تحلیل احساس در رسانه‌های اجتماعی فارسی با رویکرد شبکه عصبی پیچشی

مرتضی روحانیان، مصطفی صالحی، علی درزی و وحید رنجبر

چکیده: افزایش کاربری شهروندان از رسانه‌های اجتماعی (مانند توئیتر، فروشگاه‌های برخط و غیره) آنها را به منبعی عظیم برای تحلیل و درک پدیده‌های گوناگون تبدیل کرده است. هدف تحلیل احساس استفاده از داده‌های به دست آمده از این رسانه‌ها و کشف گرایش‌های پیدا و پنهان کاربران نسبت به موجودیت‌های خاص حاضر در متن است. در کار حاضر ما با استفاده از شبکه عصبی پیچشی که نوعی شبکه عصبی پیش‌خور است، به تحلیل گرایش نظرات در رسانه‌های اجتماعی در دو و پنج سطح و با در نظر گرفتن شدت آنها می‌پردازیم. در این شبکه عمل کانولوشن با استفاده از صافی‌هایی با اندازه‌های مختلف بر روی بردارهای جملات ورودی اعمال می‌شود و بردار ویژگی حاصل به عنوان ورودی لایه نرم بیشینه برای دسته‌بندی نهایی جملات به کار می‌رود. شبکه‌های عصبی پیچشی با پارامترهای مختلف با استفاده از معیار مساحت زیر منحنی و بر روی مجموعه داده جمع‌آوری شده از رسانه‌های اجتماعی فارسی ارزیابی شدند و نتایج به دست آمده نشان‌دهنده بهبود کارایی آنها در گستره رسانه‌های اجتماعی نسبت به روش‌های سنتی یادگیری ماشین به خصوص بر روی داده‌ها با طول کوتاه‌تر هستند.

کلیدواژه: **تحلیل احساس، رسانه‌های اجتماعی، شبکه عصبی پیچشی، شدت نظرات، متون کوتاه.**

۱- مقدمه

دانستن و تحلیل نظر دیگران همواره جزء اساسی از فرایند تصمیم‌گیری انسان‌ها در طول تاریخ بوده است. افراد در هر جامعه‌ای هنگام روبه‌روشدن با چالش‌های متفاوت از مشورت اعضای آن جامعه بهره گرفته‌اند. امروزه گسترش رسانه‌های اجتماعی به افراد جامعه با نگاه‌های مختلف فرصت داده تا در فضای عمومی نظرات خود را درباره پدیده‌های گوناگون با هم به اشتراک بگذارند. ۶۹٪ کاربران بالغ اینترنت از رسانه‌های اجتماعی به عنوان محلی برای بحث درباره موضوعات مختلف و اطلاع از نظرات دیگران استفاده می‌کنند [۱]. محتوای تولیدشده از فعالیت در این رسانه‌ها که ۲۸٪ از حضور برخط کاربران را شامل می‌شود [۲]، می‌تواند منبع عظیم داده برای تحلیل و درک رفتار افراد در مواجهه با پدیده‌های مختلف باشد.

تحلیل احساس، شاخه‌ای از پردازش زبان طبیعی است که به تحلیل گرایش‌های مردم نسبت به موجودیت‌های خاص و ویژگی‌های مرتبط به آنها به صورت خودکار می‌پردازد. این موجودیت‌ها می‌توانند محصولات، سرویس‌ها، مجموعه‌ها، افراد، اتفاقات یا موضوعات مختلف باشند. تحلیل احساس متمرکز بر نظراتی در زبان طبیعی است که به صورت مشخص یا ضمنی دارای جهت‌گیری منفی، خنثی یا مثبت هستند. جملاتی که دارای جهتگیری می‌باشند، جملات نسبی[1] نامیده می‌شوند که در مقابل جملات عینی[2] قرار دارند که واقعیات را بدون اعمال نظر گوینده بیان می‌کنند. در تحلیل احساس، ما به بررسی جملات نسبی و جملات عینی که بیانگر اتفاقات و واقعیت‌هایی با بار منفی یا مثبت هستند می‌پردازیم.

در [۳] نویسندگان این نظرات را در ۵ مرتبه از نظر شدت جهت‌گیری قرار دادند: مثبت احساسی، مثبت منطقی، خنثی، منفی منطقی و منفی احساسی.

پژوهش‌های زیادی روش‌های معمول در یادگیری ماشین را در امر تحلیل احساس مورد بررسی قرار داده‌اند [۴]. رویکردهای معمول در این پژوهش‌ها اغلب بر پایه الگوریتم‌های بانظارت[3] و ویژگی‌های استخراج‌شده به صورت غیر خودکار بوده است [۵] تا [۷]. این گزینش ویژگی معمولاً به صورت دستی انجام می‌شود و بسته به موضوع و نوع متن متفاوت است. به همین دلیل مدل‌ها وابسته به متن بوده و حالت کلی و عمومی[4] ندارند. روش‌های یادگیری عمیق در سال‌های اخیر به عنوان مجموعه روش‌هایی با تعمیم‌پذیری بالا مورد توجه پژوهشگران حوزه پردازش زبان طبیعی بوده‌اند و استفاده از آنها در تحلیل احساس به خصوص برای زبان انگلیسی رایج شده است. امروزه الگوریتم‌های سنتی یادگیری ماشینی به مرور جای خود را به روش‌های یادگیری عمیق در تحلیل احساس می‌دهند. دلیل آن این است که این روش‌ها امکان این را دارند که بدون دخالت انسانی، ویژگی‌های پیچیده فراوانی درباره داده استخراج کنند. لازمه استفاده از مدل‌های یادگیری عمیق، داشتن داده آموزش کافی، زمان و منابع رایانشی مناسب برای آموزش درست مدل شبکه عصبی است [۸] تا [۱۰].

ما در این مقاله برای تحلیل احساس متن فارسی، از شبکه‌های عصبی پیچشی (CNN)[5] که نوعی شبکه عصبی پیش‌خور[6] و چند لایه هستند، استفاده می‌کنیم. در این شبکه‌ها برای به دست آوردن خروجی به جای اتصال هر نورون لایه ورودی به لایه خروجی، بر روی داده ورودی

این مقاله در تاریخ ۲۱ مهر ماه ۱۳۹۷ دریافت و در تاریخ ۱۷ آذر ماه ۱۳۹۸ بازنگری شد.

مرتضی روحانیان، کارشناس ارشد، دانشکده علوم و فنون نوین، دانشگاه تهران، تهران، ایران، (email: rohanian@ut.ac.ir).

مصطفی صالحی (نویسنده مسئول)، دانشکده علوم و فنون نوین، دانشگاه تهران، تهران، ایران، (email: mostafa_salehi@ut.ac.ir).

علی درزی، دانشکده ادبیات و علوم انسانی، دانشگاه تهران، تهران، ایران، (email: alidarzi@ut.ac.ir).

وحید رنجبر، استادیار، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران، (email: vranjbar@yazd.ac.ir).

