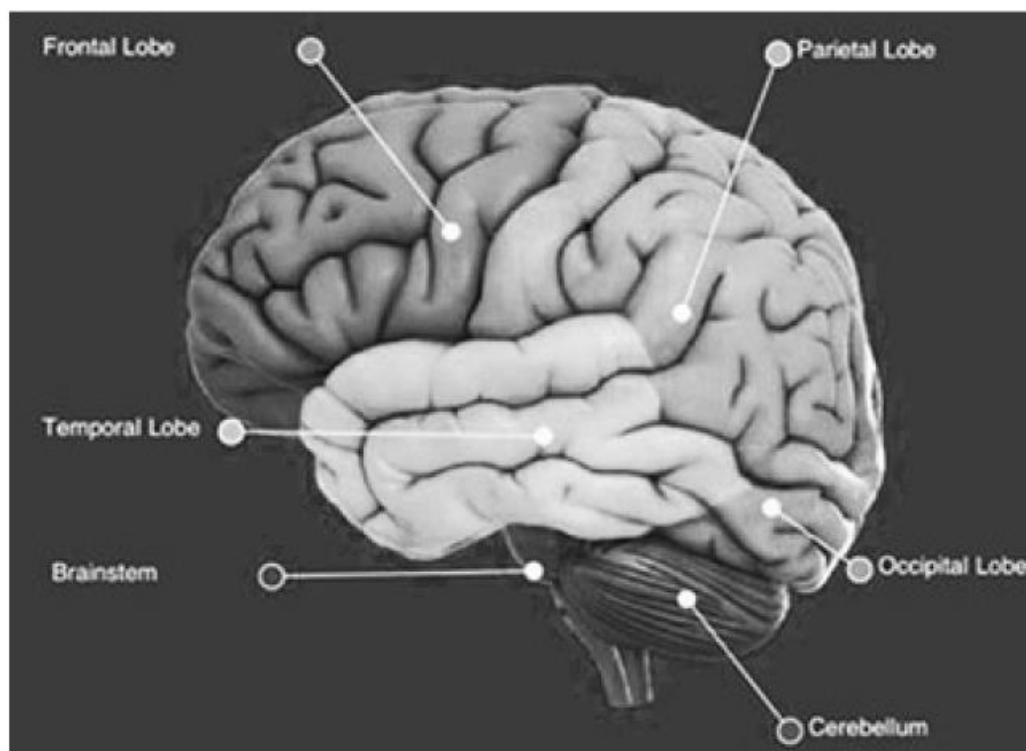
^۷Occipital

^۸Temporal

- لوب گیجگاهی مربوط به شناخت محرک شنوایی، گفتار، ادراک و خاطرات می‌باشد.

II. مخچه

مخچه در بخش پایینی پشت سر قرار دارد که به دو نیمکره تقسیم می‌شود و همچنین دومین ساختار بزرگ مغز است. مخچه شامل بیش از نیمی از نورون‌های مغز است که یکی از نواحی حسی مغز است که مسئول کنترل حرکت، ادراک حسی^۱ و هماهنگی است. همچنین با حرکات ارادی ماهیچه‌ای، مهارت‌های حرکتی ظریف، طرز ایستادن و تنظیم تعادل مرتبط است [۳۶].



شکل ۳-۱. نواحی مختلف مغز [۳۵]

III. ساقه‌ی مغز

^۱Sensory perception

ساقه‌ی مغز در پایین سر قرار گرفته است و مخچه را به نخاع^۱ متصل می‌کند. ساقه‌ی مغز کنترل پنل اصلی بدن است. فعالیت‌های حیاتی بدن را که شامل تنفس، هوشیاری، حرکت چشم و دهان، مخابره کردن پیام‌های حسی (درد، گرما و غیره)، ضربان قلب، فشار خون و گرسنگی می‌باشد کنترل می‌کند.

۳ - ۱ - ۲ فیزیولوژی عصبی مغز انسان

مغز انسان از حدود ۱۰۰ میلیون سلول عصبی به نام نورون تشکیل شده است که شارژ الکتریکی مغز توسط این نورون‌ها نگه‌داشته می‌شود. نورون‌ها خصوصیات مشترک و قسمت‌هایی مشابه سلول‌های دیگر دارند ولی تنها ویژگی الکتروشیمیایی به آن‌ها اجازه انتقال سیگنال‌های الکتریکی و پیام‌هایی به یکدیگر از فواصل زیاد را می‌دهد. نورون‌ها سه قسمت اصلی دارند: جسم سلولی^۲، آکسون^۳ و دندریت^۴ [۳۶].

نورون‌ها در حالت استراحت پتانسیل اندکی از دیگر نورون‌ها دریافت می‌کنند. در سطح غشاء سلولی آن‌ها اختلاف پتانسیل ثابتی وجود دارد به گونه‌ای که پتانسیل قسمت داخلی حدود ۵۰ میکرو ولت نسبت به سطح غشاء پایین‌تر است. این اختلاف پتانسیل که پتانسیل استراحت نامیده می‌شود در نتیجه تفاوت غلظت یون‌های مثبت و منفی در داخل و خارج غشاء سلولی ایجاد می‌شود. از آنجایی که غشاء سلولی مرتباً در حال دریافت تحریکات الکتریکی از دیگر نورون‌هاست، هیچ‌گاه در حالت سکون قرار نمی‌گیرد. نورون بطور

^۱Spinal cord

^۲Cell body

^۳Axon

^۴Dendrite

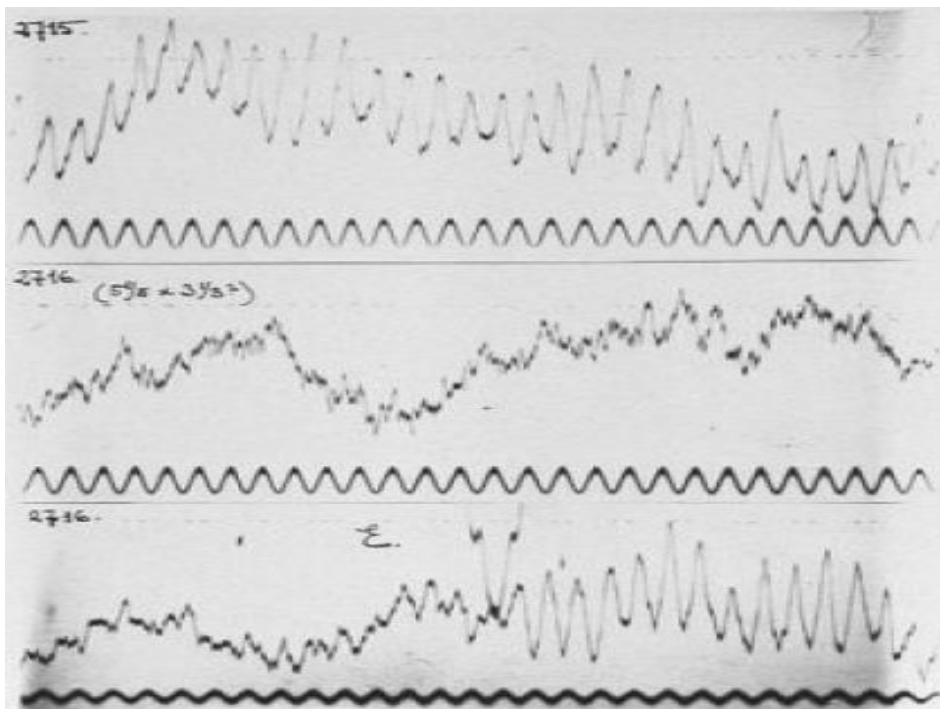
مداوم تحریک می‌شود که برانگیخته شدن نورون باعث تغییر پتانسیل غشاء می‌شود. بدین ترتیب سیگنال‌هایی با شکل موج‌های خاص در سراسر سلول عصبی انتشار می‌یابند که این مسئله خود می‌تواند باعث تحریک نورون‌های مجاور شود. پتانسیل الکتریکی که در سطح سر قابل ثبت است، اثر تحریک یک نورون واحد نیست و برای اینکه یک موج الکتریکی قابل ثبت باشد باید بخش بزرگی از بافت‌های عصبی بطور هم‌زمان تحریک شده و جریان الکتریکی ایجاد کنند. وجود میلیون‌ها نورون که هر کدام با شدت‌های مختلف تحریک می‌شوند، شکل موج‌های بسیار متنوعی را در سطح مجمله ایجاد می‌کنند که می‌توان آن‌ها را توسط الکترودهای پوستی دریافت و ضبط کرد. یکی از روش‌های ضبط این فعالیت‌ها استفاده از EEG می‌باشد که در ادامه با آن آشنا می‌شویم.

۳ - ۱ - ۳ الکتروانسفالوگرافی

EEG تکنیک اندازه‌گیری پتانسیل‌هایی است که فعالیت الکتریکی مغز را منعکس می‌کند. EEG به آسانی قابل دسترس است و چگونگی فعالیت مغز در طول زمان را ثبت می‌کند. این روش بطور گسترده‌ای توسط دانشمندان و پزشکان برای مطالعه فعالیت‌های مغزی و تشخیص اختلالات عصبی مورد استفاده قرار می‌گیرد. ماشین EEG در ابتدا توسط هنس برگر در سال ۱۹۲۹ به جهان معرفی شد [۳۷]. برگر یک پزشک اعصاب و روان از دانشگاه Gena در آلمان بود که از یک کلمه آلمانی "elektrenkephalogramm" برای توصیف نمایش گرافیکی جریان‌های الکتریکی تولید شده در مغز استفاده کرد. او ادعا کرد که جریان‌های الکتریکی تولید شده بسته به وضعیت فعالیت مغز مثل خواب، بیهوشی و صرع تغییر می‌کند.

این یک ایده انقلابی بود که باعث بوجود آمدن شاخه‌ای جدید از علم پزشکی به نام فیزیولوژی عصبی شد. شکل ۲-۳ اولین داده‌های ضبط شده EEG توسط هنس برگر را نمایش می‌دهد. برگر متوجه شد که تغییرات ریتمیک این داده‌ها بر اساس وضعیت هوشیاری فرد متفاوت است.

طی آزمایش EEG تعدادی دیسک کوچک به نام الکتروود در مکان‌های مختلف روی پوست سر قرار می‌گیرند. سپس هر الکتروود به یک تقویت‌کننده (یک تقویت‌کننده به ازای هر جفت الکتروود) و یک ماشین ضبط EEG متصل است. در نهایت سیگنال‌های الکتریکی از مغز به خطوط پر موج روی صفحه کامپیوتر (برای ثبت نتایج) تبدیل می‌شوند. داده‌های EEG بسته به کاربرد آن‌ها می‌توانند با رزولوشن‌های مکانی متفاوت شامل ۱ تا ۲۵۶ الکتروود ضبط شوند. هر جفت الکتروود یک کانال را تشکیل می‌دهد. ثبت سیگنال توسط یک الکتروود EEG تک کاناله و ثبت سیگنال توسط بیشتر از یک الکتروود EEG چند کاناله نامیده می‌شود.

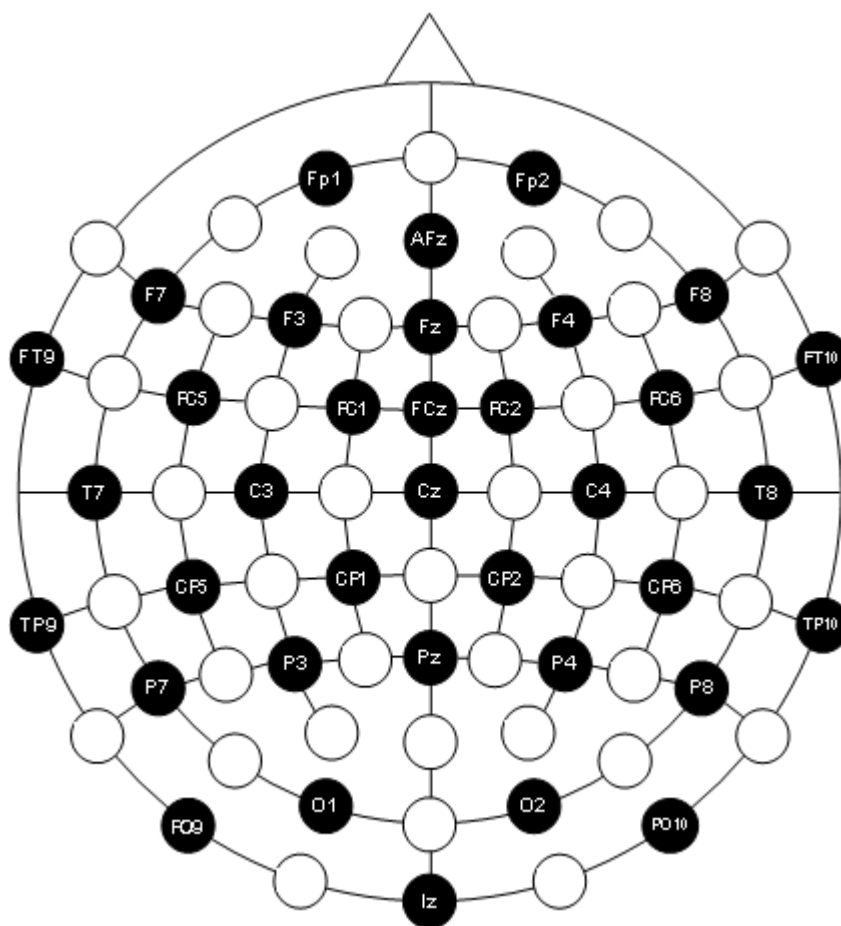


شکل ۲-۳. اولین داده‌های ضبط شده سیگنال EEG توسط هنس برگر [۳۸]

محل قرارگیری الکترودها مهم است زیرا بخش‌های مختلف قشر مخ مسئول پردازش فعالیت‌های مختلف هستند. روش استاندارد قرارگیری الکترودهای روی پوست سر سیستم بین‌المللی ۱۰-۲۰ است [۳۹]. ۱۰ و ۲۰ فاصله واقعی بین الکترودهای همسایه هستند که ۱۰٪ یا ۲۰٪ از فاصله کل جلو تا عقب یا چپ تا راست کاسه سر هستند. مکان الکترودها توسط دو نقطه تعیین می‌شود: یکی نقطه بین پیشانی و بینی^۱ که هم سطح با چشم‌ها می‌باشد و دیگری برجستگی پشت جمجمه^۲ که روی خط میانی در پشت سر قرار دارد. شکل ۳-۳ مکان الکترودها طبق سیستم بین‌المللی ۱۰-۲۰ را نشان می‌دهد. هر مکان از یک حرف برای شناسایی لوب و یک عدد برای شناسایی مکان و موقعیت نیمکره استفاده می‌کند. حروف P, T, C, F, O به ترتیب سمبل‌هایی برای آهیانه، مرکزی، گیجگاهی، پیشانی و پس سری هستند. البته باید توجه شود که "لوب مرکزی" وجود ندارد و این نام‌گذاری صرفاً برای مشخص کردن یک ناحیه خاص است. Z به الکترودهای واقع در خط میانی اشاره می‌کند. اعداد زوج به الکترودهای قرار گرفته در نیمکره راست و اعداد فرد به الکترودهای نیمکره چپ اشاره می‌کنند. از آنجایی که ولتاژ سیگنال EEG اختلاف بین ولتاژهای دو الکترودها را نشان می‌دهد، چگونگی نمایش EEG برای خواندن از ماشین EEG به روش‌های مختلفی تنظیم می‌شود. آرایش الکترودها به عنوان یک مونتاژ تلقی می‌شود. EEG می‌تواند توسط مونتاژهای زیر کنترل شود:

^۱Nasion

^۲Inion



شکل ۳-۳. سیستم بین‌المللی قرارگیری الکترود ۱۰-۲۰ [۴۰]

مونتاژ دوقطبی: یک جفت الکترود یک کانال را تشکیل می‌دهد. هر کانال (به شکل موج) اختلاف بین دو الکترود مجاور [۴۱]، [۴۲] را نشان می‌دهد. کل مونتاژ شامل چندین کانال می‌باشد. برای مثال کانال FP1-F3 نشان‌دهنده اختلاف ولتاژ بین الکترود FP1 و الکترود F3 می‌باشد. کانال بعدی در مونتاژ F3-C3 اختلاف ولتاژ بین F3 و C3 را نشان می‌دهد.

مونتاژ ارجاعی: هر کانال نشان‌دهنده اختلاف بین یک الکترود مشخص و یک الکترود مرجع از پیش تعیین شده می‌باشد [۴۱]، [۴۲]. یک مکان استاندارد برای این مرجع وجود ندارد. ولی این الکترود مکانی متفاوت با الکترودهای ثبت EEG دارد.

مونتاز میانگین مرجع: خروجی‌های حاصل از همه تقویت‌کننده‌ها جمع می‌شوند و میانگین محاسبه می‌شود. این سیگنال به عنوان مرجع برای هر کانال استفاده می‌شود [۱۸].

مونتاز لاپلاسی: هر کانال نشان‌دهنده اختلاف بین یک الکتروود و میانگین وزندار الکتروودهای اطراف است [۱۸]. سیگنال‌های EEG دیجیتالی هستند و همواره در یک مونتاز (عمدتاً مونتاز ارجاعی) ذخیره می‌شوند. از آنجایی که هر مونتاز می‌تواند به صورت ریاضی از بقیه ساخته شود، EEG می‌تواند به عنوان یک ماشین EEG در هر مونتازی دلخواه نمایش داده شود.

الگوهای EEG برای درک و فهم فعالیت‌های مغز مهم و حیاتی هستند. این الگوها توسط ویژگی‌های و آزمایش باندهای فرکانسی مرتبط با فعالیت‌های ذهنی یا حالت‌های هوشیاری قابل شناسایی هستند. باندهای فرکانسی به ۵ دسته تقسیم می‌شوند. در قسمت بعدی متعارف‌ترین الگوهای سیگنال EEG را در شرایطی که افراد در حالت‌های هوشیاری، خواب، رنج کشیدن از آشفتگی مغزی و یا تجربه حالت‌های احساسی هستند را توضیح می‌دهیم.

۳ - ۱ - ۳ ماهیت ریتم‌های فرکانسی سیگنال EEG

فرکانس یکی از مهم‌ترین ویژگی‌های سیگنال‌های EEG است. ریتم‌های EEG ناشی از فعالیت گروهی نورون‌ها است که مجموع رخداد‌های سیناپسی آن به اندازه کافی بزرگ شده‌اند. بطور کلی ۵ ریتم فرکانسی عمده در EEG تعریف شده است [۴۱]، [۴۲] که در ادامه با آن‌ها آشنا می‌شویم.

- دلتا (δ) در محدوده بین ۰.۵ تا ۴ هرتز قرار می‌گیرد. بیشترین دامنه و آهسته‌ترین موج‌ها در آن قابل مشاهده است. موج دلتا با خواب سنگین یا ناهنجاری‌های شدید ذهنی در حالت بیداری مرتبط است.

- تتا (θ) در محدوده بین ۴ تا ۸ هرتز با دامنه بیشتر از ۲۰ میکرو ولت قرار می‌گیرد. تتا از استرس، مخصوصاً خستگی و ناامیدی و موارد ناخودآگاه ذهنی، الهام‌بخشی خلاقانه و مراقبه عمیق نشأت می‌گیرد.
- آلفا (α) فرکانس بین ۸ تا ۱۳ با دامنه ۵۰-۳۰ میکرو ولت را در برمی‌گیرد که بطور عمده در نواحی پشت سر ظاهر می‌شود. زمانی که چشم فرد بسته است یا در حالت استراحت قرار دارد. آلفا معمولاً به فعالیت‌های شدید ذهنی، استرس و تنش مربوط می‌شود. فعالیت‌های آلفا ضبط شده از نواحی حسی و حرکتی به عنوان میو (μ) شناخته می‌شوند.
- بتا (β) در محدوده فرکانسی ۱۳ تا ۳۰ قرار می‌گیرد که معمولاً با دامنه پایین و فرکانس‌های متغیر متقارن در دو طرف ناحیه جلویی سر دیده می‌شود. زمانی که مغز برانگیخته می‌شود و با فعالیت‌های ذهنی درگیر می‌شود امواج بتا تولید می‌گردد. امواج بتا معمولاً با چیزهای فعال، توجه فعال یا تمرکز روی دنیای بیرون یا حل مشکلات واقعی مرتبط است.
- گاما (γ) در محدوده فرکانسی ۳۰ هرتز به بالا قرار می‌گیرد. بیشترین فرکانس این امواج حدود ۸۰ یا ۱۰۰ هرتز می‌باشد که با کارهای مختلف شناختی و فعالیت‌های حرکتی مرتبط است.

۲-۳ مروری بر دسته‌بندی سیگنال EEG

دسته‌بندی سیگنال EEG نقش مهمی در تحقیقات علمی و پزشکی دارد. دسته‌بندی سیگنال EEG در فرآیندهای شناختی و تشخیص از جمله تشخیص بیماری‌های مغزی، احساسات و غیره بسیار مهم می‌باشد. در آزمایش‌هایی که به بررسی عملکرد مغز از طریق اندازه‌گیری سیگنال EEG می‌پردازند، حجم زیادی از

داده تولید می‌شود. برای رسیدن به بهترین عملکرد ممکن، بایستی از تعداد کمتری داده که خصوصیات سیگنال را توصیف می‌کنند، استفاده کرد. طبق اصول شناسایی الگو، فرآیند تشخیص شامل دو مرحله اصلی می‌باشد: استخراج ویژگی و دسته‌بندی. در مرحله استخراج ویژگی، ویژگی‌ها و مشخصاتی که بر اساس آن هر سیگنال دسته‌بندی می‌شود باید تعریف شود. در مرحله دسته‌بندی بر اساس ویژگی‌های انتخاب شده سیگنال، دسته‌بند کلاس متعلق به سیگنال را تعیین می‌کند. از خروجی سیستم دسته‌بند می‌توان برای پی بردن به اینکه چه اتفاقی در دنیای واقعی باعث بوجود آمدن سیگنال ورودی شده است، استفاده کرد.

۳-۲-۱ تشخیص احساس با دسته‌بندی سیگنال EEG

این تحقیق روی دسته‌بندی سیگنال EEG در زمینه تشخیص احساسات تمرکز می‌کند. همانطور که در قسمت ۲-۲-۷ گفته شد دسته‌بندی احساسات هنوز یک مسئله چالش برانگیز برای محققان می‌باشد، زیرا تعریف درست و واحدی از حالت‌های احساسی مختلف وجود ندارد. همچنین در مسئله دسته‌بندی احساسات، نتایج در بازه گسترده‌ای تغییر می‌کند زیرا هر فردی تعریف و دسته‌بندی جداگانه‌ای از احساسات را در نظر می‌گیرد. الگوریتم‌های دسته‌بندی زیادی در زمینه دسته‌بندی احساسات مورد استفاده قرار گرفته‌اند، از جمله: ماشین بردار پشتیبان^۱ (SVM)، K نزدیکترین همسایگی^۲ (KNN)، شبکه عصبی مصنوعی^۳ (ANN)، بیز ساده^۴ شبکه بیزین^۵ (BNT)، آنالیز خطی مجزا^۶ (LDA). برخی از تحقیقات انجام شده در

^۱Support vector machine

^۲K nearest neighbor

^۳Artificial neural network

^۴Naïve bayes

^۵Bayesian network

^۶Linear discriminant analysis

زمینه دسته‌بندی سیگنال‌های EEG جهت تشخیص احساسات آورده شده است. که بر اساس نوع دسته‌بند مورد استفاده طبقه‌بندی شده‌اند.

- دسته‌بندی با استفاده از SVM

محققان مقاله [۴۳] از مدل سه بعدی valence-arousal-dominance برای تشخیص و ارزیابی احساسات استفاده کرده‌اند. محرک‌های صوتی منتخب از پایگاه داده IAPS برای ضبط سیگنال EEG استفاده شده است و شرکت‌کنندگان با خود ارزیاب SAM احساسات خود را ارزیابی کردند. در روش ارائه شده تبدیل فوریه گسسته اعمال و پنج باند فرکانسی استخراج شدند. بر اساس نتایج آزمایش‌ها از دو کانال FC6 و F8 برای استخراج ویژگی استفاده شد. چگالی طیف توان (PSD) برای باندهای فرکانسی محاسبه سپس نسبت توان باندهای آلفا به بتا به عنوان ویژگی به دسته‌بند SVM داده شد. نرخ دسته‌بندی این روش بین ۷۳/۶۴ تا ۷۵/۱۷ درصد تغییر می‌کند.

کلسترا و همکارانش [۴۴] روشی برای دسته‌بندی با استفاده از سیگنال EEG و سایر سیگنال‌های فیزیولوژی شامل GSR، BVP، ریتم تنفس، دمای پوست و EMG ارائه داده‌اند. آن‌ها از مدل دوبعدی احساس valence-arousal برای ارزیابی احساسات استفاده کردند. در آزمایش ترتیب داده شده توسط کلسترا و همکارانش، افراد بعد از تماشای ویدیوها احساسات خود را با معیارهای valence، arousal و liking ارزیابی نمودند. در مرحله پیش‌پردازش فیلتر میان‌گذر با فرکانس قطع بین ۰.۵ تا ۳۵ هرتز اعمال شد، سپس ویژگی‌های چگالی طیف توان PSD و الگوی مکانی مشترک CSP^۲ استخراج و دسته‌بندی با استفاده از دسته‌بند SVM انجام شد. نتایج حاصل از دسته‌بند SVM با ویژگی‌های PSD و CSP برای دسته‌بندی دو کلاسه معیار

^۲power spectral density

^۳Common spatial pattern

valence به ترتیب ۵۸/۸ و ۵۸/۸ ، arousal ۵۱/۹ و ۵۵/۷ و liking ۴۹/۴ و ۴۸/۸ می‌باشد.

روش پیشنهادی [۴۵] با استفاده از ویژگی‌ها آنتروپی تقریبی (ApEn)^۱ و آنتروپی موجک (EW)^۲ ارائه شده است. در این تحقیق از پایگاه داده IAPS برای برانگیختن احساسات و ۵ کانال FP1، FP2، T3، T4 و Pz برای ضبط سیگنال EEG استفاده شده است. در روش پیشنهادی آن‌ها ابتدا فیلتر میان گذر با فرکانس‌های قطع ۰/۲ و ۳۵ عمل شد و بعد از آن ۴ باند فرکانسی دلتا، تتا، آلفا و بتا استخراج شد. در مرحله استخراج ویژگی ویژگی‌های آنتروپی تقریبی و آنتروپی موجک محاسبه و نرمال شدند. از SVM با تابع پایه شعاعی (RBF)^۳ برای دسته‌بندی استفاده شد که نرخ دسته‌بندی به ۷۳/۲۵ برای دسته‌بندی ۲ احساس آرام خنثی و هیجان‌زده منفی رسید.

پژوهشگران در مقاله [۴۶] از کانال‌های FP1، AF3، F7، P3، P7، Pz، PO3، O1، CPz، C4، T8 و FC6 از پایگاه داده DEAP استفاده نموده‌اند. باند گامای بدست آمده از تبدیل موجک برای استخراج ویژگی انتخاب شده است. پارامترهای آنتروپی، همبستگی متقابل و ضرایب مدل AR به عنوان ویژگی ورودی دسته‌بندی SVM چند هسته‌ای در نظر گرفته شده است. نتایج برای ترکیب کانال‌های مختلف بررسی شده و ترکیب هفت کانال C4، CP2، O1، PO3، Pz، P3 و P7 بهترین نتیجه را ارائه کرده‌اند. نرخ دسته‌بندی برای این هفت کانال ۹۴/۰۹ بدست آمده است.

محققین [۴۷] دسته‌بندی چهار حالت احساسی در فضای دو بعدی arousal و valence شامل خوشحالی، خنثی، ناراحتی و ترس را انجام داده‌اند. برای پیش‌پردازش از فیلتر میان‌گذر باترورث با فرکانس قطع ۰/۵،

^۱Approximate

Wavelet entropy

^۳Radial basis function

تا ۱۰۰ هرتز اعمال شده است. ویژگی‌های آنتروپی شانون و آنتروپی رنی از باندهای بدست آمده از تبدیل موجک استخراج شده است. در پایان ۸۴/۷۹ برای دسته‌بندی چهار کلاس توسط SVM با کرنل تابع پایه شعاعی^۱ RBF بدست آورده‌اند.

در تحقیق [۴۸] از ۱۸ کانال پایگاه داده IAPS برای دسته‌بندی پنج حالت احساسی شامل ناراحت، ترسیده، خوشحال، آرام و خنثی استفاده نموده‌اند. پارامترهای اچ جورث^۲ به عنوان ویژگی از باندهای فرکانسی بدست آمده از تبدیل فوریه سریع محاسبه شده است. در روش آن‌ها نتایج دسته‌بندی دو کلاسی ۷۰٪ بدست آمده است. نتایج بدست آمده برای دسته‌بندی نمودن پنج کلاس مناسب نبوده و بیانگر دشوار بودن چنین دسته‌بندی می‌باشد.

• دسته‌بندی با استفاده از SVM و LDA

محققان در [۴۹] ۴ حالت احساسی مثبت/هیجان‌زده، منفی/هیجان‌زده، مثبت/آرام، منفی/آرام را مورد بررسی قرار دادند. پایگاه داده مورد استفاده آن‌ها کلیپ‌های صوتی IAPS می‌باشد. جهت استخراج باندهای فرکانسی تبدیل فوریه بکار گرفته شد و دو باند آلفا و بتا برای استخراج ویژگی مورد استفاده قرار گرفتند. نسبت توان باند بتا به آلفا در ۴ کانال AF3، AF4، F3 و F4 بعنوان arousal محاسبه شد. سپس مقدار valence از اختلاف دو نسبت توان باند آلفا در F3 به توان باند بتا در F4 و توان باند آلفا در F4 به توان باند بتا در F3 بدست آمد. دسته‌بندی valence و arousal به دو کلاس توسط دو دسته‌بند SVM و LDA انجام شد. نتایج بدست آمده حاکی از بهتر بودن عملکرد SVM نسبت به LDA بود. نرخ دسته‌بندی ۷۷/۸۲ و

^۱Radial basis function

^۲Hjorth

۸۸/۱۱ به ترتیب برای valence و arousal توسط SVM بدست آمد.

تحقیق [۵۰] دسته‌بندی سه حالت احساسی مثبت و منفی و خنثی را انجام داده است. در این تحقیق از چهار کانال F3، F4، FP1 و FP2 استفاده شده است. برای پیش‌پردازش فیلتر میان‌گذر ۴ تا ۳۲ هرتز اعمال شد و باندهای فرکانسی تتا و آلفا و بتا برای استخراج ویژگی انتخاب شدند. دو نوع ویژگی در این تحقیق استخراج شده است: ویژگی PSD با استفاده از روش والش و مؤلفه‌های EMD. از دو دسته‌بند LDA و SVM برای دسته‌بندی استفاده شد که نتایج SVM بالاتر از روش LDA بود. همچنین نتایج بدست آمده از EMD پیچیدگی محاسباتی و زمان کمتری داشت.

• دسته‌بندی با استفاده از ANN

پژوهش [۵۱] از پایگاه داده IAPS برای دسته‌بندی سه احساس هیجان‌زده-مثبت، هیجان‌زده-منفی و آرام بودن استفاده کرده است. پس از تماشای تصاویر از شرکت‌کننده‌ها خواسته شد تا احساسات را ارزیابی نمایند. در این تحقیق از هفت کانال F2، F1، FC2، FC1، Pz، Fz و Cz استفاده شده است. ویژگی‌های ERP استخراج شده شامل ویژگی‌های P300، P200، N100، N200 و P100 می‌باشد. P100 و N100 به ترتیب اولین اوج^۱ مثبت و منفی مشاهده شده از ۸۰ تا ۱۲۰ میلی‌ثانیه می‌باشند. P200 دومین اوج مثبت در بازه حدود ۲۰۰ میلی‌ثانیه می‌باشد که بین ۱۵۰ تا ۲۷۵ متغیر است. N200 اوج منفی در بازه حدود ۲۰۰ تا ۳۵۰ میلی‌ثانیه و P300 اوج مثبت مشاهده شده در بازه ۳۰۰ میلی‌ثانیه بین ۲۵۰ تا ۵۰۰ می‌باشد. در این تحقیق از دسته‌بند بیز ساده و ANN برای دسته‌بندی احساسات در محور Valence استفاده شد. نتایج

^۱Peak.

حاصل از دسته‌بندی توسط ANN با یک لایه مخفی و دو لایه مخفی به ترتیب ۵۷٪ و ۱۰۰٪ و نتیجه بدست آمده توسط بیز ساده ۵۶٪ بود.

ریچد و همکارش [۵۲] از پایگاه داده DEAP و چهار حالت احساسی آرام/منفی، آرام/مثبت، هیجان‌زده/منفی و هیجان‌زده/مثبت برای آنالیز احساسات استفاده کردند. در روش آن‌ها از فیلتر لاپلاسین^۱ و فیلتر میان‌گذر با فرکانس‌های قطع ۴ و ۴۵ هرتز استفاده شد. پس از پیش‌پردازش، از تبدیل موجک برای استخراج ویژگی پنج باند فرکانسی از سیگنال‌ها استفاده گردید. از دو باند فرکانسی آلفا و تتا استخراج شده از کانال FP1 ویژگی‌های آنتروپی و انرژی محاسبه شد. میانگین نتایج بدست آمده از دسته‌بند ANN با سه لایه مخفی برای دسته‌بندی چهار حالت احساسی برابر ۹۰٪ بود.

روشی که در مقاله [۵۳] ارائه شده است؛ روشی مبتنی بر ویژگی‌های حوزه زمان - فرکانس بدست آمده از تبدیل موجک برای تشخیص مقدار valence از سیگنال‌های EEG می‌باشد. در این مقاله با استفاده از الگوریتم ژنتیک بهترین حالت برای انتخاب کانال‌ها بدست آمده است. ویژگی‌های cross level wavelet feature groups محاسبه و با استفاده از دسته‌بند شبکه عصبی احتمالاتی دسته‌بندی شد. در روش آن‌ها نرخ صحت ۹۸٪ و ۹۰٪ به ترتیب برای دسته‌بندی دو کلاسه و سه کلاسه valence بدست آمد.

• دسته‌بندی با استفاده از KNN

پژوهشگران [۵۴] از پایگاه داده DEAP برای دسته‌بندی دو حالت احساسی استرس و آرام بودن با استفاده از معیارهای Valence و Arousal استفاده کرده‌اند. در روش پیشنهادی آن‌ها ویژگی‌های آماری میانگین و انحراف معیار و همچنین PSD از کانال‌های FP1، FP2، F3 و F4 محاسبه شدند. نتایج بدست آمده از

^۱Laplace

دسته‌بند KNN با استفاده از ویژگی PSD برابر ۷۰/۱ می‌باشد.

- دسته‌بندی با استفاده از KNN ، LDA و QDA

در مقاله [۵۵] برای دسته‌بندی valence و arousal به دو و سه دسته از پایگاه داده DEAP استفاده شده است. ضرایب رگرسیون خودکار AR بعنوان ویژگی محاسبه شده‌اند. انتخاب ویژگی با استفاده از روش انتخاب رو به جلو ترتیبی (SFS) انجام شده است. سه دسته‌بند KNN، LDA و QDA برای دسته‌بندی استفاده شده و با یکدیگر مقایسه شده‌اند. نتایج بدست آمده با استفاده از انتخاب ویژگی SFS ۱۰ تا ۱۵ درصد افزایش در نرخ دسته‌بندی نسبت به حالت بدون انتخاب ویژگی داشته است، همچنین دسته‌بند KNN از دو دسته‌بند دیگر نتایج بهتری ارائه می‌دهد. بهترین نتایج بین ۷۲/۳۳ و ۷۴/۲ برای دسته‌بندی دو کلاسه valence و arousal و ۶۱/۱ و ۶۵/۱۶ برای دسته‌بندی سه کلاسه بدست آمده است.

در تحقیق [۵۶] دسته‌بندی پنج حالت احساسی تنفر، خوشحالی، تعجب، ترس و خنثی انجام شده است. برای حذف نویز از فیلتر لاپلاسین و ZMUV برای نرمال‌سازی داده‌ها استفاده شده است. ویژگی‌های توان، انرژی، انحراف معیار، واریانس و آنتروپی از باندهای بدست آمده از روش تبدیل موجک محاسبه و برای دسته‌بندی استفاده شده است. دسته‌بندی با استفاده از KNN و LDA به صورت اعتبارسنجی با ۵ زیر نمونه انجام شده است که نتایج بیانگر عملکرد بهتر KNN با بیشینه نرخ دسته‌بندی ۷۹/۱۷ می‌باشد.

محققان مقاله [۵۷] از محرک‌های صوتی و تصویری برای برانگیختن احساسات استفاده کردند و پنج حالت احساسی (تنفر، ترس، تعجب، خنثی و خوشحال) برای دسته‌بندی در نظر گرفتند. در روش پیشنهادی آن‌ها

^۱Sequential forward selection

پیش‌پردازش توسط فیلتر لاپلاسین و استخراج باندهای فرکانسی با استفاده از تبدیل موجک انجام شد. نتیجه بدست آمده از دسته‌بند KNN و LDA به ترتیب برابر با ۸۳/۲۶ و ۷۵/۲۱ می‌باشد.

- دسته‌بندی با استفاده از BNT

محققان در [۵۸] روشی بر مبنای ویژگی PSD ارائه دادند که حالات احساسی را در دو محور valence و arousal به دو و سه کلاس دسته‌بندی می‌کند. در روش آن‌ها نرخ دسته‌بندی برای دو کلاس valence و arousal به ترتیب ۶۶/۶ و ۶۶/۴ و برای سه کلاس valence و arousal به ترتیب ۵۳/۴ و ۵۱/۰ می‌باشد. در روش ارائه شده [۵۹] سیگنال‌های EEG به پنج باند فرکانسی تقسیم شده و PSD آن‌ها محاسبه می‌شود. سپس محاسبات بر اساس نسبت توان هر باند به توان کل محدوده فرکانسی انجام می‌شود. دسته‌بند BNT دسته‌بندی را برای ۵ حالت احساسی خوشحال، ناراحت، عصبانی، متعجب و خنثی با نرخ صحت میانگین ۴۳٪ انجام می‌دهد.

- دسته‌بندی با استفاده از بیز ساده

روش ارائه شده در [۶۰] به این صورت می‌باشد که ابتدا شش باند فرکانسی از سیگنال EEG و مقادیر توان آن‌ها به عنوان ویژگی استخراج می‌شود. همچنین ویژگی‌های مربوط به سیگنال خارجی مثل ضربان قلب، فشار خون، دما و تنفس نیز استخراج می‌شود. دسته‌بند بیز ساده با نرخ صحت ۵۴٪ قادر به دسته‌بندی احساسات بر مبنای arousal می‌باشد.