1. Subjective Opinions
2. Objective Opinions
3. Supervised Learning
4. Generic
5. Convolutional Neural Networks
6. Feedforward

(بردارهای کلمات حاصل از جاسازی کلمات[1]) با استفاده از صافی‌های مختلف عمل کانولوشن صورت می‌گیرد. از آنجا که شبکه‌های پیچشی توان استخراج ویژگی از واحدهای زبانی با طول‌های مختلف را دارند، استفاده از آنها نتایج بهتری نسبت به روش‌های سنتی یادگیری ماشین برای زبان‌ها با منابع فراوان به بار می‌آورد [۱۱]. در این پژوهش شبکه‌های عصبی پیچشی برای اولین بار برای تحلیل احساس زبان فارسی بر روی داده‌های جمع‌آوری شده از اخبار و توئیتر به کار رفته است. داده‌ها دارای دو نوع برچسب‌گذاری دو و پنج‌تایی هستند و دارای کاربری و طول متغیرند. نتایج به دست آمده در این مقاله نشان می‌دهد که این شبکه‌ها برای زبان‌ها با منابع محدود مثل فارسی، کارایی بهتری نسبت به روش‌های سنتی یادگیری ماشین دارند. در دسته‌بندی داده‌های متنی با طول کم این بهبود نتایج تا ۱۲٪ رسیده است. مهم‌ترین دستاوردهای ما در این مقاله به صورت زیر است:

- استفاده از شبکه‌های عصبی پیچشی برای دسته‌بندی جملات در متون فارسی که با توجه به اطلاعات ما قبلاً برای زبان فارسی انجام نشده است.
- تحلیل احساس در پنج سطح مختلف برای زبان فارسی و در نظر گرفتن شدت قطبیت.
- تحلیل احساس بر روی واحدهای زبانی با طول‌های متفاوت و بررسی روش پیشنهادی بر روی گستره‌ای از متون فارسی. در این راستا مجموعه داده تحلیل احساس با برچسب‌زنی داده جمع‌آوری شده از توئیتر (متون محاوره با طول‌های متفاوت) و سایت‌های خبری فارسی (متون رسمی) تهیه شده است.

در ادامه ما در بخش ۲ به پژوهش‌های مرتبط با کار تحلیل احساس و در بخش ۳ به معرفی روش پیشنهادی می‌پردازیم. در بخش ۴ نتایج حاصل از روش پیشنهادی را گزارش و درباره آنها بحث می‌کنیم و بخش ۵ به نتیجه‌گیری و کارهای آتی اختصاص دارد.

۲- کارهای مرتبط

بیشتر مطالعات مرتبط با تحلیل احساس در گذشته بر اساس الگوریتم‌های یادگیری بانظارت انجام گرفته است که نیاز به تهیه داده برچسب‌خورده دارند. مدل بیز ساده[2]، ساده‌ترین و پراستفاده‌ترین الگوریتم احتمالاتی برای دسته‌بندی است و بر مبنای قضیه بیز کار می‌کند. این مدل احتمالات پسین رویدادها را محاسبه کرده و برچسبی که بیشترین احتمال پسین را دارد به رویداد نسبت می‌دهد. دسته‌بندی‌کننده پرکاربرد دیگر آنتروپی بیشینه (MEC)[3] است. آنتروپی بیشینه مدل احتمالاتی است که کار دسته‌بندی را می‌توان با آن انجام داد. این روش بر پایه مدل نمایی[4] و اصل حداکثر آنتروپی[5] است [۱۲]. استفاده از این روش تجربه‌های موفقی در کار پردازش زبان طبیعی از جمله در تحلیل احساس به ارمغان آورده است [۱۳]. این روش در اکثر (و نه در همه) مواقع نسبت به مدل بیز ساده برتری دارد [۴]. ماشین بردار پشتیبان (SVM)[6] برای کار دسته‌بندی اسناد بر مبنای موضوعات مشابه بسیار مفید است [۱۴]. روش SVM یک مدل یادگیری بانظارت است که کار آن دسته‌بندی‌کردن اشیا در کلاس‌های مختلف با استفاده از ویژگی‌های استخراج‌شده است. این دسته‌بندی با ایجاد ابرصفحه‌ای میان نمونه‌های هر کلاس و حداکثرکردن فاصله نمونه‌ها از این صفحه صورت می‌گیرد [۱۵]. برتری این روش نسبت به دیگر روش‌های مطرح یادگیری ماشین آن است که در مورد داده‌های ورودی پیش‌فرضی ندارد و به جای تکیه بر ارزش‌های احتمالاتی، سعی دارد تا بهینه‌ترین دسته‌بندی را با داده‌های موجود انجام دهد و نتایج به دست آمده از آن در تحلیل احساس برتری محسوسی به دیگر روش‌های یادگیری ماشین در زبان انگلیسی دارد [۶].

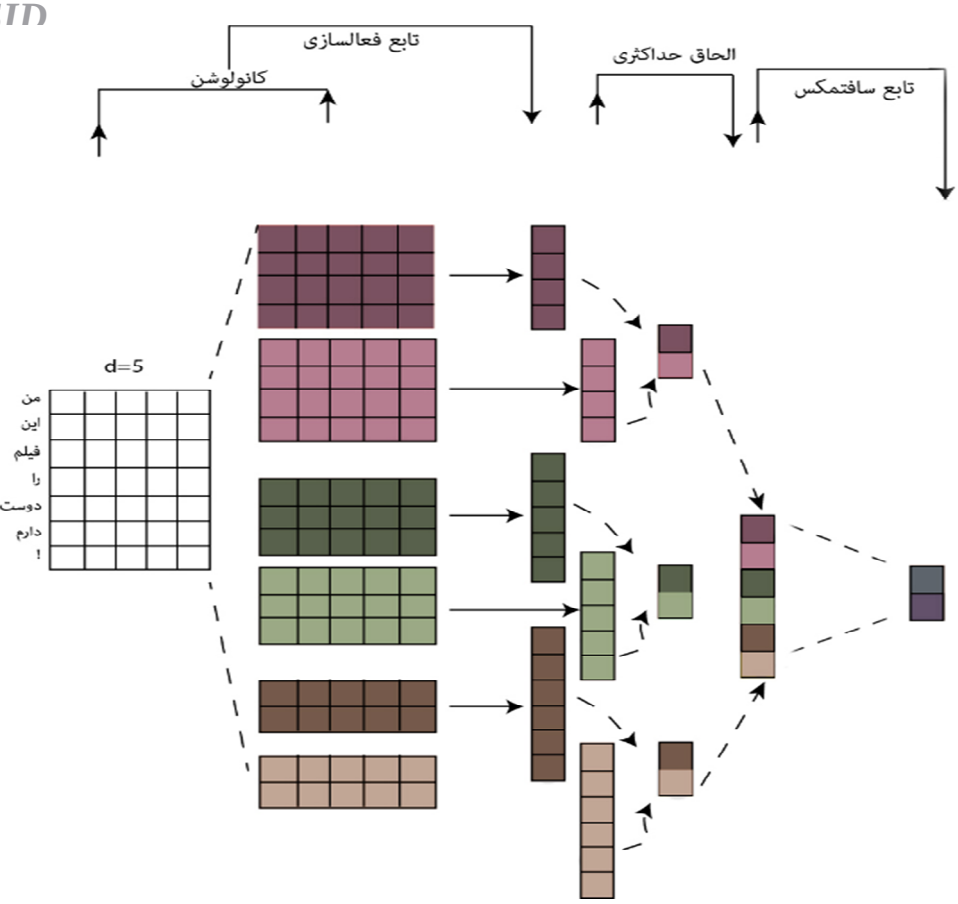
در سال‌های اخیر روش‌های یادگیری عمیق به خصوص شبکه‌های عصبی بازگشتی (RNN)[7] در تحلیل احساس برای زبان انگلیسی [۱۶]، چینی [۱۷] و آلمانی [۱۸] در میان زبان‌های مختلف، با استفاده از بردارهای مختلف نمایش کلمات کاربرد زیادی داشته است. آنها برای درک و کنترل ترکیب معنایی در کارهای پیچیده‌ای مانند تحلیل احساس مفید هستند. شبکه‌های RNN برای داده‌هایی با قابلیت تبدیل به مقادیر متوالی به کار می‌روند و با استفاده از ایده اشتراک‌گذاری پارامترها برای رسیدن به وزن‌های مطلوب، توانایی پردازش توالی‌هایی با طول‌های متفاوت را دارند [۱۹]. با وجود این که استفاده از آنها در تحلیل احساس برای زبان انگلیسی با نتایجی بهتر از روش‌های یادگیری بانظارت همراه بوده است [۹]، با رشد ساختار شبکه‌های RNN، ابعاد ماتریس‌ها در مرحله بازپخش به صورت توانی رشد می‌کنند و در عمل استفاده از آنها غیر ممکن می‌شود [۲۰].