۳ - ۲ - ۲ الگوریتم یادگیری عمیق در دسته‌بندی احساسات

همانطور که گفته شد روش پیشنهادی ما استفاده از الگوریتم یادگیری عمیق در تشخیص احساسات از

سیگنال EEG می‌باشد. از آنجایی که الگوریتم یادگیری عمیق موضوع نسبتاً جدیدی در یادگیری ماشین محسوب می‌شود تعداد تحقیقات انجام شده با استفاده از آن محدود می‌باشد که در ادامه مهم‌ترین کارهای انجام شده در این زمینه آورده شده است.

در [۶۱] روشی را بر مبنای شبکه‌های باور عمیق ارائه دادند که بطور اتوماتیک قادر به استخراج ویژگی از داده‌های EEG در چهار کانال به صورت غیر نظارتی می‌باشد. شبکه باور عمیق از چندین ماشین بولترمن محدود ساخته شده است که روی هم قرار گرفته‌اند. در این تحقیق از پایگاه داده DEAP استفاده شده است. این شبکه در نهایت سه دسته‌بند برای دسته‌بندی مقادیر valence, arousal و liking به دو کلاس ارائه می‌دهد. نرخ صحت به ترتیب ۶۰٫۹، ۵۱٫۲ و ۶۸٫۴ می‌باشد.

در [۱۲] یک شبکه باور عمیق برای دسته‌بندی احساسات از سیگنال EEG ارائه شده است که شبکه باور عمیق از واحدهای کدگذار خودکار پشته‌ای^۱ (SAE) با استفاده از روش یادگیری سلسه مراتبی پیاده‌سازی شده است. چگالی طیف توان ۳۲ کانال محاسبه شده است. و هم‌چنین PCA برای پیدا کردن مهم‌ترین ویژگی‌ها بکار گرفته شده است. نرخ صحت برای دسته‌بندی valence و arousal به سه کلاس به ترتیب ۴۲٫۵۳ و ۵۱٫۰۵ می‌باشد.

^۱stacked autoencoder

فصل ۴

روش پیشنهادی

در این فصل ابتدا تعدادی از پایگاه داده‌های کاربرد در زمینه تشخیص احساسات با استفاده از سیگنال EEG را به اختصار معرفی می‌کنیم. در بخش‌های بعدی استخراج و انتخاب ویژگی توضیح داده می‌شود. در نهایت الگوریتم یادگیری عمیق که برای طبقه‌بندی احساسات مورد استفاده قرار گرفت بررسی می‌شود.

۴ - ۱ معرفی پایگاه داده‌های موجود

یکی از عوامل مهم در تحقیقات علمی و آزمایشگاهی انتخاب پایگاه داده مناسب می‌باشد. پایگاه داده‌های زیادی در زمینه تشخیص احساس وجود دارد که اکثراً شامل مواردی چون حالات چهره یا صدا هستند. ولی تعداد معدودی پایگاه داده وجود دارد که شامل سیگنال EEG باشد. در این قسمت در ابتدا دو پایگاه داده مهم و کاربرد برای تحریک احساس برای ضبط سیگنال EEG و همچنین ساخت پایگاه داده‌های شامل سیگنال EEG بطور مختصر توضیح داده می‌شود و سپس چند پایگاه داده کاربرد در زمینه تشخیص احساس که شامل سیگنال EEG باشد معرفی می‌شود:

۴ - ۱ - ۱ IAPS

سیستم بین‌المللی تصاویر احساسی^۱ (IAPS) برای تهیه و تولید یک مجموعه از محرک‌های احساسی اصولی برای تحقیقات در زمینه احساس و توجه بوجود آمده است. هدف توسعه یک مجموعه استاندارد محرک احساسی قابل دسترسی بین‌المللی از عکس‌های رنگی می‌باشد که شامل محتوا در بازه وسیعی از حوزه‌های

^۱International Affective Picture System

معنایی می‌باشد. IAPS (که EYE-APS تلفظ می‌شود) همراه با IADS، ANEW و دیگر مجموعه‌های محرک احساسی توسط مرکز احساس و توجه NIMH، CSEA دانشگاه فلوریدا برای تهیه ابزار استاندارد که جهت استفاده محققان این زمینه، تولید و توزیع شده‌اند.

وجود این مجموعه‌ها از محرک‌های احساسی که به صورت اصولی رتبه‌بندی و تحلیل شده‌اند باید بتواند:

(۱) کنترل بهتر آزمایش‌ها با انتخاب محرک احساسی را فراهم کند.

(۲) امکان مقایسه نتایج بین تحقیقات مختلف را در شرایط یکسان یا متفاوت فراهم کند.

شرکت‌کنندگان، دانشجویان مذکر و مؤنثی بودند که در کلاس فلسفه دانشگاه فلوریدا حضور داشتند و در گروه‌های ۸ تا ۲۵ نفره قرار گرفتند. تصاویر استفاده شده در این پایگاه تصاویر:

- شامل منتخبی از طیف وسیعی از تصاویر در سه بعد مدل احساسی می‌باشند

- همه تصاویر رنگی هستند.

- تصاویر انتخاب شده به راحتی قابل فهم و قابل ارتباط برقرار کردن سریع می‌باشند.

تلاش زیادی شد که تصاویر همه ابعاد فضای احساسی ۳ بعدی را در برگیرد به همین خاطر از تصاویر ناخوشایند مثل آلودگی یا کودکان گرسنه در این پایگاه داده استفاده شد.

در تهیه این پایگاه داده برای ارزیابی احساسات از خودارزیاب آدمکی (SAM) [۶۲] استفاده شده است.

شرکت‌کنندگان با استفاده از SAM به تصاویر در سه بعد arousal، valence، dominance امتیاز دادند و

در نهایت هر تصویر با اساس میانگین مقادیر در هر یک از سه بعد در پایگاه داده قرار داده شد. [۶۳]

سیستم بین‌المللی صوت‌های دیجیتالی احساسی^۱ IADS این مجموعه نیز همانند مجموعه IAPS توسط مرکز احساس و توجه NIMH، CSEA دانشگاه فلوریدا برای تهیه ابزار استاندارد، در دسترس محققان قرار گرفته است.

شرکت‌کننده‌ها، دانشجویان مذکر و مؤنث کلاس فلسفه دانشگاه فلوریدا بودند که گروه‌های ۶ تا ۲۵ نفره را تشکیل دادند. سعی بر آن شده بود که در هر گروه هر دو جنسیت حضور داشته باشند. حداقل ۱۰۰ شرکت‌کننده به هر فایل صوتی امتیاز دادند که تقریباً نیمی از آن‌ها مؤنث بودند. صوت‌های انتخاب شده :

- شامل منتخبی از کل فضای ۳ بعدی احساسی می‌باشد.
 - صوت‌ها به راحتی قابل فهم هستند که امکان برقراری ارتباط سریع را فراهم می‌آورند.
- در تهیه این پایگاه داده نیز از SAM برای ارزیابی فایل‌های صوتی در سه بعد arousal, valence, dominance استفاده شد.

در نهایت هر صوت با اساس میانگین مقادیر در هر یک از سه بعد در پایگاه داده قرار داده شد. [۶۴]

¹International Affective Digitized Sounds.

eNTERFACE نام کارگاه‌های تابستانی هستند که توسط شبکه اروپایی برتری سیمیلار^۱ سازماندهی می‌شوند که نوع جدیدی از کارگاه‌های اروپایی به شمار می‌روند. این پایگاه داده توسط ساواران و همکارانش [۶۵] جمع‌آوری شده است که شامل دو مجموعه می‌باشد. اولین مجموعه شامل سیگنال‌های فیزیولوژی خارجی، الکتروانسفالوگرافی EEG، طیف‌سنجی مادون قرمز نزدیک تابعی^۲ (fNIRS) و فیلم‌های ویدئویی چهره ۵ نفر از شرکت‌کنندگان مذکر می‌باشد. مجموعه دوم فقط شامل fNIRS و فیلم‌های ویدئویی مربوط به چهره ۱۶ نفر از شرکت‌کنندگان مرد و زن می‌باشد. هر دو مجموعه با استفاده از پاسخ فوری به تصاویر احساسی از پایگاه داده IADS ضبط شده‌اند.

۴ - ۱ - ۴ MAHNOB-HCI

این پایگاه داده علاوه بر تشخیص احساس، برای کاربرد برچسب‌گذاری محتوا نیز استفاده می‌شود. در این پایگاه داده به هر یک از ۳۰ شرکت‌کننده بخش‌هایی از فیلم و تصاویر نشان داده می‌شود، درحالی‌که افراد توسط ۶ دوربین، یک میکروفون، یک ردیاب نگاه چشم و هم‌چنین سنسورهای اندازه‌گیری کننده EEG (۳۲ کانال)، ECG، دامنه تنفس و دمای پوست نظارت و کنترل می‌شوند. این پایگاه داده شامل دو قسمت می‌باشد. در قسمت اول با نمایش بخش‌هایی از فیلم از شرکت‌کنندگان خواسته شده است که حالت احساسی خود را بر مبنای معیارهای valence و arousal بیان کنند. در قسمت دوم تصاویر یا بخشی از فیلم به همراه یک برچسب در پایین صفحه نمایش پخش می‌شوند. در برخی موارد

^۱SIMILAR European Network of Excellence

^۲functional near infra-red spectroscopy

برچسب به درستی حالت احساسی فیلم را توصیف می‌کند و در برخی موارد دیگر برچسب با آیتم پخش شده مطابقت ندارد. بعد از هر آیتم از شرکت‌کننده خواسته می‌شود تا در صورت موافقت با برچسب بکار برده شده دکمه سبز روی صفحه نمایش و در غیر این صورت دکمه قرمز را فشار دهند. در طول انجام آزمایش داده‌های صوتی، تصویری، سیگنال‌های سیستم عصبی مرکزی و داده‌های فیزیولوژیکی همگام با سنسورها ضبط می‌شوند. پایگاه داده به صورت رایگان برای کاربردهای تحقیقاتی در دسترس می‌باشد [۶۶].

۴ - ۱ - ۵ DEAP^۱

DEAP یک پایگاه داده چندمدلی است که با هدف ایجاد یک سیستم تطابقی پیشنهادکننده ویدیو کلیپ مطابق با حالت و احساس شخص و همچنین تشخیص احساس تهیه شده است. در این سیستم پاسخ‌های جسمانی کاربر به احساسات تبدیل می‌شود. احساسات کاربر در زمان تماشای ویدیو کلیپ‌ها به سیستم پیشنهادکننده کمک می‌کند که در ابتدا سلیقه کاربر را تشخیص دهد، سپس یک ویدیو کلیپ متناسب با حالت کنونی احساس فرد را به او پیشنهاد کند.

در پایگاه داده‌های موجود تاکنون از ویدیو کلیپ استفاده نشده است و غالباً روی تصاویر یا موسیقی یا بخش‌هایی از فیلم انجام شده است. این پایگاه داده از فضای ۳ بعدی احساسی و خود ارزیاب SAM استفاده می‌کند که در آن valence از غیرفعال تا فعال، arousal از ناخوشایند تا خوشایند و dominance از احساس ضعیف تا احساس قدرتمند و تحت کنترل تغییر می‌کنند.

^۱Database for Emotion Analysis using Physiological Signals

در این پایگاه داده ویدیو کلیپ به عنوان محرک بصری برای استخراج کردن احساسات مختلف استفاده شد. یک مجموعه بزرگ و جامع از ویدیو کلیپ‌های مختلف به منظور برای هر ویدیو، یک دقیقه محرک‌تر و موثرتر انتخاب شد.

این پایگاه داده بیشترین تعداد شرکت‌کننده در میان همه پایگاه داده‌های موجود که از سیگنال‌های فیزیولوژیکی تشکیل شده‌اند، را دارد. در این آزمایش ۳۲ شرکت‌کننده سالم (۵۰ درصد مؤنث)، سن بین ۱۹ تا ۳۷ سال (میانگین ۲۶,۹ سال) حضور داشتند که EEG و سیگنال‌های فیزیولوژیکی دیگر همانند ECG, EMG, GSR و BVP آن‌ها در حین تماشای ۴۰ ویدیو کلیپ منتخب ضبط شد. شرکت‌کننده‌ها هر ویدیو را با استفاده از SAM که شامل معیارهای valence, arousal, dominance, like/dislike, familiarity می‌باشد ارزیابی کردند. برای ۲۲ شرکت‌کننده فیلم حالت چهره نیز ضبط شد. این پایگاه داده شامل همه داده‌های سیگنال‌های ضبط شده، فیلم‌های ضبط شده از چهره زیرمجموعه‌ای از شرکت‌کنندگان، امتیاز اختصاص داده شده شرکت‌کنندگان می‌باشد. لیستی از ۱۲۰ فیلم استفاده شده اولیه که در نهایت ۴۰ مورد از آن‌ها برای آزمایش استفاده شد و لینک‌های YouTube آن‌ها نیز در پایگاه داده می‌باشد. ۴۰ ویدیو کلیپ شامل مراحل زیر برای نمایش می‌باشد:

- یک نمایش ۲ ثانیه‌ای از شماره آزمایش برای اطلاع شرکت‌کنندگان از فرآیند
- ضبط ۵ ثانیه‌ای از حالت استراحت
- نمایش یک دقیقه‌ای از ویدیو کلیپ
- خودارزیابی SAM

در این پایگاه داده سیگنال EEG با نرخ نمونه‌برداری ۵۱۲ هرتز از ۳۲ الکتروود که طبق سیستم بین‌المللی ۱۰-۲۰ روی سر قرار گرفته‌اند، ضبط شده است. این مجموعه هم‌چنین شامل سیگنال‌های فیزیولوژیکی از

جمله: پاسخ گالوانیک پوستی، دامنه تنفسی، دمای پوست، ثبت ضربان قلب، اندازه‌گیری فشار خون، اندازه‌گیری فعالیت ماهیچه‌های مربوط به گردن و مربوط به لب‌خند و EOG می‌باشد [۶۶].

در این پایان‌نامه از پایگاه داده DEAP استفاده شده است. تنها از داده EEG ۳۲ کانالی بدون هیچ سیگنال فیزیولوژیکی خارجی بهره گرفتیم. ۳۲ کانال استفاده شده عبارتند از:

Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, O2

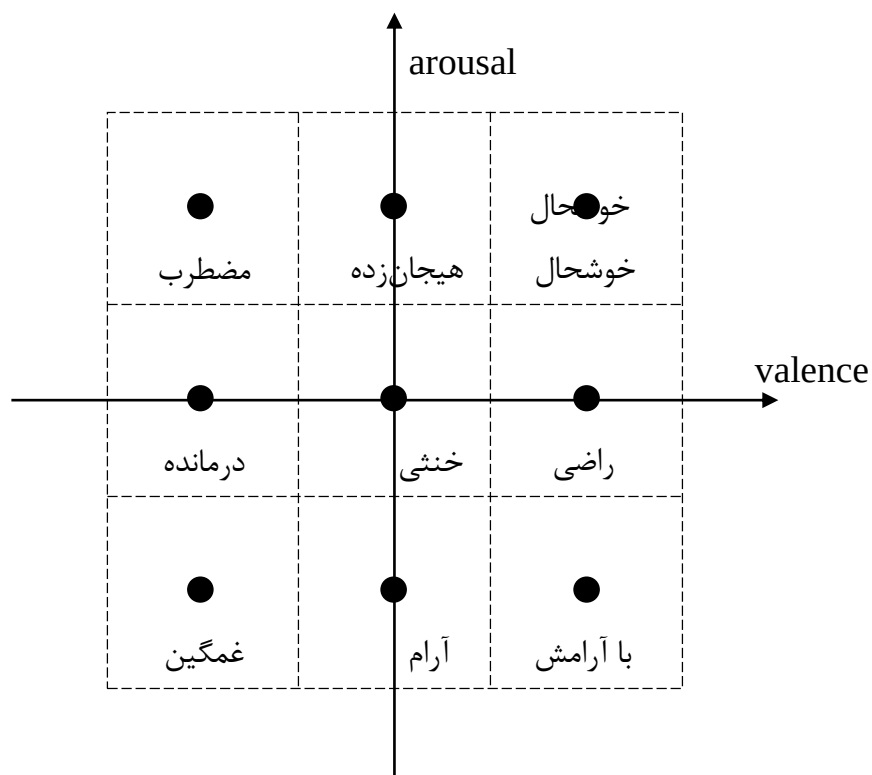
پیش‌پردازش‌هایی که در این پایگاه داده انجام شده است شامل موارد زیر می‌باشد:

- تغییر نرخ نمونه‌برداری سیگنال EEG ضبط شده به ۱۲۸ هرتز
- حذف مصنوعات و نویزهای مربوط به EOG به گونه‌ای که در [۶۶] ذکر شده است
- اعمال فیلتر میان‌گذر با فرکانس قطع ۴ تا ۴۵ هرتز
- حذف ۳ ثانیه ابتدایی سیگنال در هر کانال
- تقسیم‌بندی سیگنال‌های ۶۰ ثانیه‌ای متناظر با هر آزمایش

۴-۲ نحوه دسته‌بندی احساسات در روش پیشنهادی

همانطور که بیان شد در نهایت بعد از تماشای ویدیو کلیپ‌ها افراد یک ارزیابی از سطح arousal, valence, dominance و liking توسط خودارزیابی SAM انجام دادند (شکل ۴-۲). افراد اعداد ۱ تا ۹ را برای نشان دادن حالت‌های احساسی خود در هر معیار انتخاب کردند. در این تحقیق هر یک از مقیاس‌ها ۱-۹ valence و arousal به سه سطح تقسیم‌بندی شده است. سطح ۱-۳ valence به حالت احساسی "منفی"، ۴-۶ به "خنثی" و ۷-۹ به "مثبت" نگاشت شده است. مقیاس ۱-۳ arousal به حالت احساسی "غیرفعال"، ۴-۶

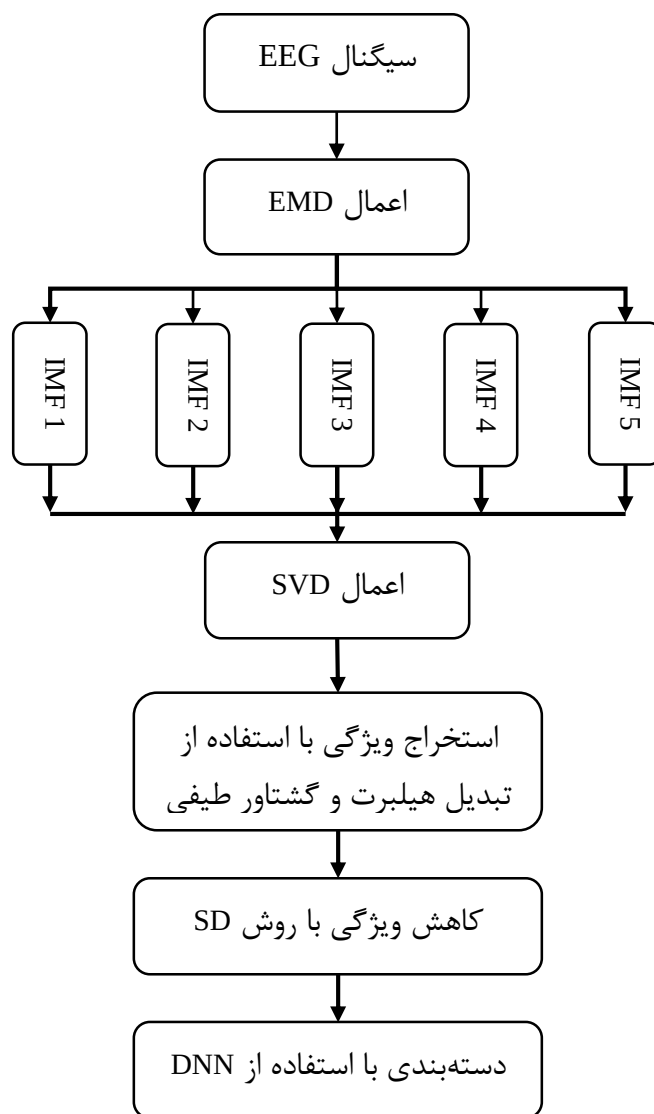
به "خنثی" و ۷-۹ به "فعال نگاشت شده است. طبق این تقسیم‌بندی سیستم ۹ حالت احساسی را طبقه-بندی می‌کند که شامل: خوشحال، راضی، با آرامش، هیجان‌زده، خنثی، آرام، ناراحت، مضطرب و بیچاره می‌باشد که در شکل ۴-۳ نشان داده شده است.



شکل ۴-۲. خود ارزیاب SAM برای حالات احساسی [۶۲]
شکل ۴-۳. دسته‌های حالات احساسی

۴ - ۱ روش پیشنهادی

در روش پیشنهادی ابتدا سیگنال EEG توسط تجزیه حالت تجربی (EMD) تجزیه می‌شود. EMD سیگنال را به مجموعه‌ای از توابع حالت ذاتی (IMF) تجزیه می‌کند. انتخاب بهترین و مؤثرترین اجزای IMF توسط الگوریتم تجزیه مقدار منفرد (SVD) انجام می‌شود به این منظور اجزای IMF متناظر با بیشترین مقادیر ویژه انتخاب شده، سپس بقیه اجزا حذف می‌شود. IMFها با استفاده از اجزا با بیشترین مقادیر ویژه بازسازی می‌شوند. بدست آوردن طیف IMF توسط تبدیل هیلبرت انجام می‌شود. در مرحله استخراج ویژگی از گشتاور طیفی که از طیف IMF استخراج می‌شود به عنوان ویژگی استفاده می‌شود. پس از آن تعداد ویژگی‌ها را با روش SD کاهش می‌دهیم. در نهایت از نوعی شبکه باور عمیق که واحد سازنده آن ماشین بولترمن محدود می‌باشد به منظور دسته‌بندی سیگنال به احساسات مختلف بر اساس مدل‌های دوبعدی در فضای arousal و valence استفاده شده است.



شکل ۴-۴. روش پیشنهادی

جزئیات روش پیشنهادی در ادامه فصل توضیح داده شده است.

۴ - ۲ استخراج ویژگی

در مقوله استخراج ویژگی از سیگنال EEG مشکلاتی وجود دارد: اول از همه اینکه سیگنال EEG غیرخطی (منظور از غیرخطی وجود ناخالصی‌های حاصل از نویز می‌باشد) و غیر ایستا می‌باشد. برای استفاد از روش‌های تحلیل سیگنال مثل تبدیل فوریه و موجک معمولاً سیگنال به صورت ایستا یا قطعه‌های ایستا در نظر گرفته می‌شود. ویژگی‌های استخراج شده با استفاده از روش‌های متداول خصوصیات غیرخطی یا غیر ایستایی سیگنال اولیه را حفظ نمی‌کنند. هم‌چنین برای پیدا کردن رخدادهای گذرا، روش‌های پردازش سیگنال متداول مثل آنالیز فوریه ارتباط بین رخداد یک اتفاق و زمان آن را فراهم نمی‌کنند.

در زمینه HMI و محاسبات احساسی ما هم‌چنان در آغاز راه فهم کامل ساختار سیگنال‌های EEG و ساختارهای عصبی مرتبط با احساسات هستیم. در مقوله تشخیص حالت‌های احساسی انسان‌ها با استفاده از سیگنال EEG مدرک معتبری که نشان دهد چه کانال‌هایی و چه محدوده فرکانسی برای مسئله دسته‌بندی الگو در یادگیری ماشین مناسب است، وجود ندارد. ما در این تحقیق سعی داریم تا مسائل مربوط به کانال و محدوده فرکانسی را تا حد ممکن برطرف کنیم. از این‌رو روشی بر مبنای EMD برای تجزیه سیگنال و استخراج ویژگی ارائه شده است.

در تحقیقات سال‌های اخیر نشان داده شده است روش آنالیز زمان-فرکانس بر اساس تبدیل هیلبرت-هانگ^۱ (HHT) و استفاده از فرکانس لحظه‌ای^۲ (IF) برای آنالیز سیگنال‌های غیر ایستا و غیر خطی مؤثر و مناسب می‌باشد [۶۷]. برای محاسبه فرکانس لحظه‌ای معنادار از یک سیگنال چندجزئی از طریق یک سیگنال

^۱Hilbert-Huang Transform

^۲Instantaneous Frequency

تحلیلی^۱ (AS) یا تبدیل هیلبرت، یک سیگنال چندجزیی بایستی به مجموعه‌ای از توابع تک جزیی تبدیل شود [۶۸]. چندجزیی به مواردی اشاره می‌کند که چندین اکسترمم بین دو نقطه متوالی گذرنده از صفر برای یک سیگنال نوسانی وجود دارد که نشان دهنده وجود چندین جزء فرکانسی در هر محدوده زمانی می‌باشد. سیگنال‌های EEG ذاتاً چندجزیی هستند. اگرچه عناصر این اجزای فرکانس‌هایی که با هم وجود دارند، در طول زمان تغییر می‌کنند که برای سیگنال‌های غیر ایستا معمولی و متداول می‌باشد. این روش با روش‌های بر مبنای فوریه که دامنه و فرکانس برای هارمونیک‌های هر محدوده زمانی ثابت می‌باشد، متفاوت است.

۴ - ۲ - ۱ تجزیه حالت تجربی^۲ (EMD)

در روش پیشنهادی EMD [۶۷] برای تجزیه سیگنال EEG به سیگنال‌های تک جزیی استفاده شده است. EMD یک روش مبتنی بر داده^۳ می‌باشد که قادر است خصوصیات غیر ایستایی و غیر خطی بودن سیگنال را حفظ کند. در مقایسه با دیگر الگوریتم‌های تحلیلی زمان-فرکانس، مزیت مهم الگوریتم EMD این است که الگوریتم تجزیه به یک پایه از پیش تعریف شده مثل موجک یا FFT نیاز ندارد. همانند روش موجک، EMD سری‌های زمانی اولیه را به حالت‌های نوسانی به نام توابع حالت ذاتی^۴ (IMF) تجزیه می‌کند ولی

^۱Analytical signal

^۲Empirical Mode Decomposition (EMD)

^۳Data-driven

^۴Intrinsic Mode Functions

برخلاف موجک، به مجموعه‌ای از موجک‌های از پیش تعیین شده محدود نیست. بعلاوه در موجک تعیین موجک مناسب برای یک آنالیز خاص یک چالش می‌باشد که EMD از این مشکل مستثنی می‌باشد. در این تحقیق استفاده از روش‌های HHT و EMD در آنالیز سیگنال‌های غیرخطی و غیر ایستا مثل سیگنال‌های EEG با موضوع تشخیص احساسات آزمایش و ارزیابی می‌شود. در ادامه الگوریتم EMD توضیح داده می‌شود.

۴-۲-۱- فرآیند تجزیه

EMD یک روش تطابقی تجزیه سیگنال می‌باشد که هر سیگنال پیچیده‌ای را به مجموعه‌ای از IMF ها تجزیه می‌کند. دو شرط اصلی که در حین فرآیند تجزیه باید به آن‌ها توجه شود عبارتند از [۶۹]:

۱. برای هر IMF، تعداد اکسترمم‌ها و تعداد نقاط عبور کننده از صفر باید حداکثر یک واحد اختلاف داشته باشند.
۲. مقدار میانگین منحنی تعریف شده توسط بیشینه‌ها و کمینه‌های محلی در هر نقطه باید صفر باشد، این مسئله نشان می‌دهد که بیشینه‌ها و کمینه‌ها به صورت متقارن در طول میانگین محلی (صفر) قرار گرفته‌اند.

الگوریتم EMD برای یک سری زمانی $x(t)$ در ادامه آورده شده است.

الگوریتم ۱. الگوریتم EMD

۱. پیدا کردن محل اکسترمم‌های $x(t)$

۲. درونیابی (با استفاده از درونیابی اسپلاین مکعبی) بین همه‌ی کمینه‌ها (هم‌چنین بیشینه‌ها) برای

بدست آوردن پوش پایین $e_{min}(t)$ و پوش بالا $e_{max}(t)$

$$۳. \quad m(k) = \frac{[e_{min}(t) + e_{max}(t)]}{2} \quad \text{محاسبه میانگین محلی}$$

۴. کم کردن میانگین محلی از سیگنال برای بدست آوردن "حالت نوسانی"، $s_1(t) = x(t) -$

$$m(t)$$

۵. در نظر گرفتن $s_1(t)$ به عنوان داده جدید و تکرار مراحل ۱ تا ۴، m بار تا زمانی که $s_{1.m}(t)$

به یک IMF تبدیل شود. در این مرحله وجود دو شرط ذکر شده برای IMF بررسی می‌شود.

$$\text{سپس داریم: } c_1(t) = s_{1.m}(t)$$

۶. بدست آوردن باقیمانده با کم کردن IMF بدست آمده از سیگنال اصلی $r_1(t) = x_1(t) -$

$$c_1(t)$$

۷. باقیمانده $r_1(t)$ به عنوان داده جدید در نظر گرفته می‌شود و الگوریتم از اول تکرار می‌شود تا

زمانی که باقیمانده $r_m(t)$ یک عدد ثابت، یک تابع تک جزئی یا یک تابع شامل تنها یک اکستریم

باشد به منظور هیچ IMF دیگری از آن استخراج نشود.

۸. در نهایت سیگنال اولیه می‌تواند به عنوان مجموعه‌ای از IMFها و باقیمانده به شکل زیر بیان

شود.

$$x(t) = \sum_{i=1}^m c_i(t) + r_m(t)$$

^۱Cubic spline

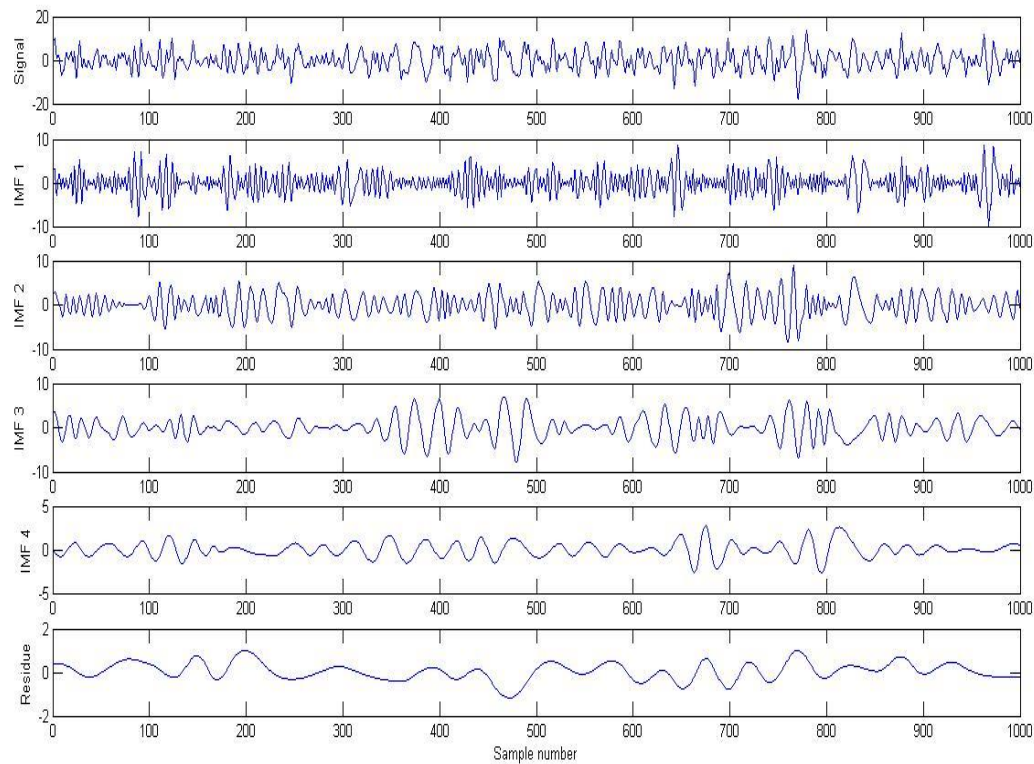
^۲Lower envelope

^۳Upper envelope

^۴Oscillatory mode

که در آن m تعداد IMF ها، $c_i(t)$ ، امین i IMF و $r_m(t)$ باقیمانده نهایی می باشد.

شکل ۴-۵ نتیجه اعمال EMD روی سیگنال EEG ثبت شده در یک کانال را نشان می دهد.



شکل ۴-۵. نمونه‌ای از نتیجه اعمال EMD روی سیگنال EEG ثبت شده در یک کانال

۴ - ۲ - ۱ - اسپکتروم هیلبرت-هانگ:

هانگ و همکاران [۶۸] بیان کرده‌اند که برای بدست آوردن یک فرکانس لحظه‌ای (IF) معنادار فیزیکی از طریق سیگنال تحلیلی، شرط مهم و ضروری داشتن یک سیگنال تک جزئی و با میانگین محلی صفر می باشد، هم‌چنین این جزء بایستی نسبت به میانگین صفر متقارن باشد. به عبارت دیگر ساختن یک سیگنال تحلیلی برای محاسبه IF ها تنها با سیگنال‌های تک جزئی کار می کند نه با موج‌های پیچیده. با اعمال EMD قبل از

ساختن سیگنال تحلیلی و تبدیل هیلبرت، قادر به استخراج فرکانس‌های لحظه‌ای معنادار از سیگنال‌های چندجزئی هستیم. بعد از تجزیه سیگنال به IMFها توسط EMD، هر نوع اطلاعات و خصوصیات مربوط به حوزه زمان حفظ می‌شود. اطلاعات محلی در حوزه فرکانس و شناسایی ساختارهای محلی پنهان در سیگنال اولیه توسط تبدیل هیلبرت قابل استخراج می‌باشد. اطلاعات محلی توسط اسپکتروم هیلبرت قابل توصیف می‌باشد. این اطلاعات شامل دامنه و فرکانس لحظه‌ای می‌باشد. [۷۰]

برای تعریف ریاضی، یک سیگنال تحلیلی را در نظر می‌گیریم که قسمت موهومی تبدیل هیلبرت، $\hat{x}$ ، از سیگنال $x(t)$ می‌باشد [۷۱].

پس داریم:

$$z(t) = x(t) + j\hat{x}(t) \quad (۱-۴)$$

$$\hat{x}(t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(t)}{s - t} dt \quad (۲-۴)$$

که در آن P مقدار Cauchy principal value می‌باشد.

با تبدیل به فرم توانی داریم:

$$z(t) = a(t)e^{i\varphi(t)} \quad (۳-۴)$$

که دامنه $a(t)$ و فاز $\varphi(t)$ به صورت زیر تعریف می‌شوند.

$$a(t) = \sqrt{x^2(t) + \hat{x}^2(t)} \quad (۴-۴)$$

$$\varphi(t) = \arctan\left(\frac{\hat{x}(t)}{x(t)}\right) \quad (۵-۴)$$

فرکانس لحظه‌ای سیگنال $x(t)$ به صورت زیر محاسبه می‌شود:

$$IF(t) = \frac{d\varphi(t)}{dt} = \frac{x(t)\dot{\hat{x}}(t) - \dot{x}(t)\hat{x}(t)}{x^2(t) - \hat{x}^2(t)} \quad (۶-۴)$$

۴ - ۲ - ۲ تجزیه مقدار منفرد (SVD):

SVD روشی برای تشخیص و مرتب کردن ابعاد داده‌ها به نحوی است که داده‌ها بیشترین تغییر را داشته باشند که در این صورت می‌توان بهترین تقریب از داده‌های اصلی با ابعاد کمتر را پیدا کرد. در نتیجه SVD می‌تواند به عنوان روشی برای کاهش بعد داده‌ها در مواردی چون بازسازی سیگنال یا فشرده‌سازی تصاویر استفاده شود. به این صورت که SVD عناصر متناظر با مقادیر کوچک را حذف کرده و تنها از عناصر متناظر با مقادیر منفرد بزرگ را برای بازسازی استفاده می‌کند.

در روش پیشنهادی به داده‌های با ابعاد بالا که همان IMF های بدست آمده از سیگنال EEG می‌باشند SVD اعمال می‌شود که آن‌ها را به یک فضای با ابعاد کمتر کاهش می‌دهد. در واقع یک ساختار از IMF ها با وضوح بیشتر ارائه می‌دهد که از بیشترین تأثیر تا کمترین تأثیر مرتب شده‌اند.

تجزیه مقدار منفرد بر پایه قوانین و خواص جبر خطی، یک ماتریس $m \times n$ مانند X را به ضرب ۳ ماتریس U_{mm} ، Σ_{mn} و V_{nn}^T شامل اجزای سازنده‌اش تجزیه می‌کند [۷۲].

$$X_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \quad (۷-۴)$$

^۱Singular value decomposition

که در آن $X_{m \times n}$ ماتریس اصلی، $U_{m \times m}$ ماتریسی شامل بردارهای ویژه چپ^۱ $\Sigma_{m \times n}$ ماتریسی شامل مقادیر منفرد^۲ و $V_{n \times n}$ ماتریسی شامل بردارهای ویژه راست^۳ می‌باشد. ماتریس‌های U و V متعامد و ماتریس Σ یک ماتریس قطری است. ستون‌های U و V بردارهای ویژه ماتریس‌های $X^T X$ و XX^T می‌باشند. عناصر قطری ماتریس Σ جذر مقادیر ویژه $X^T X$ و XX^T می‌باشند. X ماتریس ورودی و T علامت ترانپوز است. تجزیه مقدار منفرد را می‌توان بصورت زیر بیان نمود [۷۳].

$$X = \sum_{i=1}^r \sigma_i u_i v_i^T \quad (۸-۴)$$

که u_i ، i امین بردار ویژه XX^T ، v_i ، i امین بردار ویژه $X^T X$ ، σ_i ، i امین مقدار منفرد و r نیز بیانگر مرتبه ماتریس است. حاصلضرب $u_i v_i^T$ که ماتریسی هم بعد با ماتریس اصلی را تشکیل می‌دهد، i امین تصویر مشخصه می‌باشد. پس از تجزیه ماتریس داده به تصاویر مشخصه سازنده‌اش، با هر تعداد از این تصاویر مشخصه، داده قابل بازسازی است. هر چه از تصاویر بیشتری در بازسازی ماتریس بهره گرفته شود، حاصل به ماتریس اولیه نزدیک‌تر خواهد بود. در روش پیشنهادی پس از اعمال SVD ماتریسی هم بعد با ماتریس اولیه IMF‌ها تشکیل می‌شود که در آن از ۶ تصویر مشخصه برای بازسازی سیگنال استفاده شده است.