شبکه‌های پیچشی که کولوبرت و دیگران [۲۱] در ابتدا برای کاربرد در بینایی رایانه‌ای ارائه کرده‌اند، اخیراً در بسیاری از کارهای پردازش زبان طبیعی مانند تجزیه نحوی، تجزیه سطحی، برچسب‌زنی نقش معنایی[8] و قطعه‌بندی[9] مورد استفاده قرار گرفته است. استفاده از شبکه‌های پیچشی در تحلیل احساس نیز برای زبان‌ها با منابع فراوان مورد استفاده قرار گرفته و باعث بهبود قابل توجه دقت و کاهش زمان مرحله آموزش نسبت به دیگر روش‌های یادگیری عمیق شده است [۱۰].

پژوهش‌های حوزه تحلیل احساس در زبان فارسی معمولاً یا با استفاده از روش‌های مبتنی بر قاعده هستند یا مبتنی بر پیکره [۲۲]. برای بهبود نتایج معمولاً از پیش‌پردازش نظرات و ویژگی‌های لغت‌نامه استفاده شده است [۲۳]. بصیری و همکاران [۲۴] یک چارچوب مبتنی بر لغت‌نامه ارائه کردند که به صورت بدون نظارت با استفاده از قواعد از پیش تعیین شده و لغت‌نامه تعریف‌شده جهتگیری متون محاوره را تشخیص می‌دهد. استفاده از SVM برای تحلیل احساس در زبان فارسی بر روی داده مربوط به نقد فیلم، منجر به نتایج بهتری نسبت به روش‌های دیگر یادگیری ماشین شده است [۲۵]. بازدهی این روش‌ها وابسته به کیفیت برچسب‌دهی در پیکره‌ها و شیوه گزینش ویژگی‌ها پیش از شروع کار دسته‌بندی است. روشنفکر و همکاران [۲۶] برای اولین بار از شبکه‌های عصبی LSTM برای تشخیص احساس متون فارسی استفاده کردند و توانستند نسبت به روش‌های یادگیری سنتی نتایج بهتری داشته باشند، اما این نوع شبکه‌ها برای آموزش نیاز به داده‌های خیلی زیادی دارند. همچنین آنها در کار خود فقط دو سطح از احساس را در نظر گرفتند و از جاسازی ساده کلمات استفاده کردند.

به طور کلی مزایای استفاده از یادگیری عمیق شامل موارد زیر

1. Word Embedding
2. Naïve Bayes Classifier
3. Maximum Entropy Classifier
4. Exponential Model
5. Principle of Maximum Entropy
6. Support Vector Machine
7. Recurrent Neural Networks
8. Semantic Role Labeling
9. Chunking



شکل ۱: نمایش یک شبکه عصبی پیچشی برای دسته‌بندی جملات.

است [۲۷]:

- احتیاجی نیست ویژگی‌ها به صورت دستی تهیه شوند. در یادگیری عمیق به جای استخراج دستی ویژگی‌ها معمولاً از جاسازی کلمات استفاده می‌شود که در آنها اطلاعات مربوط به بافت متنی وجود دارند.
- با استفاده از شبکه‌های عصبی یادگیری، انتخاب ویژگی‌ها[۱] و نمایش آنها می‌تواند هم با یادگیری بانظارت و هم بدون نظارت[۲] صورت گیرد.
- در تحلیل احساس با متن‌های گوناگونی از لحاظ سبک نوشتار و بافت معنایی روبه‌رو هستیم. انعطاف و تعمیم‌پذیری روش‌های یادگیری عمیق، اجازه می‌دهد تا با مشکل عدم تعمیم‌پذیری مدل کمتر روبه‌رو شویم.

۳- راهکار پیشنهادی

قبل از استفاده از داده‌ها در روش پیشنهادی، ابتدا توسط ابزارهای موجود، پیش‌پردازش‌هایی نظیر بهنجارسازی، توکن‌بندی و جداسازی بر روی داده‌های ورودی انجام می‌شود. همچنین برای تحلیل احساس زبان‌ها با منابع محدود مثل فارسی با استفاده از یادگیری عمیق نیاز به مجموعه‌ای از بردارها برای نمایش کلمات داریم که آنها را با استفاده از جاسازی کلمات روی مجموعه ویکی‌پدیای فارسی به دست می‌آوریم. این بردارها به عنوان داده ورودی شبکه برای استخراج ویژگی‌ها به کار می‌روند.

شبکه‌های عصبی پیچشی (CNN) نوعی خاص از شبکه‌های عصبی برای پردازش داده هستند که با بردارهای کلمات مانند یک تور[۳] برخورد می‌کنند [۱۰] . صافی‌های[۴] هر لایه کانولوشن بر روی طول ماتریس‌های حاصل از بردارهای ورودی حرکت می‌کند. عرض صافی‌ها به اندازه عرض بردار ورودی (بعد بردار کلمات) و طول آنها معمولاً بین ۲ تا ۵ کلمه است. از بردارهای حاصل نگاشت‌های ویژگی[۵] حاصل می‌شوند که با استفاده از لایه الحاق حداکثری[۶] تبدیل به یک بردار نهایی می‌شوند و از آن به عنوان ورودی لایه آخر برای دسته‌بندی جملات استفاده می‌شود.

۳-۱ کانولوشن

پایه اصلی شبکه پیچشی بردارهای کلمات ورودی $x \in R^s$ هستند که در آن s ابعاد بردارها است و هر سند ورودی به صورت ماتریس $d \in R^{n \times s}$ نمایش داده می‌شود که n تعداد کلمات آن است و هر سطر ماتریس، بردار یک کلمه را نمایندگی می‌کند.

لایه کانولوشن که هدف از آن استخراج ویژگی‌های محلی از رشته‌های حرفی داخل جمله است، تعمیمی از رویکرد پنجره است که در آن چند صافی با اندازه مشخص کل جمله را می‌پیمایند و نتایج حاصل از کانولوشن روی ماتریس‌های مختلف را با هم ترکیب می‌کنند. عرض این صافی‌ها به اندازه ماتریس ورودی و طول آنها قابل تعیین است. هر کدام از این صافی‌ها همان طور که در شکل ۱ مشخص است، یک ماتریس

1. Representation Learning
2. Unsupervised Learning
3. Grid
4. Filters
5. Feature Maps
6. Max-Pooling

هستند که مجموع حاصل‌ضرب عناصر آن و ماتریس ورودی، ماتریس جدیدی است که بردار کلمه ورودی را با توجه به بافتش تغییر می‌دهد. ما یک لایه کانولوشن $W \in R^{h \times s}$ تعریف می‌کنیم که در آن h تعداد کلماتی است که می‌خواهیم روی آن کانولوشن انجام دهیم (طول صافی). اگر کانولوشن را با عملگر * نشان دهیم هر بار پیمایش صافی روی بردار برابر است با

$$W * d_{j:j+h-1} = \sum_{i=j}^{j+h-1} \sum_{k=0}^{s-1} W_{i,k} d_{i,k} \quad (1)$$

سپس صافی به طول h کلمه، $d_{j:j+h-1}$، را با استفاده از تابع غیر خطی f به عدد حقیقی c_j نگاشت می‌کنیم

$$c_i = f(d_{j:j+h-1} + b) \quad (2)$$

که b در آن یک عدد حقیقی است که میزان تمایل[1] را نشان می‌دهد. با اعمال کانولوشن بر روی تمام سند با استفاده از W، یک بردار ویژگی حاصل می‌شود

$$c(W) = [c_1, c_2, \ldots, c_{n-h+1}] \quad (3)$$

هر شبکه می‌تواند دارای مقادیر متفاوت برای نوع صافی‌ها و ماتریس‌های وزنی باشد که به هر کدام کانال[2] گفته می‌شود. در حین مرحله آموزش یک شبکه عصبی پیچشی بر پایه کاربری خاص ایجاد می‌شود که عناصر ماتریس‌های صافی‌ها یاد گرفته می‌شوند. چون هر عنصر ورودی و صافی باید به طور جداگانه ذخیره‌سازی شوند، معمولاً فرض می‌کنیم که این عناصر جز در نقاط محدودی که مقادیر آنها ذخیره شده‌اند، در بقیه نقاط دارای مقدار صفر هستند.

۳-۲ الحاق حداکثری

هدف از لایه الحاق حداکثری، یکسان‌سازی طول بردارهای جملات[3] و کاهش ابعاد بردار خروجی در عین حفظ اطلاعات مهم است. برای مثال اگر ۵۰۰ صافی وجود داشته باشد، بعد از اعمال الحاق حداکثری، برداری ۵۰۰بعدی داریم که به عنوان ورودی با طول ثابت برای بیشینه نرم[4] استفاده می‌شود. همچنین اگر هر صافی را دارای اطلاعات مربوط به یک ویژگی خاص ورودی بدانیم با استفاده از الحاق حداکثری می‌توانیم پیش‌بینی کنیم که آیا این ویژگی در جمله وجود داشته یا نه. کار الحاق حداکثری در این لایه این است که بیشترین مقدار هر ویژگی را از میان صافی‌های مختلف برگزیند

$$\hat{c}_W = \max c(W)_i \quad (4)$$

این روش نسبت به میانگین‌گیری بهتر است چرا که در دسته‌بندی، همه کلمات به یک اندازه مهم نیستند و اهمیت نسبی آنها در لایه الحاق حداکثری لحاظ می‌شود. در نهایت از همه صافی‌ها یک بردار ویژگی سراسری[5] به دست می‌آید که ورودی لایه بعد محسوب می‌شود

$$\hat{C}_W = [\hat{c}_{W^1}, \ldots, \hat{c}_{W^k}]^T \quad (5)$$

در (۵)، $T = \{1, \ldots, k\}$ است. اندازه بردار ویژگی سراسری برای جملات مختلف ثابت است. در لایه الحاق، ما عمل الحاق حداکثری را برای ترکیب ویژگی‌های مختلف حاصل از لایه کانولوشن و به وجود آوردن یک بردار با بعد ثابت انجام می‌دهیم. برای رسیدن به بردار جمله خروجی صافی‌های لایه کانولوشن را پیوند زنجیره‌ای می‌دهیم. هرچه تعداد صافی‌ها بیشتر باشد، تعداد نگاشت‌های ویژگی بیشتر می‌شود و از هر نگاشت ویژگی، بیشترین مقدار در مرحله الحاق حداکثری، انتخاب و به بردار ویژگی سراسری اضافه می‌شود. پس می‌توان گفت که اندازه این بردارها به تعداد نگاشت‌های ویژگی و تعداد کانال‌ها وابسته است. در این مرحله اطلاعات مربوط به مکان قرارگیری ویژگی‌های مختلف را با توجه به در نظر نگرفتن ترتیب کلمات از دست می‌دهیم.

۳-۳ بیشینه نرم

در لایه آخر از تابع بیشینه نرم که نوع گسترش داده شده رگرسیون لجستیک است برای دسته‌بندی بردار ویژگی به دست آمده استفاده می‌کنیم. اگر $U \in R^{2 \times k}$ و $b^U \in R^2$ پارامترهای لایه نرم بیشینه باشند و ورودی وزن‌دار برابر باشد با

$$y_j = U_j \hat{c}_W + b_j^U \quad (6)$$

که در آن c_w بردار ورودی، U_j ردیف jام U، b_j^U عنصر jام b^U و Y_j برچسب jام در ماتریس d است. اندازه این لایه برابر با تعداد برچسب‌ها است. احتمال برچسب خروجی برابر است با

$$P(Y_j = 1 \mid d, W, b^U) = \frac{e^{y_j}}{\sum_i e^{y_i}} \quad (7)$$

اگر D مجموعه ماتریس‌های داده آموزش باشد، برای آموزش دسته‌بندی دوتایی باید (۸) و (۹) را به ترتیب برای داده‌های منفی و مثبت کمینه کرد

$$-\sum_{d \in D} \log(P(Y_{POS}^d - 1 \mid d, W, b^U)) \quad (8)$$

$$-\sum_{d \in D} \log(P(Y_{POS}^d - 2 \mid d, W, b^U)) \quad (9)$$

پارامترهای شبکه عصبی پیچشی (d, W, b^U) که (۸) و (۹) را کمینه می‌کنند با محاسبه گرادیان از طریق روش پس‌انتشار به دست می‌آیند. شکل کلی معماری CNN پیشنهادی در شکل ۱ قابل مشاهده است.

۳-۴ جاسازی کلمات

بر خلاف بسیاری از روش‌های دیگر در یادگیری ماشین، ورودی در اینجا به صورت متنی نیست. این به آن معناست که برای تهیه ورودی، متن باید به بردارهای ویژگی یا به عبارت دیگر به جاسازی‌های کلمات تبدیل شوند که این جاسازی‌ها، اگر درست استخراج بشوند خود حاوی اطلاعات بافتی و معنایی متن هستند [۲۸].