۴ - ۲ - ۳ گشتاور طیفی^۴

گشتاور یک معیار اندازه‌گیری کمی از شکل مجموعه‌ای از نقاط می‌باشد. گشتاور خصوصیت عددی از توزیع

^۱Left vectors of eigen value

^۲Singular Value

^۳Right vectors of eigen value

^۴Spectral moment

احتمال می‌باشد. در این تحقیق از گشتاور طیفی برای توصیف ویژگی‌های هر IMF استفاده شده است. این نوع از تحلیل با طیف سیگنال به عنوان توزیعی از انرژی رفتار می‌کند. گشتاور طیفی اطلاعات را مستقیماً از طیف بازیابی می‌کند. یکی از فواید گشتاور طیفی عدم حساسیت آن به تغییر فاز در سیگنال می‌باشد. فیلترها عموماً فاز و یا دامنه طیف را برهم می‌زنند. یکی دیگر از فواید آن، محاسبات مستقیم و قابل فهم می‌باشد. گشتاورها با مرتبه‌های مختلف تفسیر آماری مختلفی دارند که نشان دهنده توزیع داده‌ها با استفاده از میزان پراکندگی، تقارن و تمرکز حول میانگین می‌باشند که به عنوان میانگین، واریانس، چولگی^۱ و کشیدگی^۲ بیان شوند. به عبارت دیگر واریانس نشان دهنده پراکندگی داده‌ها، چولگی نشان دهنده میزان عدم تقارن توزیع داده‌ها و کشیدگی معیاری از تیزی منحنی در نقطه بیشینه توزیع داده‌ها می‌باشد [۷۴]. در روش پیشنهادی از گشتاور مرتبه دوم و چهارم به عنوان ویژگی استفاده شده است. گشتاور طیفی مرتبه i ام از طریق فرمول زیر قابل محاسبه می‌باشد [۷۵]:

$$M_i = \int_0^{\infty} f^i W_s(f) df \quad i = 0.1.2 \dots \quad (9-4)$$

که در آن f فرکانس و W_s ، $response PSD$ می‌باشد و i مرتبه گشتاور طیفی را نشان می‌دهد.

^۱skewness

^۲kurtosis

۴ - ۳ کاهش ویژگی

در فرآیند دسته‌بندی ویژگی‌هایی که اطلاعاتی در مورد کلاس‌ها داشته باشند می‌توانند به تفکیک کلاس‌ها کمک نمایند. در مقابل ویژگی‌هایی که اطلاعاتی در مورد کلاس ندارند، تنها سبب افزایش بعد ویژگی و کاهش کارایی دسته‌بندی می‌شود. ویژگی‌های بی‌ربط یا اضافه باعث پیچیدگی فرآیند طبقه‌بندی و ایجاد خطا در فرآیند پیش‌بینی می‌شود. از این‌رو روش‌های کاهش یا انتخاب ویژگی برای بدست آوردن ویژگی‌های مناسب برای دسته‌بندی ارائه شده‌اند.

بطور کلی دو روش برای کاهش ویژگی‌ها وجود دارد. روش اول انتخاب ویژگی^۱ می‌باشد که بین ویژگی‌های استخراج شده ویژگی‌های مناسب برای دسته‌بندی انتخاب می‌کنند. در روش دوم که تبدیل ویژگی^۲ نامیده می‌شود ویژگی‌های جدیدی با استفاده از ویژگی‌های اصلی ساخته می‌شود. روش‌های انتخاب ویژگی معمولاً بر روش‌های تبدیل ویژگی رجحت دارند زیرا ارائه و تعریف اولیه ویژگی‌های استخراج شده در آن تغییر داده نمی‌شود [۷۶].

روش‌های انتخاب ویژگی به دو دسته کلی تقسیم می‌شوند: فیلتر و پوشه‌ای^۳. در روش‌های بر مبنای فیلتر بدون اجرای دسته‌بند، ویژگی‌ها با استفاده از پارامترهای آماری محاسبه می‌شوند. روش‌های پوشه‌ای مرتبط-ترین ویژگی‌ها را با ارزیابی زیرمجموعه‌های ویژگی‌های مختلف الگوریتم دسته‌بندی انتخاب می‌کنند. روش-های پوشه‌ای کارایی بهتری دارند ولی هزینه محاسباتی بیشتری نسبت به روش‌های بر مبنای فیلتر دارند

^۱Feature selection

^۲Feature transformation

^۳wrapper

که به دلیل آموزش چندباره الگوریتم دسته‌بندی می‌باشد [۷۷]. در این تحقیق از روش وابستگی آماری^۱ (SD) که یک روش بر مبنای فیلتر می‌باشد جهت کاهش ویژگی استفاده شده است که در ادامه توضیح داده شده است.

۴ - ۳ - ۱ وابستگی آماری (SD) بین ویژگی‌ها و برچسب‌ها

روش SD تعیین می‌کند که آیا مقادیر یک ویژگی به برچسب‌های کلاس مربوطه وابسته هستند یا بصورت تصادفی همزمان اتفاق افتاده‌اند. مقدار هر ویژگی در ابتدا در یکی از سطح‌های Q پله‌بندی شده قرار می‌گیرد. مقیاس پله‌بندی کردن به گونه‌ای در نظر گرفته می‌شود که همه‌ی سطرها شامل تعداد مقادیر تقریباً برابر در کل مجموعه می‌شود. انتخاب سطرها به این طریق، به جای مقیاس‌های پله‌بندی کننده یکسان متداول، برای اعتبار آماری بخشیدن به وقوع سطوح‌های مختلف پله‌بندی کننده می‌باشد. همبستگی آماری بین مقادیر ویژگی Y و برچسب کلاس Z طبق فرمول زیر محاسبه می‌شود [۷۸].

$$SD = \sum_{y \in Y} \sum_{z \in Z} p(y, z) \frac{p(y, z)}{p(y)p(z)} \quad (۱۰-۴)$$

هرچه مقدار SD بزرگتر باشد، وابستگی بین مقادیر و برچسب‌های کلاس بیشتر می‌شود. در مواردی که ویژگی مستقل از برچسب کلاس باشد، SD مقدار کمینه ۱ می‌گیرد. این روش اندازه‌گیری با روش اطلاعات متقابل (MI) شباهت‌هایی دارد.

$$MI = \sum_{y \in Y} \sum_{z \in Z} p(y, z) \log\left(\frac{p(y, z)}{p(y)p(z)}\right) \quad (۱۱-۴)$$

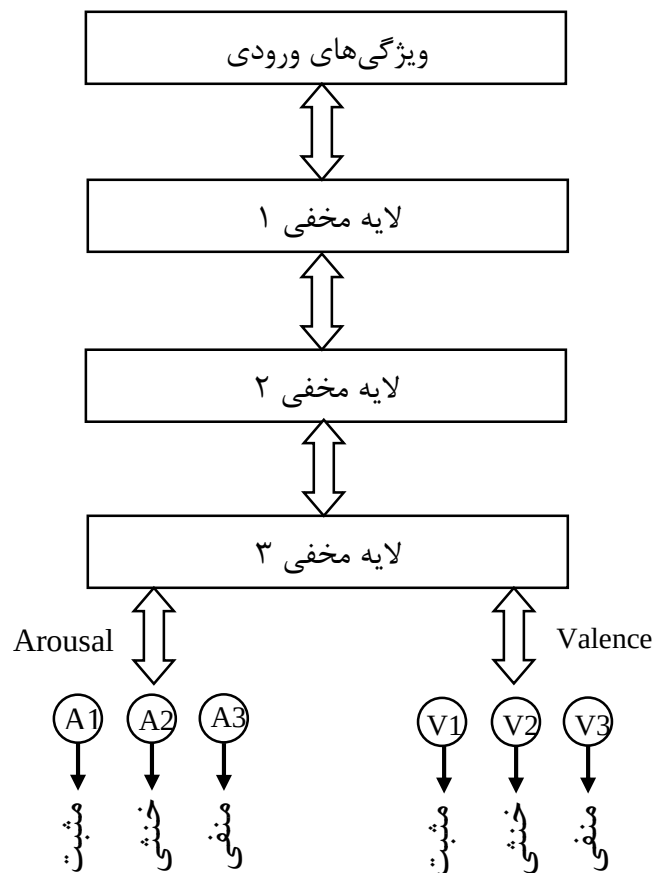
^۱Statistical Dependency

روش SD به علت نبود لگاریتم به سطح‌های پله‌بندی کننده حساس‌تر است. با این وجود معیار MI در آزمایش‌ها برای مقایسه استفاده شده است. روش‌های SD و MI به امتیازدهی و رتبه‌بندی ویژگی‌ها منتهی می‌شود که طبق آن تعدادی از ویژگی‌ها با بیشترین نرخ دسته‌بندی انتخاب می‌شوند.

اگر D_{FS} تعداد ویژگی‌های انتخاب شده باشد، تعداد سطح‌های پله‌بندی کننده برای این روش به صورت تجربی $L_S = \lfloor |D_{FS}|/10 \rfloor$ انتخاب شده است که به این معنا که هر ستون شامل ۱۰ نمونه داده می‌باشد.

۴-۴ دسته‌بندی با شبکه عصبی عمیق

روش پیشنهادی سیستم تشخیص احساس با ساختار عمیق از سه لایه RBM و دو دسته‌بند softmax ساخته شده است (شکل ۴-۶). سیستم دسته‌بندی احساس را با دسته‌بندی معیارهای valence و arousal به صورت جدا انجام می‌دهد. دو دسته‌بند softmax، یکی برای valence و دیگری برای arousal در نظر گرفته شده است.

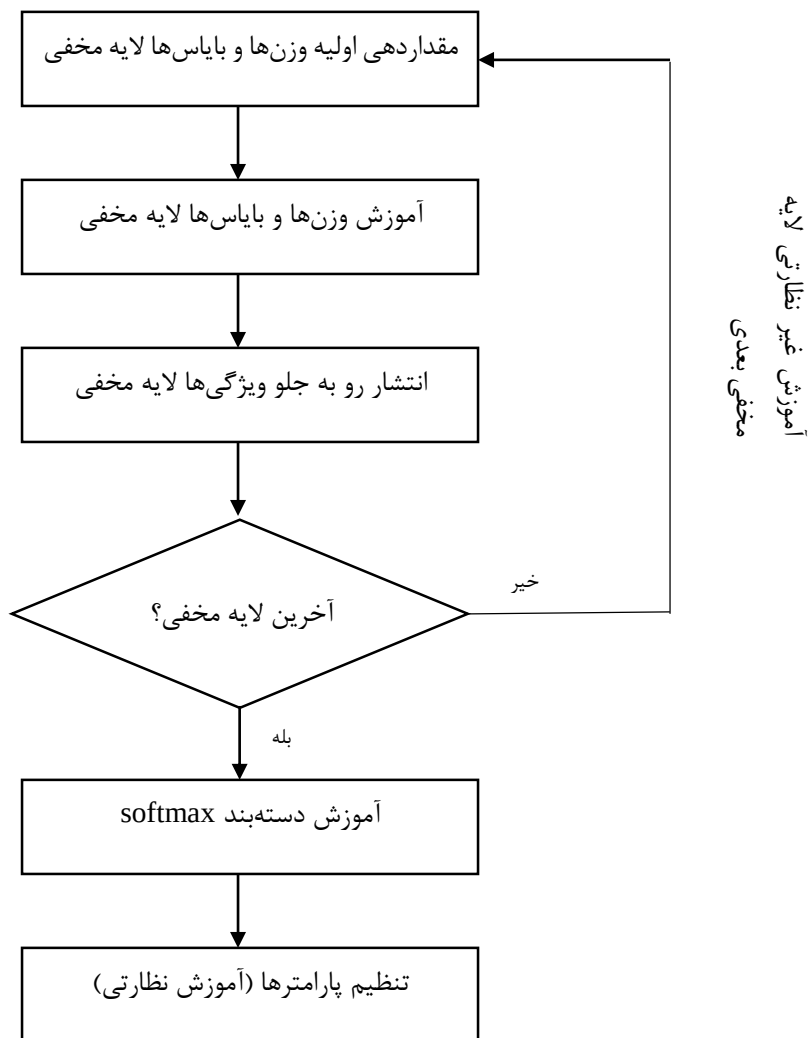


شکل ۴-۶. DNN با سه لایه مخفی و دو دسته‌بند softmax

DNN از تکنیک پیش‌آموزش لایه‌ای حریصانه استفاده می‌کند که از لایه ورودی آغاز و تا لایه softmax ادامه دارد. اولین RBM با استفاده از ویژگی‌های ورودی آموزش داده می‌شود. ویژگی‌های بدست آمده از لایه اول برای پیش‌آموزش لایه دوم انتشار می‌یابند. در این مرحله ویژگی‌ها از ویژگی‌های بدست آمده از لایه اول ساخته می‌شود. این‌رویه تا لایه آخر ادامه پیدا می‌کند تا همه پارامترها و وزن‌ها بدست آیند. برای بدست آوردن پارامترهای لایه softmax از یادگیری نظارتی استفاده شده است. ویژگی‌های خروجی از آخرین لایه مخفی به عنوان ویژگی‌های لایه softmax استفاده می‌شود. که در این تحقیق از خودارزیاب حالات احساسی افراد شرکت‌کننده به عنوان برچسب واقعی استفاده شده است.

بعد از اینکه شبکه پارامترهای بایاس و وزن‌ها را یاد می‌گیرد الگوریتم آموزش نظارتی نهایی انجام می‌شود. در واقع فرآیند آموزش با همه لایه‌ها و دسته‌بندهای softmax به عنوان یک مدل مجزا انجام می‌شود که وزن‌های همه لایه‌ها را با تکنیک انتشار رو به عقب بهبود می‌بخشد. تکنیک انتشار رو به عقب برای کاهش خطای دسته‌بندی استفاده شده است.

خلاصه‌ای از فرآیند آموزش DNN در شکل ۴-۷ نشان داده شده است.



شکل ۴-۷. خلاصه‌ای از فرآیند آموزش DNN

فصل ۵

آزمایش‌ها و نتایج

در این فصل عملکرد روش پیشنهادی با انجام آزمایش‌ها مختلف مورد بررسی قرار گرفت. نتایج بر روی دسته‌بندی دو معیار valence و arousal از پایگاه داده DEAP ارزیابی شده است. همانطور که در بخش ۴ - ۲ توضیح داده شد هر کدام از این معیارها بین ۱ تا ۹ مقداردهی شده‌اند که در سه بازه، سه تایی برای دسته‌بندی در نظر گرفته شده‌اند. در ابتدا روش اعتبار سنجی^۱ و دقت دسته‌بندی توضیح داده شده است. سپس چگونگی پیاده‌سازی روش پیشنهادی در پایگاه داده DEAP با جزئیات توضیح داده شده است. در ادامه نتایج حاصل از دسته‌بندی آورده شده است. آزمایش‌هایی که برای دستیابی به بهترین نتیجه انجام شد، شامل بدست آوردن مناسب‌ترین کانال‌ها از سیگنال EEG و انتخاب مجموعه‌ای مناسب از ویژگی‌ها می‌باشد. پس از آن روش پیشنهادی با معیارهای مختلفی شامل ارزیابی روش با استفاده ویژگی‌های مختلف، مقایسه با دیگر دسته‌بندها و همچنین مقایسه روش پیشنهادی با سایر کارهای انجام شده، مورد ارزیابی قرار گرفت.

۵ - ۱ اعتبار سنجی متقابل

برای ارزیابی روش از اعتبارسنجی متقابل استفاده شد. اعتبارسنجی متقابل روشی است با قابلیت تشخیص اینکه نتایج تا چه اندازه قابل تعمیم و مستقل از داده‌های آموزشی می‌باشد. این تکنیک بطور ویژه در کاربردهای پیش‌بینی با هدف تعیین میزان مفید بودن مدل پیشنهادی مورد استفاده قرار می‌گیرد. یکی از روش‌های اعتبارسنجی متقابل تکنیک K-زیر نمونه^۲ می‌باشد. در این نوع اعتبارسنجی داده‌ها به K

^۱Cross validation

^۲K-fold

زیرمجموعه افراز می‌شوند. از این K زیرمجموعه، هر بار یکی برای اعتبارسنجی و K-1 تای دیگر برای آموزش بکار می‌روند. این‌روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یکبار برای آموزش و یکبار برای اعتبارسنجی بکار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود. بطور معمول از ۱۰ زیر نمونه استفاده می‌شود. از مزیت‌های روش k زیر نمونه این است که همه داده‌ها هم به عنوان تست و هم آموزش مورد استفاده قرار می‌گیرند و هر داده تنها یک بار به عنوان تست استفاده می‌شود. در این تحقیق نیز k، ۱۰ در نظر گرفته شده است.

۵-۲ نرخ صحت

برای ارزیابی میزان کارایی دسته‌بند از نرخ تشخیص درست و یا مفهوم معکوس آن نرخ خطا استفاده شده است.

$$\text{نرخ صحت} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۱-۵)$$

که در آن:

TP: تشخیص‌های درست از میان مواردی که باید جزء یک کلاس تشخیص داده شوند

TN: تشخیص‌های درست از میان موارد غیر متعلق به یک کلاس

FP: تشخیص‌های نادرست از میان مواردی که باید جزء یک کلاس تشخیص داده شوند

FN: تشخیص‌های نادرست از میان موارد غیر متعلق به یک کلاس می‌باشند.

۵-۳ پیاده‌سازی روش پیشنهادی

روش پیشنهادی به صورت زیر پیاده‌سازی می‌شود:

سیگنال EEG استفاده شده از مشارکت ۳۲ فرد سالم (۵۰ درصد مؤنث) ثبت شده است. داده‌ها شامل ۳۲ کانال و ۸۰۶۴ نمونه می‌باشد. برای هر فرد ۴۰ آزمایش شامل تماشای ویدیو کلیپ‌های مختلف، برای ضبط سیگنال EEG انجام شده است. قبل از اعمال روش پیشنهادی به منظور افزایش کیفیت سیگنال پیش‌پردازش‌هایی به شرح زیر انجام شده است.

- تغییر نرخ نمونه‌برداری سیگنال EEG ضبط شده به ۱۲۸ هرتز
- حذف مصنوعات و نویزهای مربوط به EOG به گونه‌ای که در [۶۶] ذکر شده است
- اعمال فیلتر میان‌گذر با فرکانس قطع ۴ تا ۴۵ هرتز
- حذف ۳ ثانیه ابتدایی سیگنال در هر کانال
- تقسیم‌بندی سیگنال‌های ۶۰ ثانیه‌ای متناظر با هر آزمایش
- در مرحله اول پس از پیش‌پردازش دو کانال F3 و F4 برای مراحل بعدی پردازش انتخاب می‌شوند. طول داده مربوط به هر کلاس که در مراحل بعدی مورد پردازش قرار می‌گیرد برابر ۶۰ ثانیه یا ۷۶۸۰ نمونه می‌باشد.
- مرحله دوم شامل پنجره‌بندی کردن سیگنال می‌باشد. برای این کار تعداد پنجره‌ها ۵ در نظر گرفته شده است. بنابراین طول پنجره از تقسیم طول سیگنال بر ۵ بدست می‌آید. سیگنال به ۵ پنجره با طول ۱۵۳۶ نمونه تقسیم‌بندی می‌شود. بنابراین داده حاصل برای مرحله بعد که از مجموع دو کانال بدست می‌آید از ۴۰ آزمایش با اندازه 10×1536 تشکیل شده است.

- در سومین مرحله EMD اعمال می‌شود. ۱۰ سیگنال انتخاب شده با طول ۱۵۳۶ نمونه توسط EMD تجزیه می‌شود. برای هر سیگنال ۵، IMF محاسبه می‌شود و مجموعاً برای دو کانال ۱۰، IMF بدست می‌آید. در نهایت برای هر کدام از آزمایش‌ها ۵۰×۱۵۳۶ داده بدست می‌آید.
- در مرحله چهارم SVD روی ماتریس شامل IMF‌ها اعمال می‌شود. اجزای مؤثر هر کدام از IMF‌ها توسط SVD انتخاب می‌شود، به منظور اجزای IMF متناظر با بیشترین مقادیر ویژه انتخاب شده، سپس بقیه اجزا حذف می‌شود. IMF‌ها با استفاده از اجزا با بیشترین مقادیر ویژه بازسازی می‌شوند و ماتریسی با ابعاد اولیه ماتریس IMF‌ها ساخته می‌شود. در این تحقیق از اجزای متناظر با شش مقدار ویژه بزرگتر برای بازسازی سیگنال استفاده شده است.
- قدم پنجم شامل استخراج ویژگی گشتاور طیفی است. در این مرحله ابتدا طیف IMF‌ها توسط تبدیل هیلبرت بدست می‌آید. سپس گشتاور طیفی مرتبه ۲ و ۴ از طیف IMF‌ها محاسبه می‌شود. در واقع برای هر کدام از آزمایش‌ها با اندازه ۵۰×۱۵۳۶ دو ویژگی محاسبه می‌شود؛ در نتیجه برای هر آزمایش ۱۰۰ ویژگی استخراج می‌شود.
- در قدم ششم برای انتخاب ویژگی‌های مناسب با استفاده از روش SD، ۴۰ ویژگی از ۱۰۰ ویژگی برای هر آزمایش انتخاب می‌شود.
- قدم هفتم شامل جداسازی داده آموزش و تست می‌باشد. در این مرحله ۱۰٪ داده‌ها برای تست و ۹۰٪ داده‌ها برای آموزش در نظر گرفته می‌شود.
- در قدم هشتم دسته‌بندی با استفاده از DNN انجام می‌شود. برای انجام دسته‌بندی داده‌ها به DBN داده می‌شود. DBN شامل یک لایه ورودی و ۳ لایه مخفی و یک لایه خروجی شامل دو دسته‌بند softmax برای هر کدام از معیارهای valence و arousal می‌باشد. که نتایج حاصل از دسته‌بندی برای valence و arousal به سه دسته توسط نرخ صحت بدست می‌آید.

- در قدم نهم روش اعتبارسنجی متقابل ۱۰ زیر نمونه اعمال می‌شود. نتیجه نهایی از میانگین ۱۰ نرخ صحت بدست آمده، حاصل می‌شود.

۵ - ۳ - ۱ تنظیم پارامترهای اولیه شبکه DNN

در این قسمت پارامترهای ساختار عمیق مورد استفاده در آزمایشات شرح داده می‌شود. مقادیر پارامترها با توجه به مطالب گفته شده توسط هینتون [۷] تعیین شده است اما مقادیر برخی پارامترها به طور تجربی و شهودی انتخاب شده است. وزن‌ها و بایاس‌ها بگونه‌ای مقداردهی شدند که دارای میانگین صفر و انحراف معیار ۰/۰۱ باشند.

بر اساس روابط اصلاح وزن‌های RBM محاسبه گرادیان می‌تواند توسط یک نمونه آموزشی انجام شود. اما معمولاً بهتر است داده آموزشی به دسته‌های کوچکتر^۱ تقسیم شود و از این گروه‌های کوچکتر برای اصلاح وزن‌ها استفاده شود. هینتون پیشنهاد می‌کند که اندازه دسته‌های کوچک برابر با تعداد کلاس‌ها باشد، از این رو تعداد دسته‌های کوچک برابر ۳ در نظر گرفته شده است. تضعیف وزن^۲ با اضافه کردن یک عنصر اضافه به گرادیان معمولی صورت می‌گیرد. عنصر اضافه مشتق تابعی است که وزن‌های بزرگ را جریمه می‌کند. مقدار تضعیف وزن ۰/۰۰۰۲ تعیین شده است. ممنتوم^۳ یک روش ساده برای افزایش سرعت یادگیری و جلوگیری از گیر افتادن در کمینه محلی است.

^۱Minibatch

^۲Weight decay

^۳Momentum

مقدار ممنتم ۰/۵ برای ۱۰ تکرار اول و ۰/۹ برای ادامه کار در نظر گرفته شده است. نرخ یادگیری^۱ برابر ۰/۰۰۱، تعداد تکرار در مرحله پیش آموزش ۳۰۰، در مرحله آموزش نهایی ۴۰۰ می‌باشد. تعداد واحدهای لایه مخفی DNN نیز ۵۰ بوده است.

جدول ۵-۱. دسته‌بندی سه کلاسه arousal و valence

نرخ صحت دسته‌بندی		فرد	نرخ صحت دسته‌بندی		فرد
arousal	Valence		arousal	Valence	
۸۸.۸۹	۸۶.۶۷	فرد ۱۷	۷۵.۵۶	۵۰	فرد ۱
۸۱.۱۱	۸۲.۲۲	فرد ۱۸	۶۲.۲۲	۵۵.۵۶	فرد ۲
۶۲.۲۲	۵۵.۵۶	فرد ۱۹	۵۰	۶۸.۸۹	فرد ۳
۸۷.۷۸	۶۳.۳۳	فرد ۲۰	۵۱.۱۱	۷۱.۱۱	فرد ۴
۶۳.۳۳	۵۵.۵۶	فرد ۲۱	۶۶.۶۷	۶۵.۵۶	فرد ۵
۶۶.۶۷	۳۷.۷۸	فرد ۲۲	۷۶.۶۷	۵۴.۴۴	فرد ۶
۷۴.۴۴	۴۶.۶۷	فرد ۲۳	۴۴.۴۴	۷۲.۲۲	فرد ۷
۷۵.۵۶	۶۳.۳۳	فرد ۲۴	۶۸.۸۹	۶۰	فرد ۸
۵۷.۷۸	۳۴.۴۴	فرد ۲۵	۹۷.۷۸	۸۲.۲۲	فرد ۹
۴۵.۵۶	۵۷.۷۸	فرد ۲۶	۸۶.۶۷	۶۶.۶۷	فرد ۱۰
۵۱.۱۱	۵۷.۷۸	فرد ۲۷	۵۶.۶۷	۶۶.۶۷	فرد ۱۱
۵۶.۶۷	۵۷.۷۸	فرد ۲۸	۷۲.۲۲	۵۲.۲۲	فرد ۱۲
۵۰	۳۳.۳۳	فرد ۲۹	۷۴.۴۴	۵۳.۳۳	فرد ۱۳

^۱Learning rate

فرد ۱۴	۵۳.۳۳	۴۷.۷۸	فرد ۳۰	۸۱.۱۱	۹۱.۱۱
فرد ۱۵	۴۰	۷۵.۵۶	فرد ۳۱	۷۴.۴۴	۴۴.۴۴
فرد ۱۶	۷۳.۳۳	۷۸.۸۹	فرد ۳۲	۶۳.۳۳	۷۳.۳۳
			Valence		Arousal
میانگین			۶۰.۵۲		۶۷.۳۶

۵ - ۴ نتایج روش پیشنهادی در دسته‌بندی ۳ کلاسه

همانطور که قبلاً نیز اشاره شد برای تشخیص احساسات از دو معیار valence و arousal استفاده می‌شود. در این تحقیق هر کدام از این معیارها به ۳ دسته تقسیم می‌شوند که نتایج در جدول ۵-۱ آورده شده است.

۵ - ۵ انتخاب کانال‌ها

همانطور که قبلاً گفته شد در پایگاه داده DEAP سیگنال EEG از ۳۲ کانال ضبط شده است. همه‌ی کانال‌ها شامل اطلاعات مورد نیاز برای تشخیص احساسات نمی‌باشد. از این‌رو تنها از تعدادی کانال‌ها برای مراحل اصلی پردازش دسته‌بندی مورد استفاده قرار می‌گیرند. نتایج دسته‌بندی معیار arousal با استفاده از ترکیب کانال‌های متفاوت در جدول ۵-۲ آورده شده است. این ترکیبات از برخی مقالاتی که از پایگاه داده DEAP استفاده کرده‌اند، استخراج شده است.

تحقیقات انجام شده نشان می‌دهد که لوب پیشانی مغز با احساسات در ارتباط می‌باشد [۷۹]. نتایج بدست آمده در [۵۴] نیز نشان می‌دهد که کانال‌های F3 و F4 در مقایسه با سایر کانال‌ها اطلاعات بیشتری در رابطه با احساسات فرد منعکس می‌کنند. با کمتر شدن تعداد کانال‌ها زمان اجرا و بار محاسباتی سیستم تشخیص بطور چشمگیری کاهش پیدا می‌کند. علاوه بر این استفاده از کانال‌های کمتر مزیت بزرگی در طراحی و پیاده‌سازی محصولات تجاری می‌باشد.

جدول ۵-۲. نتایج انتخاب کانال‌های متفاوت

میانگین نرخ دسته‌بندی arousal	میانگین نرخ دسته‌بندی valence	کانال‌ها	مقاله
۶۲,۳۴	۵۸,۴۳	FP1, AF3, F7, P7, P3, Pz, PO3, O1, CP2, C4, T8, FC6	[۴۶]
۶۶,۹۲	۵۹,۷۶	FP1, FP2, F3, F4	[۵۴]
۶۵,۷۲	۶۳,۷۲	۳۲ کانال	[۱۲], [۵۵]
۶۷,۳۶	۶۰,۵۲	F3, F4	کانال‌های پیشنهادی

۵ - ۶ ارزیابی روش با اعمال روش‌های کاهش ویژگی مختلف

مسئله دیگری که در این تحقیق مورد بررسی قرار گرفت مقوله کاهش ویژگی‌ها می‌باشد. در ۴ - ۳ انواع روش‌های کاهش ویژگی توضیح داده شد و دلیل انتخاب روش SD که یک روش بر مبنای فیلتر می‌باشد شرح داده شده است. مقایسه‌ای بین روش‌های مختلف به منظور انتخاب بهترین و مؤثرترین الگوریتم انجام

شده است. روش‌هایی که برای مقایسه انتخاب شده‌اند عبارتند از: تحلیل مؤلفه‌ای مستقل^۱ (PCA) و تحلیل تفکیک‌کننده خطی^۲ (LDA) از دسته روش‌های تبدیل ویژگی و انتخاب ویژگی ترتیبی رو به جلو^۳ (SFS) از دسته روش‌های پوشه‌ای. هدف ما از مقایسه چهار روش فوق، تعیین میزان اطلاعات مفیدی است که بعد از کاهش ویژگی برای دسته‌بند وجود دارد و این مقدار بر اساس نرخ صحت اندازه‌گیری می‌شود. به منظور مقایسه دقیق‌تر، کاهش ویژگی بصورت یکسان برای هر چهار روش ذکر شده صورت گرفت. با استفاده از این تکنیک‌ها، تعداد ویژگی‌ها تا ۴۰ ویژگی کاهش داده و معیارهای ارزیابی عملکرد شامل مدت زمان اجرا و نرخ صحت اندازه‌گیری شدند. نتایج حاصل از دسته‌بندی معیار arousal در اعمال هر یک از روش‌های کاهش ویژگی برای یک شرکت‌کننده در جدول ۳-۵ آورده شده است.

جدول ۳-۵. ارزیابی روش با اعمال روش‌های کاهش ویژگی مختلف

روش کاهش ویژگی	میانگین نرخ دسته‌بندی arousal	زمان اجرا (ثانیه)
PCA	۶۲,۵۰	۳۱,۱
LDA	۶۰,۷۷	۲۹
SFS	۵۰,۴۲	۳۹,۵
SD	۶۷,۳۶	۲۵,۳
بدون کاهش ویژگی	۶۴,۱۳	۳۱,۷

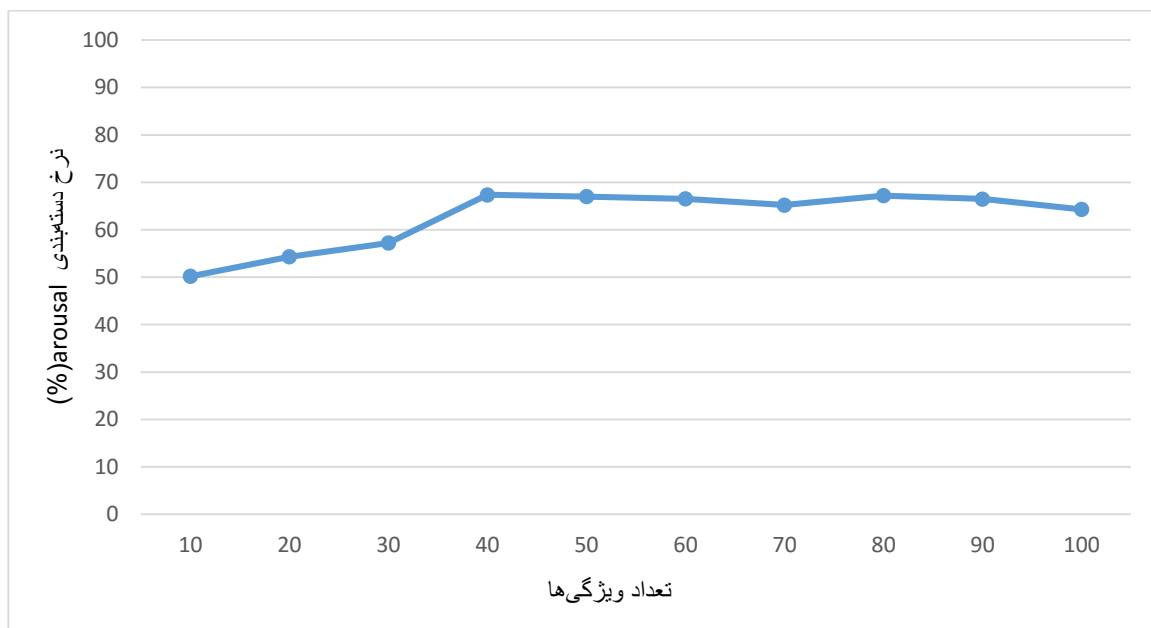
^۱Principle component analysis

^۲Linear discriminant analysis

^۳Sequential forward selection

مقایسه نتایج نشان داد که روش SD زمان اجرای کمتری نسبت به سایر روش‌ها دارد و نسبت به حالت ۱۰۰ ویژگی (بدون کاهش ویژگی) زمان اجرا با روش پیشنهادی ۲۰ درصد کمتر می‌باشد، همچنین نرخ صحت دسته‌بند نیز افزایش یافته است. بنابراین ۴۰ ویژگی بدست آمده از روش SD از روش‌های دیگر دارای اطلاعات مفید بیشتری می‌باشند و میزان اطلاعات مفیدی که در این روش کاهش ویژگی از دست می‌رود از سه روش دیگر کمتر است.

دلیل انتخاب ۴۰ ویژگی بعنوان تعداد بهینه ویژگی‌ها با ارزیابی روش با تعداد مختلف ویژگی‌های کاهش یافته نشان داده می‌شود. نتایج حاصل از تعداد ویژگی‌های مختلف در شکل ۵-۱. تعیین تعداد بهینه ویژگی ارائه شده است.



شکل ۵-۱. تعیین تعداد بهینه ویژگی

همانگونه که در شکل ۵-۱. تعیین تعداد بهینه ویژگی فوق مشاهده می‌شود ۴۰ ویژگی بهترین نتیجه را در دسته‌بندی می‌دهد. از ویژگی ۴۰ تا ۱۰۰ تقریباً نرخ صحت دسته‌بندی ثابت می‌باشد از این‌رو تعداد ۴۰ انتخاب می‌شود زیرا زمان کمتر و پیچیدگی محاسباتی کمتری را نسبت به تعداد بالاتر ویژگی‌ها دارد.

۵ - ۷ ارزیابی روش با استفاده از ویژگی‌های مختلف

در این تحقیق کارایی روش پیشنهادی با استفاده از ویژگی‌های مختلف مورد آزمایش و ارزیابی قرار گرفت. ویژگی‌های مورد استفاده عبارتند از: گشتاور طیفی، ضرایب رگرسیون خودکار، آنتروپی، ویژگی‌های آماری شامل میانگین، واریانس، و دو ویژگی ضریب تغییر^۲ و شاخص نوسان^۳. در ادامه توضیح مختصری از ویژگی‌های فوق ارائه می‌شود. نتایج حاصل از دسته‌بندی با استفاده از هر یک از ویژگی‌ها در جدول ۴-۵ آورده شده است.

- گشتاور طیفی در فصل چهار توضیح داده شده است.

- ضرایب رگرسیون خودکار:

مدل و ضرایب رگرسیون خودکار (AR) در نمایش و مدل کردن خصوصیات و اطلاعات موجود در سیگنال قابلیت بالایی دارند. مدل AR کاربردهای فراوانی در پردازش سیگنال EEG دارد. در مدل AR هر نمونه از مجموع نمونه‌های وزندار قبلی طبق فرمول ۲-۵ بدست می‌آید. مرتبه مدل AR توسط تعداد وزن‌ها تعیین می‌شود که ضرایب AR نامیده می‌شوند.

$$x(t) = \sum_{i=1}^p a_i x(t-i) \quad i = 1, 2, \dots, p \quad (2-5)$$

که در آن P مرتبه مدل و ضرایب AR توسط a_i مشخص می‌شوند. ضرایب AR می‌توانند به عنوان ویژگی به کار برده شوند. برای بدست آوردن این ضرایب، عموماً داده EEG به بلاک‌هایی پنجره‌بندی می‌شود. سپس

^۱Autoregression coefficient

^۲Coefficient of variation

^۳Fluctuation index

هنگامی که مقدار t در سرتاسر پنجره داده جابه‌جا می‌شود، معادلات مدل بی‌شماری بدست می‌آید که به ما اجازه می‌دهد تا ضرایب بهینه AR را محاسبه کنیم.

• آنتروپی:

هر سیگنال شامل دو قسمت است، آنتروپی که قسمت قابل پیش‌بینی سیگنال تصادفی غیرمنظم و قسمت اصلی اطلاعات می‌باشد. افزونگی^۱ که قسمت قابل پیش‌بینی، منظم و قسمت زائد اطلاعات می‌باشد. در داده‌های متغیر مانند سیگنال که دارای تغییرات زیادی است، از این داده‌ها می‌توان اطلاعات مفید و ارزشمند استخراج کرد که به نام آنتروپی شناخته می‌شوند که بی‌نظم و کاملاً تصادفی هستند [۸۰]. در یک داده متغیر، فرض کنید که متغیر تصادفی X حالت‌های سیستم را بیان می‌کند. مقدار متغیرهای X به صورت زیر هستند.

$$xX = \{x_1, x_2, \dots, x_n\} \quad (n > 0) \quad (3-5)$$

و احتمال هر یک برابر است با:

$$P = \{p_1, p_2, \dots, p_n\} \quad 0 \leq p_i \leq 1, i = 1, 2, \dots, n \quad (4-5)$$

که باید داشته باشیم:

$$\sum_{i=1}^n p_i = 1 \quad (5-5)$$

بنابراین اطلاعات آنتروپی از این سیستم توسط فرمول زیر بدست می‌آید:

$$H = - \sum_{i=1}^n p_i \ln p_i \quad (6-5)$$

^۱Redundancy

یکی از این آنتروپی‌های مهم، آنتروپی طیف است. آنتروپی طیف، برآوردی از میزان پیچیدگی سری زمانی را ارائه می‌کند. تبدیلات مختلفی برای بدست آوردن طیف سیگنال وجود دارد. یکی از این روش‌ها تبدیل فوریه است که در آن چگالی طیف سیگنال بدست می‌آید. این تبدیل توزیع چگالی طیف را به صورت تابعی از فرکانس سیگنال بیان می‌کند.