این بردارها با آموزش‌دادن شبکه عصبی بر مبنای پیکره متنی حاصل می‌شوند و فرایندی زمان‌بر هستند. با استفاده از یادگیری عمیق، یک مدل زبانی برای یادگیری نمایش توزیع‌یافته کلمات[6] بر اساس سه ایده کلی زیر می‌توان ارائه کرد [۲۹]:

- هر کلمه در پیکره به یک بردار ویژگی sبعدی حاوی اعداد حقیقی متناظر می‌شود.
- تابع احتمال توأم برای کلمات با استفاده از این نمایش‌های برداری بیان می‌شود.

1. Bias
2. Channel
3. Sentence Vector
4. Softmax
5. Global Feature Vector
6. Distributed Representation of Words

جدول ۱: الگوهای مشخص برای توییت‌ها.

نتیجه	مثال	محتوا
⟨کاربر⟩	@username	نام کاربری
⟨آدرس⟩	http://t.co/url	URLs
⟨عدد⟩	2007/۱۳۸۹	اعداد
⟨هشتگ⟩	#tweet	هشتگ‌ها
یا	|	خط مورب

- یادگیری بردارهای ویژگی کلمات و پارامترهای تابع احتمال به طور هم‌زمان انجام می‌شوند.

هر کدام از جاسازی‌های کلمات می‌تواند ابعادی به دلخواه کاربر داشته باشد. بعد بالاتر به معنای این است که اطلاعات بیشتری ضبط شده ولی در عین حال با زیادشدن بعد، هزینه‌های محاسباتی نیز افزایش می‌یابد.

٤- ارزیابی راهکار پیشنهادی

برای ارزیابی راهکار پیشنهادی و مقایسه نتایج آن با روش‌های سنتی یادگیری ماشین، آن را با پارامترهای مختلف بر روی مجموعه داده رسانه‌های اجتماعی فارسی آزمایش می‌کنیم.

٤-١ مجموعه داده

برای آموزش و ارزیابی مدل از مجموعه داده‌های مختلف استفاده شده است:

- مجموعه داده سنتی‌پرس: این مجموعه ۱۱۰۰ نظر درباره محصولات است که از فروشگاه برخط دیجی‌کالا جمع آوری شده که شامل جملات فارسی با برچسب‌های حاوی بار معنایی است که در پردازش زبان طبیعی و به طور مشخص در زمینه تحلیل احساس یا عقیده‌کاوی کاربرد دارد [۳۰].
- مجموعه داده توئیتر فارسی: ۱۱۱۶۰ توئیت جمع‌آوری شده و برچسب خورده در دو و پنج سطح برای ارزیابی و آموزش در بازه بین فروردین تا تیر ۱۳۹۶ (برچسب‌ها به صورت دستی و توسط افراد متخصص زبان فارسی ایجاد شده است).
- مجموعه داده اخبار: ۱۶۴۳ خبر فارسی از ۳۰ منبع خبر فارسی در بازه بین فروردین تا تیر ۱۳۹۶ جمع‌آوری شده و خبرها نیز به صورت دستی برچسب خورده است. برای مجموعه داده توئیتر، ۱۵۰۰۰ توییت از سایت شبکه اجتماعی توئیتر جمع‌آوری شد. ابتدا برای پردازش به آرایه‌ای از توکن‌ها تبدیل شدند. به دلیل این که ما تنها به تحلیل زبان فارسی می‌پردازیم حروف زبان‌های دیگر از متون حذف شدند. توییت‌ها با الگوی مشخص‌شده در جدول ۱ اصلاح شدند. استفاده از این الگوها برای یکسان‌سازی توکن‌های داده‌ها و بردارهای پیش‌آموخته بوده است.

توییت‌ها و اخبار با استفاده از ۳ نفر و در پنج دسته برچسب‌گذاری شدند. داده‌هایی که حداقل ۲ نفر به آن یک برچسب داده بودند برای آموزش و ارزیابی پنج‌تایی در نظر گرفته شدند. اگر ۲ نفر به یکی از داده‌ها برچسبی از یک نوع جنسیت ولی با شدت متفاوت داده بودند (مثلاً مثبت احساسی و مثبت منطقی)، آن داده برای دسته‌بندی دوتایی به کار گرفته شد. در نهایت ۹۹۶۳ توییت و ۱۱۵۴ خبر برای دسته‌بندی پنج‌تایی و ۱۱۱۶۰ توییت و ۱۶۴۳ خبر برای دسته‌بندی دوتایی باقی ماند. در جدول ۲ اطلاعات مربوط به تعداد توکن‌ها و پراکندگی آنها در دسته‌های مختلف نمایش داده شده است.

ما از بردارهای پیش‌آموخته ویکیپدیای فارسی برای دستیابی به بردارهای جاسازی کلمات استفاده کردیم که بر روی تعداد همآیی کلمات در متن، آموزش دیده شده است [۳۱].

۲۰٪ هر کدام از مجموعه داده‌ها به عنوان داده ارزیابی و ۸۰٪ بقیه به عنوان داده آموزش در نظر گرفته شده است.

٤-٢ پیاده‌سازی

برای استفاده کامل از منابع محاسباتی GPU، پیاده‌سازی در محیط Keras و Theano [۳۲] که چارچوبی سطح بالا در زبان پایتون برای پیاده‌سازی شبکه‌های عصبی عمیق است صورت گرفت. برای تنظیم پارامترها از تابع جستجوی شبکه‌ای scikit-learn [۳۳] که می‌تواند با استفاده از تمام ترکیبات پارامترهای احتمالی بهترین عملکرد را شناسایی کند، استفاده شده است. آموزش همراه با توقف اولیه صورت گرفته است به این معنا که اگر زیان اعتبارسنجی[1] بعد از پنج دور افزایش نیابد، پردازش متوقف می‌شود. برای آموزش همه مجموعه داده‌ها از صافی‌هایی سه‌تایی و ابعاد ۱۰۰ برای بردارهای ورودی استفاده شده است. در مرحله آموزش تعداد تکرار برابر ۱۰۰ در نظر گرفته شد. تعداد دسته‌ها دو و پنج در نظر گرفته شده و برای تنظیم‌کردن از دراپ‌اوت[2] با نرخ ۰٫۵ در لایه ماقبل آخر استفاده شده است. آموزش با استفاده از گرادیان کاهشی اتفاقی و قانون به روز رسانی آدادلتا[3] صورت گرفته است. برای رسیدن به تحلیل مناسب شرایط تصادفی (مقداردهی اولیه برای شبکه و جاسازی کلمات و دیگر پارامترها) برای همه مدل‌ها یکسان در نظر گرفته شده است.

٤-٣ معیار ارزیابی

برای ارزیابی از مساحت سطح زیر نمودار دوبعدی که در آن نرخ تشخیص صحیح دسته مثبت روی محور Y و نرخ تشخیص غلط دسته منفی روی محور X رسم می‌شود استفاده می‌کنیم که به آن مساحت زیر منحنی (AUC)[4] می‌گوییم. هرچه عدد زیر نمودار بزرگ‌تر باشد دسته‌بندی صورت‌گرفته دقیق‌تر بوده است.