• ویژگی‌های آماری:

ویژگی‌های آماری از متداول‌ترین روش‌های انتخاب ویژگی برای دسته‌بندی‌ها می‌باشند. از دو پارامتر آماری میانگین و انحراف معیار برای مقایسه انتخاب شده است. میانگین با استفاده از فرمول زیر محاسبه می‌شود.

$$\mu_X = \frac{1}{N} \sum_{n=1}^N X(n) \quad (7-5)$$

انحراف معیار نیز از طریق فرمول زیر بدست می‌آید.

$$\sigma_X = \sqrt{\frac{1}{N} \sum_{n=1}^N (X(n) - \mu(x))^2} \quad (8-5)$$

این دو ویژگی از هر IMF محاسبه می‌شوند.

• ضریب تغییر:

یکی دیگر از ویژگی‌های استفاده شده در مقالات ضریب تغییر می‌باشد. با در نظر گرفتن میانگین μ_X و

انحراف معیار σ_X ضریب تغییر به صورت زیر قابل محاسبه می‌باشد [۸۱].

$$V_c = \frac{\sigma_X^2}{\mu_X^2} \quad (9-5)$$

• شاخص نوسان

شاخص نوسان برای اندازه‌گیری شدت تغییرات سیگنال استفاده شده است. این شاخص با استفاده از فرمول ۵-۶ محاسبه می‌شود [۸۱].

$$Fi = \frac{1}{L} \sum_{j=1}^L |IMF(j+1) - IMF(j)| \quad (۵-۱۰)$$

در این فرمول L بیانگر طول هر IMF می‌باشد.

جدول ۵-۴. مقایسه ویژگی‌های مختلف برای دسته‌بندی

ویژگی	نرخ دسته‌بندی valence	نرخ دسته‌بندی arousal
گشتاور طیفی	۶۰٫۵۲	۶۷٫۳۶
رگرسیون خودکار	۵۹	۶۱٫۸۸
آنتروپی	۵۵٫۷۸	۶۶٫۲۸
میانگین	۴۲٫۵۴	۴۰٫۳۲
انحراف معیار	۵۰	۴۵٫۷۹
ضریب تغییر	۴۷٫۱۶	۳۸٫۰۳
شاخص نوسان	۵۷٫۴۴	۴۹٫۲۷

همانطور که نرخ دسته‌بندی حاصل از ویژگی‌ها نشان می‌دهد ویژگی گشتاور طیفی به خوبی نماینده و ویژگی داده‌ها می‌باشد از این‌رو برای طبقه‌بندی داده‌های این تحقیق مورد استفاده قرار می‌گیرد.

۵ - ۸ ارزیابی روش با استفاده از دیگر دسته‌بندها

الگوریتم‌های دسته‌بندی بسیاری برای دسته‌بندی وجود دارند که در این تحقیق پرکاربردترین آن‌ها در زمینه تشخیص احساس با شبکه باور عمیق پیشنهادی DNN مقایسه می‌شوند. دسته‌بندها شامل ماشین بردار پشتیبان^۱ (SVM)، K نزدیکترین همسایگی^۲ (KNN) و شبکه عصبی^۳ (NN) می‌باشند. نتایج حاصل از اعمال دسته‌بندها در دسته‌بندی valence به صورت جدا برای هر فرد در جدول ۵ - ۵ آورده شده است.

جدول ۵-۵. نتایج مقایسه با سایر دسته‌بندها در دسته‌بندی valence

نرخ صحت دسته‌بندی valence				فرد	نرخ صحت دسته‌بندی valence				فرد
NN	KNN	SVM	DNN		NN	KNN	SVM	DNN	
۷۵,۲۴	۷۸.۸۹	۳۱.۹۰	۸۶.۶۷	فرد ۱۷	۷۱,۱۹	۶۲.۲۲	۵۰.۷۱	۵۰	فرد ۱
۶۸,۱۰	۷۸.۸۹	۴۰.۲۳	۸۲.۲۲	فرد ۱۸	۴۰,۲۴	۴۴.۴۴	۳۹.۲۸	۵۵.۵۶	فرد ۲
۵۰,۲۴	۳۵.۵۶	۳۶.۶۶	۵۵.۵۶	فرد ۱۹	۶۶,۴۳	۵۶.۶۷	۴۶.۶۶	۶۸.۸۹	فرد ۳
۵۹,۵۲	۴۶.۶۷	۴۸.۸۰	۶۳.۳۳	فرد ۲۰	۳۴,۰۵	۳۸.۸۹	۴۱.۶۶	۷۱.۱۱	فرد ۴
۵۳,۸۱	۶۲.۲۲	۳۲.۱۴	۵۵.۵۶	فرد ۲۱	۵۴,۵۲	۶۴.۴۴	۲۹.۶۷	۶۵.۵۶	فرد ۵
۴۰,۹۵	۵۰	۵۰.۹۵	۳۷.۷۸	فرد ۲۲	۵۹,۵۲	۵۴.۴۴	۴۵.۴۷	۵۴.۴۴	فرد ۶
۵۳,۱۰	۵۲.۲۲	۲۶.۱۹	۴۶.۶۷	فرد ۲۳	۳۳,۳۳	۵۶.۶۷	۵۳.۵۷	۷۲.۲۲	فرد ۷
۶۹,۲۹	۳۵.۵۶	۳۳.۵۷	۶۳.۳۳	فرد ۲۴	۵۹,۵۲	۴۸.۸۹	۴۲.۶۱	۶۰	فرد ۸
۶۳,۸۱	۴۲.۲۲	۴۸.۸۰	۳۴.۴۴	فرد ۲۵	۸۹,۲۹	۸۳.۳۳	۳۹.۲۸	۸۲.۲۲	فرد ۹

^۱Support vector machine

^۲K-nearest neighbor

^۳Neural network

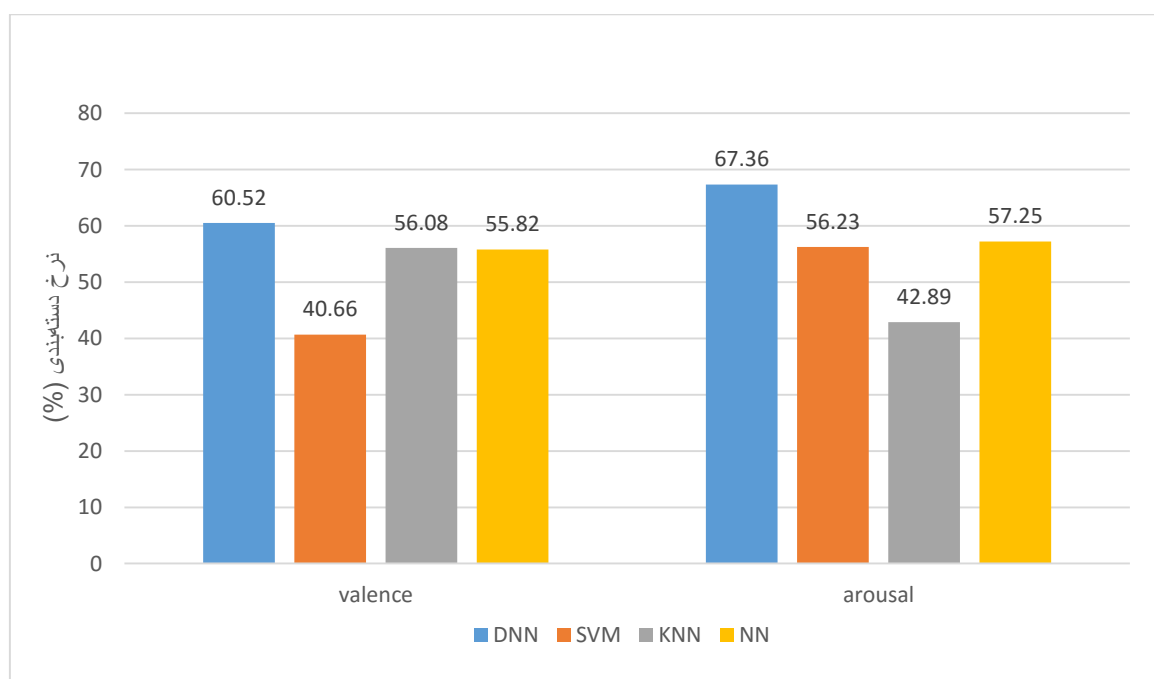
۵۰	۶۰	۳۹.۰۴	۵۷.۷۸	فرد ۲۶	۶۵.۲۴	۴۳.۳۳	۳۲.۶۱	۶۶.۶۷	فرد ۱۰
۴۵.۹۵	۴۰	۳۳.۰۹	۵۷.۷۸	فرد ۲۷	۳۱.۴۳	۴۱.۱۱	۴۲.۸۵	۶۶.۶۷	فرد ۱۱
۷۲.۸۶	۷۲.۲۲	۴۱.۶۶	۵۷.۷۸	فرد ۲۸	۵۳.۸۱	۳۶.۶۷	۴۹.۷۶	۵۲.۲۲	فرد ۱۲
۴۸.۸۱	۳۸.۸۹	۳۶.۹۰	۳۳.۳۳	فرد ۲۹	۴۴.۰۵	۶۴.۴۴	۵۲.۳۸	۵۳.۳۳	فرد ۱۳
۸۰.۲۴	۸۹.۴۴	۲۲.۶۱	۸۱.۱۱	فرد ۳۰	۴۰	۴۷.۷۸	۳۶.۴۲	۵۳.۳۳	فرد ۱۴
۴۰.۲۴	۷۱.۱۱	۳۳.۳۳	۷۴.۴۴	فرد ۳۱	۴۰.۲۴	۵۶.۶۷	۴۶.۴۲	۴۰	فرد ۱۵
۶۹.۵۲	۷۶.۱۱	۵۰.۹۵	۶۳.۳۳	فرد ۳۲	۶۱.۶۷	۶۴.۴۴	۴۴.۷۶	۷۳.۳۳	فرد ۱۶
					۵۵.۸۲	۵۶.۰۸	۴۰.۶۶	۶۰.۵۲	میانگین

جدول ۵-۶. نتایج مقایسه با سایر دسته‌بندها در دسته‌بندی arousal

نرخ صحت دسته‌بندی arousal				فرد	نرخ صحت دسته‌بندی arousal				فرد
NN	KNN	SVM	DNN		NN	KNN	SVM	DNN	
۸۱.۲۵	۸۳.۳۳	۵۲.۸۶	۸۸.۸۹	فرد ۱۷	۵۸.۷۵	۶۵.۵۶	۴۳.۵۷	۷۵.۵۶	فرد ۱
۶۶.۰۴	۶۰	۲۷.۳۸	۸۱.۱۱	فرد ۱۸	۴۹.۵۸	۳۱.۱۱	۵۲.۶۲	۶۲.۲۲	فرد ۲
۵۶.۰۴	۵۳.۳۳	۳۰	۶۲.۲۲	فرد ۱۹	۵۰	۵۶.۶۷	۴۱.۴۳	۵۰	فرد ۳
۷۶.۴۵	۸۰	۳۶.۱۹	۸۷.۷۸	فرد ۲۰	۵۰.۸۳	۳۱.۱۱	۶۲.۶۲	۵۱.۱۱	فرد ۴
۵۰.۸۳	۴۲.۲۲	۴۰.۲۵	۶۳.۳۳	فرد ۲۱	۵۵.۴۱	۴۵	۳۳.۸۱	۶۶.۶۷	فرد ۵
۵۰.۸۳	۶۱.۱۱	۳۱.۹۰	۶۶.۶۷	فرد ۲۲	۶۵.۲	۵۶.۶۷	۴۷.۳۸	۷۶.۶۷	فرد ۶
۵۹.۷۹	۶۳.۳۳	۵۱.۴۳	۷۴.۴۴	فرد ۲۳	۳۰.۸۳	۲۷.۷۸	۴۷.۸۴	۴۴.۴۴	فرد ۷
۶۱.۸۷	۶۳.۳۳	۴۶.۶۷	۷۵.۵۶	فرد ۲۴	۵۹.۵۸	۶۸.۸۹	۳۵.۷۱	۶۸.۸۹	فرد ۸
۳۶.۸۷	۴۵.۵۶	۳۰	۵۷.۷۸	فرد ۲۵	۸۹.۳۷	۹۳.۳۳	۵۰.۹۵	۹۷.۷۸	فرد ۹
۴۳.۳۳	۴۳.۳۳	۵۱.۱۹	۴۵.۵۶	فرد ۲۶	۷۴.۱۶	۷۸.۸۹	۴۹.۷۶	۸۶.۶۷	فرد ۱۰
۳۶.۴۵	۳۷.۷۸	۳۴.۷۶	۵۱.۱۱	فرد ۲۷	۵۰.۸۳	۵۲.۲۲	۵۰.۹۵	۵۶.۶۷	فرد ۱۱
۵۰.۸۳	۵۰	۳۷.۸۶	۵۶.۶۷	فرد ۲۸	۶۰.۸۳	۶۴.۴۴	۳۸.۱۰	۷۲.۲۲	فرد ۱۲
۴۳.۵۴	۳۸.۸۹	۴۵.۴۸	۵۰	فرد ۲۹	۶۳.۷۵	۶۰	۳۶.۱۹	۷۴.۴۴	فرد ۱۳

فرد ۱۴	۴۷.۷۸	۴۹.۵۲	۴۱.۱۱	۴۴.۵۸	فرد ۳۰	۹۱.۱۱	۴۶.۶۷	۷۴.۴۴	۷۹.۳۷
فرد ۱۵	۷۵.۵۶	۴۶.۶۷	۶۳.۳۳	۷۰.۴۱	فرد ۳۱	۴۴.۴۴	۳۵.۲۴	۵۲.۲۲	۳۷.۲۹
فرد ۱۶	۷۸.۸۹	۳۵	۶۵.۵۶	۶۴.۱۶	فرد ۳۲	۷۳.۳۳	۵۲.۳۸	۴۸.۸۹	۶۲.۹۱
میانگین	۶۷.۳۶	۴۲.۸۹	۵۶.۲۳	۵۷.۲۵					

با بررسی جدول ۵ - ۵ و جدول ۵ - ۶ مشاهده می‌شود که روش پیشنهادی نسبت به روش‌های دیگر بهتر عمل می‌کند و نرخ دسته‌بندی آن نسبت به سایر روش‌ها بیشتر است. این به علت ساختار عمیق DNN می‌باشد.



شکل ۵-۲. ارزیابی روش پیشنهادی با دسته‌بندی‌های مختلف

شکل ۵-۲ نشان دهنده میانگین نتایج ۳۲ شرکت کننده با استفاده از دسته‌بندی‌های مختلف می‌باشد. همانطور که مشاهده می‌شود در دسته‌بندی هر دو معیار valence و arousal شبکه باور عمیق از دسته‌بندی‌های دیگر بهتر عمل کرده است.

۵ - ۹ مقایسه روش با کارهای دیگر انجام شده در DEAP

در این بخش روش پیشنهادی با کارهای انجام شده بر روی پایگاه داده DEAP مقایسه می‌شود. برای مقایسه دو روش شبکه بیزین [۵۸] و DBN [۱۲] را انتخاب کردیم. در [۵۸] یک روش دسته‌بندی بر مبنای دسته‌بند بیزین ارائه شده است که در این روش به جای تابع پسین متداول در دسته‌بند بیزین، یک تابع وزن‌دار پسین^۱ استفاده شده است. در [۱۲] یک شبکه باور عمیق ارائه شده است که با استفاده از کدگذار خودکار پشته‌ای^۲ با روش یادگیری سلسله مراتبی پیاده‌سازی شده است. نتایج حاصل از مقایسه در جدول ۵-۷ آورده شده است.

جدول ۵-۷. نتایج مقایسه روش پیشنهادی با سایر کارها

روش‌ها	نرخ دسته‌بندی valence	نرخ دسته‌بندی arousal
شبکه بیزین [۵۸]	۵۳,۴	۵۳,۴۲
DBN [۱۲]	۵۱	۵۲,۰۳
روش پیشنهادی	۶۰,۵۲	۶۷,۳۶

^۱Weighted log posterior

^۲Stacked autoencoder

همانطور که مشاهده می‌شود روش پیشنهادی خیلی بهتر نسبت به روش‌های قبلی کار می‌کند و عمل
دسته‌بندی احساسات را انجام می‌دهد.

فصل ۶

نتیجه گیری

۶- ۱ کلیات پایان نامه

الگوریتم یادگیری عمیق یک شاخه نسبتاً جدید در یادگیری ماشین می باشد که در کاربردهای بسیاری نسبت به سایر الگوریتم های یادگیری نتایج بهتر و امیدوارکننده تری داشته است. همانطور که می دانیم حالت های احساسی در زمینه های مختلف زندگی روزمره انسان ها تأثیر می گذارند، بنابراین توانایی کشف و تشخیص حالات احساسی فرد در تعاملات انسان و ماشین ضروری می باشد. در این تحقیق از الگوریتم یادگیری عمیق برای دسته بندی احساسات با استفاده از سیگنال EEG استفاده شده است. هم چنین از خصوصیات و مزایای منحصربه فرد EMD و SVD برای تجزیه سیگنال بهره گرفتیم که روش پیشنهادی فرآیند دسته بندی احساسات را توانا تر ساخته است. فصل اول شامل مقدمه ای راجع به یادگیری عمیق و کاربردهای آن در دنیای واقعی می باشد. فصل دوم شامل مفاهیم اولیه یادگیری عمیق و احساسات می باشد. در این فصل پس از معرفی مفاهیم یادگیری عمیق، احساسات در تعاملات انسان و ماشین، مفهوم احساسات و روش های مدل کردن احساسات بیان شده است. در فصل سوم ابتدا توضیح مختصری در خصوص ساختار مغز انسان و EEG آمده است و با بررسی مطالعاتی که در گذشته با استفاده از سیگنال EEG در تشخیص احساسات انجام شده است، دنبال شده است. در فصل چهارم ابتدا پایگاه داده های پر کاربرد شامل سیگنال EEG در زمینه تشخیص احساسات را معرفی کردیم. پس از آن، ابزارهای مورد استفاده در روش پیشنهادی را بطور مفصل مورد بررسی قرار دادیم و در نهایت روش پیشنهادی بیان شده است. در فصل پنجم نتایج آزمایش ها و ارزیابی روش پیشنهادی در مقایسه با دیگر روش های دسته بندی آورده شده است. فصل ششم نیز شامل نتیجه گیری و پیشنهادهایی برای کارهای آتی در این زمینه می باشد.

۶-۲ نتیجه‌گیری و پیشنهادها

روش پیشنهادی با استفاده از EMD و SVD و الگوریتم یادگیری عمیق در زمینه تشخیص احساسات نو و بدیع می‌باشد. این روش از مزایای EMD برای تجزیه سیگنال به IMFها با در نظر گرفتن غیر ایستا بودن این نوع سیگنال‌ها بهره می‌برد، بعلاوه از SVD برای انتخاب مؤثرترین اجزای IMFها استفاده می‌کند. روش پیشنهادی از مزایای یادگیری عمیق که روش نو و کارآمدی برای دسته‌بندی می‌باشد بهره می‌برد.

مراحل انجام شده روش پیشنهادی به ترتیب در ادامه ذکر شده است. در ابتدا سیگنال EEG توسط EMD تجزیه می‌شود. EMD سیگنال را به مجموعه‌ای از توابع حالت ذاتی IMF تجزیه می‌کند. برای انتخاب مؤثرترین اجزای IMFها از الگوریتم SVD استفاده شده است. به کمک SVD می‌توان اجزای IMF متناظر با بیشترین مقادیر ویژه را انتخاب کرد. IMFها با استفاده از اجزا با بیشترین مقادیر ویژه بازسازی می‌شوند. برای بدست آوردن اسپکتروم IMFها تبدیل هیلبرت بکار گرفته شده است. پس از آن در مرحله استخراج ویژگی از گشتاور مومنتی که از IMFها استخراج می‌شود به عنوان ویژگی استفاده می‌کنیم. سپس ویژگی‌های بدست آمده با روش SD کاهش داده می‌شود. در نهایت شبکه با یادگیری عمیق برای دسته‌بندی سیگنال به احساسات مختلف بر اساس معیارهای valence و arousal استفاده شده است.

این روش در مقایسه با روش‌های دیگر با استفاده از پایگاه داده DEAP احساسات را دسته‌بندی کرده‌اند، بهبود قابل توجهی داشته است که نتایج حاصل از برای دسته‌بندی هر یک از معیارهای valence و arousal به سه دسته به ترتیب ۵۲، ۶۰ و ۳۶، ۶۷ می‌باشد.

با توجه به نتایج بدست آمده می‌توان استنباط کرد که الگوریتم یادگیری عمیق نسبت به روش‌های سطحی دیگر در بسیاری از کاربردها بهتر عمل می‌کند. در کاربرد تشخیص احساسات استاندارد برای دسته‌بندی

احساسات وجود دارد و روش‌های مختلفی برای دسته‌بندی احساسات پیشنهاد و ارائه شده است که روش پیشنهادی در این پایان‌نامه در مقایسه با مطالعات انجام شده در این زمینه بهبود داشته است. همانطور که گفته شده الگوریتم یادگیری عمیق روش جدیدی می‌باشد و هنوز پتانسیل زیادی برای بهبود و توسعه دارد. روشی که در این تحقیق استفاده شد روشی نو در کاربرد تشخیص احساسات می‌باشد و نیاز به افزایش درصد تشخیص در آن وجود دارد. یکی از مواردی که می‌تواند در کارهای آینده به بهبود عملکرد الگوریتم کمک کند توسعه روش‌هایی برای انتخاب کانال‌های مورد استفاده می‌باشد. علاوه بر این، چندین ساختار عمیق دیگر نیز از جمله کدگذار خودکار و شبکه عصبی کانولوشنال^۱ وجود دارند که می‌توان کارایی آن‌ها را در تشخیص احساس با استفاده از سیگنال EEG مورد ارزیابی و بررسی قرار داد. مورد بعدی که می‌توان به آن اشاره کرد تحقیق و مطالعه بیشتر معیارهای dominance و liking در ارزیابی و دسته‌بندی احساسات می‌باشد.

^۱Convolutional

مراجع:

- [1] L. Deng, (2012), “Three classes of deep learning architectures and their applications: a tutorial survey,” **APSIPA Trans. signal Inf. Process.**
- [2] Y. Lecun, (2013), “Marc’Aurelio Ranzato. Deep Learning Tutorial”, ICML, P 15, Atlanta, Georgia.
- [3] H. T. Ghashghaei, C. C. Hilgetag, and H. Barbas, (2007), “Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala” **Neuroimage**, **34**, **3**, pp 905–923.
- [4] Y. LeCun, S. Chopra, R. Hadsell, M. Ranzato, and F. Huang, (2006), “A tutorial on energy-based learning”, **Predict. Struct. data**, **1**, pp 1-15.
- [5] C. Poultney, S. Chopra, and Y. L. Cun, (2006), “**Efficient learning of sparse representations with an energy-based model**”, vol 1, MIT Press, USA, pp. 1137–1144.
- [6] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, (2003), “An introduction to MCMC for machine learning” **Mach. Learn.**, **50**, 1–2, pp. 5–43.

- [7] G. Hinton, (2010), “A practical guide to training restricted Boltzmann machines”, **Momentum**, 9, 1, pp 926.
- [8] M. A. Carreira-Perpinan and G. E. Hinton, (2005) “On contrastive divergence learning,” in *Proceedings of the tenth international workshop on artificial intelligence and statistics*, , pp. 33–40.
- [9] G. E. Hinton, (2002), “Training products of experts by minimizing contrastive divergence”, **Neural Comput.**, 14, 8, pp. 1771–1800.
- [10] G. E. Hinton, S. Osindero, and Y.-W. Teh, (2006), “A fast learning algorithm for deep belief nets”, **Neural Comput.**, 18, 7, pp. 1527–1554.
- [11] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, (2007), “Greedy layer-wise training of deep networks”, **Adv. Neural Inf. Process. Syst.**, 19, pp. 153.
- [12] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, (2014), “EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation”, **Sci. World J.**, 2014, pp. 1-9.
- [13] J. A. Russell, (2009), “Emotion, core affect, and psychological construction”, **Cogn. Emot.**, 23, 7, pp. 1259–1283.
- [14] J. A. Russell, (2003), “Core affect and the psychological construction of emotion”, **Psychol. Rev.**, 110, 1, pp. 145-172.
- [15] P. Salovey, (1997), "Emotional development and emotional intelligence: Educational implications", vol 1, Basic Books, NewYork, pp. 200.
- [16] J. Kim and E. André, (2008), “Emotion recognition based on physiological changes in music listening”, **Pattern Anal. Mach. Intell. IEEE Trans.**, 30, 12, pp. 2067–2083.

- [17] D. M. Badzinski, **(1991)**, “Children’s Cognitive Representations of Discourse Effects of Vocal Cues on Text Comprehension”, **Communic. Res.**, **18**, **6**, pp. **715–736**.
- [18] B. Kort, R. Reilly, and R. W. Picard, (2001), “An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion”, icalt, p. 43, Wisconsin, USA.
- [19] S. Dockray and A. Steptoe, **(2010)**, “Positive affect and psychobiological processes”, **Neurosci. Biobehav. Rev.**, **35**, **1**, pp. **69–75**.
- [20] Y. F. Birks and I. S. Watt, **(2007)**, “Emotional intelligence and patient-centred care,” **J. R. Soc. Med.**, **100**, **8**, pp. **368–374**.
- [21] J. Healey, R. Picard, and F. Dabek, (1998), “A new affect-perceiving interface and its application to personalized music selection”, Proc. Wkshp Percept. User Interfaces, ,pp. 4-6, California, USA.
- [22] B. Fehr and J. A. Russell, **(1984)**, “Concept of emotion viewed from a prototype perspective”, **J. Exp. Psychol. Gen.**, **113**, **3**, pp. **464**.
- [23] G. E. Schwartz, P. L. Fair, P. Salt, M. R. Mandel, and G. L. Klerman, **(1976)**, “Facial muscle patterning to affective imagery in depressed and nondepressed subjects”, **Scienc**, **192**, **4238**, pp. **489–491**.
- [24] P. R. Kleinginna Jr and A. M. Kleinginna, **(1981)**, “A categorized list of emotion definitions, with suggestions for a consensual definition”, **Motiv. Emot.**, **5**, **4**, pp. **345–379**.
- [25] C. Pert, **(1977)**, "Molecules of emotion: The science behind mind-body medicine", vol 1, Touchtone, New York, pp. 187.

- [26] P. Eckman, (1972), "Universal and cultural differences in facial expression of emotion," **Nebraska symposium on motivation**, 1972, 19, pp. 207–284.
- [27] R. Calvo and S. D'Mello, (2010), "Affect detection: An interdisciplinary review of models, methods, and their applications", **Affect. Comput. IEEE Trans.**, 1, 1, pp. 18–37.
- [28] A. Ortony and T. J. Turner, (1990), "What's basic about basic emotions?", **Psychol. Rev.**, 97, 3, p. 315.
- [29] Spencer, H. (1890). "**The Principles of Psychology**", Vol. 1. Appelon, New York, pp. 25.
- [30] Bradley, M. M. & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. **Journal of Behavior Therapy and Experimental Psychiatry**, 25, 1. pp 25-49.
- [31] J. A. Russell and A. Mehrabian, (1997), "Evidence for a three-factor theory of emotions," **J. Res. Pers.**, 11, 3, pp. 273–294.
- [32] J. A. Russell, (1980), "A circumplex model of affect.", **J. Pers. Soc. Psychol.**, 39, 6, pp. 1161.
- [33] O. D. Richard, E. H. Peter, and G. S. David, (2001), "**Pattern classification**", Vol 2, A Wiley-Interscience, USA, pp. 373–378.
- [34] I. C. Christie and B. H. Friedman, (2004), "Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach", **Int. J. Psychophysiol.**, 51, 2, pp. 143–153.
- [35] F. J. Gray, (2002), "**Anatomy for the Medical Clinician**," Vol 1, Victoria: Shannon Books Melbourne, pp. 30.

- [36] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, M. JO, and W. S. M. Neuroscience, (2004),“Sinauer associates”, *Inc.*, 3, 5, pp. 24.
- [37] T. F. Collura, **(1993)**, “History and evolution of electroencephalographic instruments and techniques.”, **J. Clin. Neurophysiol.**, **10**, **4**, pp. **476–504**.
- [38] Sinuly (2012) Phd. thesis, “ANALYSIS AND CLASSIFICATION OF EEG SIGNALS” philosophy depart. university of southern Qeenland, Australia
- [39] H. H. Jasper, **(1958)**, “The ten twenty electrode system of the international federation” , **Electroencephalogr. Clin. Neurophysiol.**, **10**, pp. **371–375**.
- [40] Christoph De Boeck, “electrode positions,” (2009). [Online]. Available: <http://www.staalhemel.com/2009/electrodepositions>. [Accessed: 27-Jan-2016].
- [41] E. Niedermeyer and F. H. L. da Silva, **(2005)**, "**Electroencephalography: basic principles, clinical applications, and related fields**", Vol 5, Lippincott Williams & Wilkins, USA, pp. 111.
- [42] B. J. Fisch and R. Spehlmann, **(1999)**, "**Fisch and Spehlmann’s EEG primer: basic principles of digital and analog EEG**". Elsevier Health Sciences , Louisiana, USA, pp. 35.
- [43] Y. Liu and O. Sourina, (2012), “EEG-based dominance level recognition for emotion-enabled interaction,” ICME , pp. 1039–1044, Victoria, Australia.
- [44] S. Koelstra, A. Yazdani, M. Soleymani, C. Mühl, J.-S. Lee, A. Nijholt, T. Pun, T. Ebrahimi, and I. Patras, (2010), Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos, , pp. 89–100, In:**Brain informatics**", Y. Yao, R. Son, T. Poggio, J. Liu, J. Huang, Springer Berlin Heidelberg, Germany.

- [45] S. A. Hosseini and M. B. Naghibi-Sistani, (2011), “Emotion recognition method using entropy analysis of EEG signals”, **Int. J. Image, Graph. Signal Process.**, 3, 5, pp. 30.
- [46] A. E. Vijayan, D. Sen, and A. P. Sudheer, (2015), “EEG-Based Emotion Recognition Using Statistical Measures and Auto-Regressive Modeling”, CICT, pp. 587–591, Ghaziabad, India.
- [47] V. Bajaj and R. B. Pachori, (2015), Detection of Human Emotions Using Features Based on the Multiwavelet Transform of EEG Signals, pp. 215–240, In: "**Brain-Computer Interfaces**", A. Hosseini and A. Azar, Springer.
- [48] R. Mehmood and H. Lee, (2015), “Towards emotion recognition of EEG brain signals using Hjorth parameters and SVM”, 91, pp. 24–27.
- [49] R. Ramirez and Z. Vamvakousis, (2012), Detecting emotion from EEG signals using the emotive epoc device, In: "**Brain Informatics**", F. Zannoto, S. Tsumoto, N. Taatgen, Y. Yao, Springer Berlin Heidelberg, Germany.
- [50] Khasnobish, A., Bhattacharyya, S., Singh, G., Jati, A., Konar, A., Tibarewala, D. N., & Janarthanan, R. (2013). “The Role of Empirical Mode Decomposition on Emotion Classification Using Stimulated EEG Signals“. In **Advances in Computing and Information Technology**, pp. 55-62, Springer Berlin Heidelberg.
- [51] N. Singhal and I. Engineering, (2013), “Comparison of Classifiers with Event Related Potentials as an Attribute in Emotion Quantification along Valence Axis Using EEG” , 7, 1, pp. 61–74.
- [52] T. S. Rached and A. Perkusich, (2013), “Emotion Recognition Based on Brain-Computer Interface Systems”, **Intech**.

- [53] H. Zhang, S. Warisawa, I. Yamada, (2015) ,“Emotional Valence Detection based on a Novel Wavelet Feature Extraction Strategy using EEG Signals”, HEALTHINF, pp. 52-59, Lisbon, Portugal.
- [54] T. F. Bastos-Filho, A. Ferreira, A. C. Atencio, S. Arjunan, and D. Kumar, (2012), “Evaluation of feature extraction techniques in emotional state recognition”, IHCI, pp. 1–6, Kharagpur, India.
- [55] S. Hatamikia, K. Maghooli, and A. M. Nasrabadi, **(2014)**,“The Emotion Recognition System Based on Autoregressive Model and Sequential Forward Feature Selection of Electroencephalogram Signals”, **J. Med. Signals Sens.**, **4**, **3**, pp. **194**.
- [56] M. Murugappan, R. Nagarajan, and S. Yaacob, **(2009)**, “Comparison of different wavelet features from EEG signals for classifying human emotions”, ISIEA, pp. 836–841, Kuala Lumpur, Malaysia.
- [57] M. Murugappan, N. Ramachandran, and Y. Sazali, **(2010)**, “Classification of human emotion from EEG using discrete wavelet transform”, **J. Biomed. Sci. Eng.**, **3**, **04**, pp. **390**.
- [58] S. Y. Chung and H. J. Yoon, (2012), “Affective classification using Bayesian classifier and supervised learning”, ICCASpp. 1768–1771, JeJu Island, South Korean.
- [59] K.-E. Ko, H.-C. Yang, and K.-B. Sim, **(2009)**,“Emotion recognition using EEG signals with relative power values and Bayesian network”, **Int. J. Control. Autom. Syst.**, **7**, **5**, pp. **865–870**.
- [60] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, (2006), Emotion assessment: Arousal evaluation using EEG’s and peripheral physiological signals, pp. 530–537, In: "**Multimedia Content Representation, Classification and Security**", B. Günsel, A. K. Jain, A. M. Tekalp and B. Sankur, Springer Berlin Heidelberg, Germany.

- [61] D. Wang and Y. Shang, (2013), "Modeling physiological data with deep belief networks," **Int. J. Inf. Educ. Technol.**, 3, 5, pp. 505.
- [62] J. D. Morris, (1995), "Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response", **J. Advert. Res.**, 35, 6, pp. 63–68.
- [63] Y. Liu and O. Sourina, (2013), "EEG databases for emotion recognition", Cyberworlds, pp. 302–309, Yokohama, Japan.
- [64] Bradley, M., & Lang, P. J. (1999). The International affective digitized sounds (IADS), stimuli, instruction manual and affective ratings. NIMH Center for the Study of Emotion and Attention.
- [65] O. Martin, I. Kotsia, B. Macq, and I. Pitas, (2006), "The eNTERFACE'05 audio-visual emotion database", Data Engineering Workshops, p. 8, Atlanta, GA, USA.
- [66] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, (2012), "A multimodal database for affect recognition and implicit tagging", **Affect. Comput. IEEE Trans.**, 3, 1, pp. 42–55.
- [67] N. E. Huang, (2007), An adaptive data analysis method for nonlinear and nonstationary time series: the empirical mode decomposition and Hilbert spectral analysis, pp. 363–376, In: "**Wavelet Analysis and Applications**", T. Qian, M. I. Vai and Y. Xu, Birkhäuser Basel, Switzerland.
- [68] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, (2009), "On instantaneous frequency", **Adv. Adapt. Data Anal.**, 1, 02, pp. 177–229.
- [69] P. Flandrin, G. Rilling, and P. Goncalves, (2004), "Empirical mode decomposition as a filter bank", **Signal Process. Lett. IEEE**, 11, 2, pp. 112–114.

- [70] D. Kim and H.-S. Oh, (2009), “EMD: A package for empirical mode decomposition and hilbert spectrum”, **R J.**, 1, 1, pp. 40–46.
- [71] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, (1998), “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”, **Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**, 454, 1971, pp. 903–995.
- [72] K. Baker, (2005), “Singular value decomposition tutorial”, **Ohio State Univ.**, 2005, pp. 1–24.
- [73] M. D. Sacchi, (2002), “Statistical and transform methods in geophysical signal processing,” *available ht tp://cm-gw. phys. ualberta. ca/~ sacchi/saig/book/(Last viewed April 23, 2012).*
- [74] S. M. Alam and M. I. H. Bhuiyan, (2013), “Detection of seizure and epilepsy using higher order statistics in the EMD domain”, **Biomed. Heal. Informatics, IEEE J.**, 17, 2, pp. 312–318.
- [75] J. Li and J. Chen, “Appendix F: Spectral Moments,” *Stoch. Dyn. Struct.*, pp. 383–384.
- [76] P. J. García-Laencina, G. Rodríguez-Bermudez, and J. Roca-Dorda, (2014), “Exploring dimensionality reduction of EEG features in motor imagery task classification”, **Expert Syst. Appl.**, 41, 11, pp. 5285–5295.
- [77] I. Guyon and A. Elisseeff, (2003), “An introduction to variable and feature selection”, **J. Mach. Learn. Res.**, 3, pp. 1157–1182.
- [78] J. Pohjalainen, O. Räsänen, and S. Kadioglu, (2015), “Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits”, **Comput. Speech Lang.**, 29, 1, pp. 145–171.

- [79] R. J. Davidson, D. C. Jackson, and N. H. Kalin, **(2000)**, “Emotion, plasticity, context, and regulation: perspectives from affective neuroscience.”, **Psychol. Bull.**, **126**, **6**, pp. **890**.

- [80] Kapur, J. N., & Kesavan, H. K. (1992). Entropy optimization principles and their applications, pp.. 3-20, In “**Entropy and Energy Dissipation in Water Resources**”. Singh, Vijay, **Fiorentino**, M. Norwell, MA (United States); Kluwer Academic Publishers, US.

- [81] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai, **(2013)**, “Feature extraction and recognition of ictal EEG using EMD and SVM”, **Comput. Biol. Med.**, **43**, **7**, pp. **807–816**.

Abstract

Deep learning is a new branch of machine learning, which is shown promising result compared to other learning algorithms. Deep learning algorithm are capable to represent and classify a set of data in hence of their hierarchical structure. Moreover, deep structures provides comprehensive presentation in comparison with shallow structures. Deep learning is being used in machine vision, voice recognition, natural language processing and signal processing. In this work, we used Deep learning to classify EEG signals representing different class of emotions. Emotional states influence human's daily life. Detection and recognition of emotional states is necessary in human-machine interaction in order to make these interactions more similar to human-human interaction.

In the proposed method, EEG signals are decomposed by empirical mode decomposition (EMD). EMD decomposes EEG signal into a set of intrinsic mode functions (IMFs). Then singular value decomposition (SVD) is applied to extracted IMFs. IMFs' components corresponded to higher eigen values are selected and the others are discarded. IMFs are reconstructed using components related to higher Eigen values. IMFs' spectrum is obtained using Hilbert transform. In feature extraction step, spectral moment of IMFs is calculated as features. Finally, a type of Deep belief network composed of Restricted Boltzmann Machine is being used to classify EEG signal to detect emotion in 2 dimensional model of valence and arousal.

In this research DEAP database is used to evaluate the proposed method. Based on the obtained results, it is seen that the proposed method performs better than other algorithms applied on DEAP database. The result of classification of 3 classes of valence and arousal are 60% and 68% respectively.

Keywords: Deep learning, EEG signal classification, emotion recognition, empirical mode decomposition.



Shahrood University of Technology
Faculty of Computer Engineering and information technology
Artificial intelligence group

A Deep learning method for classification of EEG data

Mina sadat hosseini

Supervisors:

Dr.pouyan

Dr.ferdosi

Associate supervisor:

Dr.mashayekhi

Februray 2016