٤-٤ نتایج

نتایج مدل‌های ما با پارامترهای متفاوت برای دو و پنج دسته در مقایسه با مدل‌های یادگیری ماشین سنتی و شبکه‌های عصبی بازگشتی در جدول ۳ آمده است. ارزیابی‌ها با معیار مساحت زیر نمودار نشان می‌دهند که مدل شبکه‌های پیچشی به طور قابل ملاحظه‌ای عملکرد بهتری در هر دو دسته نسبت به روش‌های پراستفاده سنتی یادگیری ماشین و شبکه‌های عصبی بازگشتی دارند. تعدد لایه‌ها برای پردازش که ویژگی‌های سطح بالا را برخلاف روش‌های سنتی یادگیری ماشین بدون نظارت استخراج می‌کنند و وجود لایه الحاق حداکثری که نمونه‌های مناسبی از ویژگی‌ها را انتخاب می‌کند، به بهبود نتایج کمک کرده‌اند. با وجود این که برخلاف بسیاری روش‌های دیگر یادگیری ماشین شبکه عصبی پیچشی ترتیب کلمات را محسوب نمی‌کند، صافی‌های موجود شبکه تا سه کلمه و بیشتر را در کنار هم در نظر می‌گیرند که محاسبه آن در روش‌های سنتی بسیار دور از ذهن است. این موضوع به خصوص در جملاتی که در آنها، بین بخشی از عبارتی خاص با قرارگرفتن کلمات

1. Validation Loss
2. Dropout
3. Adadelta
4. Area Under Curve

جدول ۲: تعداد توکن‌ها و پراکندگی آنها در توییت‌های مجموعه داده توئیتر فارسی پس از برچسب زنی دو سطحی و پنج سطحی.

دوتایی			پنج‌تایی				
مثبت	منفی	مثبت احساسی	مثبت منطقی	خنثی	منفی منطقی	منفی احساسی	ویژگی‌ها
۵۲۹۷	۵۸۶۳	۱۷۸۲	۱۶۹۴	۳۲۵۴	۱۵۸۶	۱۶۴۷	تعداد توییت
۸۴۹۴۵	۹۳۸۰۵	۲۷۶۵۸	۲۷۸۹۵	۵۱۰۹۸	۲۵۲۶۹	۲۶۸۴۶	تعداد توکن
۱۶٫۰۳	۱۵٫۹۹	۱۵٫۵۲	۱۶٫۴۶	۱۵٫۷	۱۵٫۹۳	۱۶٫۲۹	میانگین توکن بر توییت
۷۰۸۹	۷۸۶۵	۵۹۱۲	۵۶۷۸	۷۱۲۱	۵۱۰۵	۵۲۷۰	تعداد واژگان

جدول ۳: نتایج ارزیابی مدل‌ها برای دسته‌بندی ۲ و ۵ تایی روی داده‌های مختلف با معیار مساحت زیر نمودار.

مدل	پنج‌تایی			دوتایی		
	توئیتر	اخبار	نظرات	توئیتر	اخبار	نظرات
CNN	۰٫۴۳	۰٫۴۶	۰٫۴۳	۰٫۷۵	۰٫۷۹	۰٫۷۳
RNN	۰٫۳۹	۰٫۴۱	۰٫۴۰	۰٫۶۳	۰٫۷۴	۰٫۶۹
بیز ساده	۰٫۳۹	۰٫۴۰	۰٫۳۸	۰٫۵۲	۰٫۶۲	۰٫۵۶
آنتروپی بیشینه	۰٫۳۹	۰٫۳۹	۰٫۳۸	۰٫۶۱	۰٫۷۱	۰٫۶۹
SVM	۰٫۴۰	۰٫۴۰	۰٫۴۱	۰٫۶۳	۰٫۷۴	۰٫۶۸

جدول ۴: نتایج ارزیابی مدل‌ها برای دسته‌بندی ۲ و ۵ تایی روی داده‌های مختلف با در نظر گرفتن تأثیر تعداد صافی‌ها.

اندازه صافی	پنج‌تایی			دوتایی		
	نظرات	توئیتر	اخبار	توئیتر	اخبار	نظرات
دو	۰٫۳۹	۰٫۴۳	۰٫۴۴	۰٫۷۴	۰٫۶۸	۰٫۶۸
سه	۰٫۴۰	۰٫۴۳	۰٫۴۶	۰٫۷۵	۰٫۷۹	۰٫۶۹
پنج	۰٫۴۱	۰٫۴۲	۰٫۴۵	۰٫۶۹	۰٫۷۹	۰٫۷۲
هفت	۰٫۴۰	۰٫۴۲	۰٫۴۶	۰٫۶۳	۰٫۷۹	۰٫۷۳
نه	۰٫۴۰	۰٫۴۰	۰٫۴۳	۰٫۶۳	۰٫۷۹	۰٫۷۲

دیگر فاصله افتاده به کار می‌آید.

به طور کلی افزایش بعد و تعداد تکرار در مرحله آموزش بردارهای ورودی کلمات به بهترشدن کارایی مدل منجر می‌شود. این بردارها که از متون ویکی‌پدیای فارسی به دست آمده‌اند، توانایی ذخیره‌کردن اطلاعات مربوط به روابط معنایی موجود در بردارهای کلمات را دارند که محصول نمایش توزیع‌یافته آنهاست. این ویژگی توانایی شبکه عصبی پیچشی را برای کشف روابط معنایی ترکیبی اجزای جمله نسبت به روش‌های دیگر یادگیری ماشین افزایش می‌دهد. همچنین نتایج به دست آمده نشان‌دهنده کارایی بالاتر شبکه‌های عصبی پیچشی بر روی رسانه‌های اجتماعی با طول متن کوتاه‌تر (توئیتر) نسبت به روش‌های سنتی یادگیری ماشین و شبکه‌های عصبی بازگشتی است. اگرچه عملکرد مدل پیشنهادی در محیط زبان غیر رسمی و با علایم نگارشی و غیر نگارشی مرسوم با افت روبه‌رو می‌شود اما در تحلیل متون خبری، بردارهای عمومی به دست آمده از ویکی‌پدیای فارسی به دلیل استفاده از زبان رسمی، کارایی بهتری نسبت به دو داده دیگر نشان داده‌اند.

بدیهی است که علاوه بر بردارهای کلمات عمومی، استفاده از بردارهای کار ویژه (در اینجا بردارهای مخصوص تحلیل احساس) به افزایش دقت مدل کمک خواهد کرد. همین طور در کار حاضر برای به دست آوردن بردارهای جملات ما از تجمیع بردارهای کلمات استفاده کردیم. برای بهبود کیفیت نمایش برداری متن می‌توان از روش‌های دیگری چون بردارهای پاراگراف که قادر است متون با طول‌های مختلف را نمایندگی کند، بهره گرفت.

جدول ۵: نتایج ارزیابی مدل‌ها برای دسته‌بندی ۲ و ۵ تایی روی داده‌های مختلف با در نظر گرفتن تأثیر ابعاد بردارهای کلمات.

ابعاد بردارها	پنج‌تایی			دوتایی		
	نظرات	توئیتر	اخبار	توئیتر	اخبار	نظرات
۱۰	۰٫۴۱	۰٫۴۰	۰٫۳۳	۰٫۶۹	۰٫۶۶	۰٫۶۱
۵۰	۰٫۴۳	۰٫۴۵	۰٫۴۳	۰٫۷۵	۰٫۷۸	۰٫۷۲
۱۰۰	۰٫۴۳	۰٫۴۳	۰٫۴۶	۰٫۷۵	۰٫۷۹	۰٫۷۳
۲۰۰	۰٫۴۲	۰٫۴۵	۰٫۴۴	۰٫۷۵	۰٫۷۸	۰٫۷۳
۳۰۰	۰٫۴۲	۰٫۴۵	۰٫۴۳	۰٫۷۴	۰٫۷۸	۰٫۷۲

٤-٥ تأثیر تعداد صافی‌ها بر نتایج

در این بخش تاثیر اندازه صافی را بر روی داده‌های مختلف را اندازه گیری می‌کنیم. برای اینکار بعد بردارها را ۱۰۰ در نظر می‌گیرم و هر آزمایش را برای دقت بیشتر ۱۰۰ بار تکرار کرده و از نتایج میانگین گرفته می‌شود. اندازه‌های صافی را برابر ۲، ۳، ۵، ۷ و ۹ می‌گیریم. نتایج برای دو و پنج دسته در جدول ۴ آمده است. نتایج به دست آمده نشان می‌دهد که مدل با صافی‌های دو و سه بر روی داده توئیتر بهتر عمل می‌کند. با افزودن بر اندازه صافی‌ها کیفیت مدل بر روی توئیتر بهبود پیدا نمی‌کند. با افزایش طول جملات به خصوص در دسته‌بندی پنج‌تایی، نیاز به صافی‌های بزرگ‌تر احساس می‌شود. عملکرد مدل برای دسته‌بندی نظرات و اخبار با بزرگ‌شدن صافی و رسیدن به اندازه پنج در هر دو بخش دو و پنج‌تایی رشد می‌کند. با تغییر صافی به هفت و نه نتایج بر روی داده نظرات و اخبار بهبود محسوسی پیدا نمی‌کند. مشاهدات ما نشان می‌دهد که بهتر است پیش از آموزش، مدل صافی با اندازه مناسب از راه آزمون و خطا به دست آید. در هر مدل هر صافی را می‌توان به تنهایی یا به صورت ترکیبی با صافی‌های دارای اندازه نزدیک به کار برد.

٤-٦ تأثیر ابعاد بردارهای کلمات بر نتایج

با قراردادن اندازه صافی روی سه و تعداد تکرار ۱۰۰، ما ابعاد را برابر اعداد ۱۰، ۵۰، ۱۰۰، ۲۰۰ و ۳۰۰ قرار دادیم. بهترین عدد برای بعد انتخابی به نوع مجموعه داده بستگی دارد.

از نتایج نشان داده شده در جدول ۵ مشاهده می‌شود که عملکرد مدل با افزایش ابعاد بردارها به بیش از ۱۰ بهبود چشم‌گیری پیدا می‌کند. این موضوع بر روی داده اخبار که بیشترین طول را دارد محسوس‌تر است. می‌توان نتیجه گرفت افزایش ابعاد به بیش از ۲۰۰ در کار حاضر تأثیر چندانی در نتایج ندارد و حتی ممکن است کارایی را کاهش دهد. همین طور با افزایش ابعاد بردارهای کلمات زمان مورد نیاز برای آموزش آنها به طور قابل توجهی افزایش می‌یابد. در کاربرد به نظر می‌رسد تعداد دسته‌ها و ابعاد بردار رابطه قابل مشاهده‌ای ندارند و برای انواع دسته‌بندی‌ها بازه ۱۰۰ تا ۲۰۰ مناسب است.

پیداکردن بازه مناسب برای داده‌های آموزش با اندازه بزرگ‌تر از داده حاضر، نیاز به بررسی بازه‌های متفاوت دارد.

جدول ۶: نتایج ارزیابی مدل‌ها برای دسته‌بندی ۲ و ۵ تایی روی داده‌های مختلف با در نظر گرفتن تأثیر تعداد تکرار در آموزش بردار کلمات.

تعداد تکرار آموزش	پنج‌تایی			دوتایی		
	نظرات	توئیتر	اخبار	توئیتر	اخبار	نظرات
۲۵	۰٫۳۹	۰٫۴۰	۰٫۴۲	۰٫۷۱	۰٫۷۲	۰٫۶۹
۵۰	۰٫۴۳	۰٫۴۳	۰٫۴۳	۰٫۷۵	۰٫۷۸	۰٫۷۲
۱۰۰	۰٫۴۳	۰٫۴۳	۰٫۴۶	۰٫۷۵	۰٫۷۹	۰٫۷۳
۲۰۰	۰٫۴۴	۰٫۴۵	۰٫۴۲	۰٫۷۵	۰٫۷۸	۰٫۷۲

۴-۷ تأثیر تعداد تکرار در آموزش بردار کلمات بر نتایج

برای بررسی عدد مناسب تکرار در مرحله آموزش بردارهای کلمات ما ابعاد بردار کلمات را ۱۰۰ و اندازه صافی را ۳ در نظر می‌گیریم. اعداد تکرار ما ۲۵، ۵۰، ۱۰۰ و ۲۰۰ است.

از نتایج نمایش داده شده در جدول ۶ مشاهده می‌شود که تعداد تکرار بیشتر از ۲۵ ارتباط مشخصی با انواع داده با طول‌های مختلف ندارد. همین طور نتایج دسته‌بندی‌های مختلف با تعداد دسته مختلف ارتباط معناداری با تعداد تکرار در آموزش بردارها ندارد. تکرار بیشتر از ۱۰۰ نتایج را تغییر چندانی نمی‌دهد. بدیهی است که با افزایش تکرار، زمان طی‌شده برای آموزش بردارها افزایش می‌یابد. تکرار ۱۰۰ نتایج بهتری در مجموع دارد اما با توجه به فاصله بسیار نزدیکی که با تعداد تکرار ۵۰ دارد و با توجه به مدت زمان طی‌شده برای آموزش برای داده‌ای به حجم مجموعه ما چندان به صرفه به نظر نمی‌رسد. اصولاً برای زبان‌ها با منابع محدود تعداد تکرار بین ۵۰ تا ۱۰۰ پیشنهاد می‌شود.

۵- نتیجه‌گیری و پیشنهادها

در کار حاضر ما شبکه عصبی پیچشی با پارامترهای متفاوت را با استفاده از بردارهای کلمات بر روی رسانه‌های اجتماعی جهت تحلیل احساس متن فارسی به کار بردیم. آموزش بردارهای عمومی در شبکه عصبی با یک لایه کانولوشن کارایی بهتری نسبت به روش‌های سنتی یادگیری ماشین و شبکه‌های عصبی بازگشتی به خصوص بر روی داده‌ها با طول کوتاه نشان داد. نتایج ما نشان داد که بردارهای کلمات استخراج‌شده از داده‌های عمومی بدون توجه به نوع کاربری می‌توانند به بهبود نتایج در پردازش زبان طبیعی کمک کنند. برای تحقیقات آتی، در نظر گرفتن بردارهای کار ویژه پیشنهاد می‌شود. همچنین در کار حاضر، ما تنها به دسته‌بندی در سطح جملات پرداختیم. برای کارهای آینده می‌توان به سطوح بالاتر از جمله، مثل کل سند و تحلیل احساس بر اساس موجودیت‌های مختلف که منجر به تحلیل دقیق‌تر برای جملات و اسناد دارای نظرات با جهت‌گیری متفاوت می‌شود پرداخت.

مراجع

[1] S. Greenwood, A. Perrin, and M. Duggan, "Social media update 2016: facebook usage and engagement is on the rise, while adoption of other platforms holds steady," Pew Research Center, 2016.

[2] J. Mander, "Daily time spent on social networks rises to 1.72 hours," London: Global Web Index, 2015.

[3] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1-167, May 2012.

[4] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *Proc. of the ACL-02 Conf. on Empirical Methods in Natural Language Processing*, *Association for Computational Linguistics*, vol. 10, pp. 79-86, Philadelphia, PA, USA, 6-7 Jul. 2002.

[5] A. Tripathy, A. Agrawal, and S. Kumar Rath, "Classification of sentiment reviews using n-gram machine learning approach," *Expert Systems with Applications*, vol. 57, pp. 117-126, Sept. 2016.

[6] T. Mullen and N. Collier, "Sentiment analysis using support vector machines with diverse information sources," in *Proc. 9th Conference on Empirical Methods in Natural Language Processing*, *EMNLP'04*, vol. 4, pp. 412-418, Jan. 2004.

[7] B. Agarwal and N. Mittal, "Machine learning approach for sentiment analysis," In: *Prominent Feature Extraction for Sentiment Analysis*, pp. 21-45, Dec. 2016.

[8] S. Poria, H. Peng, A. Hussain, N. Howard, and E. Cambria, "Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis," *Neurocomputing*, vol. 261, pp. 217-230, Oct. 2017.

[9] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent twitter sentiment classification," in *Proc. 52nd Annual Meeting of the Association for Computational Linguistics*, *ACL'14*, vol. 2, pp. 49-54, Baltimore, MD, USA, Jun. 2014.

[10] Y. Kim, Convolutional Neural Networks for Sentence Classification, arXiv preprint arXiv:1408.5882, 2014.

[11] Y. Zhang and B. Wallace, A Sensitivity Analysis of (and Ppractitioners' Guide to) Convolutional Neural Networks for Sentence Classification, arXiv preprint arXiv:1510.03820, 2015.

[12] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, pp. 620, May 1957.

[13] M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," in *Proc. IEEE 4th Int. Conf. on, Computing, Communications and Networking Technologies, ICCCNT'13,* 5 pp., Tiruchengode, India, 4-6 Jul. 2013.

[14] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, Sept.. 1995.

[۱۵] م. ع. زارع چاهوکی و س. ح. ر. محمدی، "بهینه‌سازی هسته‌های چندگانه در ماشین بردار پشتیبان جفتی برای کاهش شکاف معنایی تشخیص صفحات فریب‌آمیز،" *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۶، شماره ۴، صص. ۱۳۵-۱۴۵، زمستان ۱۳۹۵.

[16] C. N. Dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in *Proc. of, the 25th Int. Conf. on Computational Linguistics,* COLING'14, pp. 69-78, Dublin, Ireland, 25–29 Aug. 2014.

[17] Y. Zhang, M. Chen, L. Liu, and Y. Wang, "An effective convolutional neural network model for Chinese sentiment analysis," in *Proc. AIP Conf. Proc.,* vol. 1836, pp. 020084, Rome, Italy, 27-29 Jan. 2017.

[18] M. Cieliebak, J. Deriu, D. Egger, and F. Uzdilli, "A twitter corpus and benchmark resources for german sentiment analysis," in *Proc. of the 5th Ine. Workshop on Natural Language Processing for Social Media, SocialNLP*, pp. 45-51, Boston, USA, Dec. 2017.

[19] R. Socher, A. Perelygin, J. Y. Wu, J. Chuang, C. D. Manning, A. Y. Ng, and C. Potts, "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proc. of the Conf. on Empirical Methods in Natural Language Processing, EMNLP'13*, vol. 1631, pp. 1631-1642, Seattle, WA, USA, 18-21 Oct. 2013.

[۲۰] م. ح. رفان، م. کمرزرین و ع. دمشقی، "بهبود دقت و پایداری RTDGPS با استفاده از مدل ترکیبی RNN و PSO،" *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۶، شماره ۱، صص. ۱۸۵-۱۹۶، بهار ۱۳۹۵.

[21] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *J. of Machine Learning Research*, vol. 12, no. 76, pp. 2493-2537, Aug. 2011.

[22] A. Bagheri and M. Saraee, "Persian sentiment analyzer: a framework based on a novel feature selection method," *International J. of Artificial Intelligence™*, vol. 12, no. 2, pp. 115-129, Nov. 2014.

[۲۳] ح. اکبریان، م. صالحی و ه. ویسی، "تعیین جهت‌گیری نظرات در رسانه‌های اجتماعی فارسی‌زبان،" *ارائه‌شده در بیست و چهارمین کنفرانس مهندسی برق ایران*، ۶ صص.، شیراز، ایران، ۲۳-۲۱ اردیبهشت ۱۳۹۵.

[24] M. E. Basiri, A. R. Naghsh-Nilchi, and N. Ghassem-Aghaee, "A framework for sentiment analysis in Persian," *Open Trans. on Information Processing*, vol. 1, no. 3, pp. 1-14, Nov. 2014.

[25] M. S. Hajmohammadi and R. Ibrahim, "A SVM-based method for sentiment analysis in persian language," in *Proc. Int. Conf. on Graphic and Image Processing, ICGIP'12,* vol. 8768, 5 pp., Singapore, Singapore, 5-7 Oct. 2013.

مرتضی روحانیان سال ۱۳۹۷ کارشناسی ارشد خود را در رشته زبان‌شناسی رایانشی دانشگاه تهران به پایان رسانید. وی هم‌اکنون به عنوان دانشجوی دکتری دانشگاه کوئین مری انگلستان مشغول به تحصیل است. زمینه‌های پژوهشی مورد علاقه ایشان زبان شناسی رایانشی و یادگیری ماشین می‌باشد.

مصطفی صالحی کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر به ترتیب در سال‌های ۱۳۸۴ و ۱۳۸۶ به پایان رسانید. او در سال ۱۳۹۱ موفق به دریافت درجه دکتری در همین رشته از دانشگاه صنعتی شریف شد. پس از آن پسادکتری خود را به‌ترتیب در سال‌های ۱۳۹۴ و ۱۳۹۵ در دانشگاه بولونیا کشور ایتالیا و دانشگاه تلکام سود پاریس (فرصت مطالعاتی) گذراند. وی هم‌اکنون به عنوان عضو هیأت علمی گروه بین‌رشته‌ای فناوری دانشگاه تهران، با مرتبه دانشیاری مشغول به فعالیت است. زمینه‌های پژوهشی ایشان شامل شبکه‌های اجتماعی مجازی و علم داده می‌باشد.

علی درزی زبان‌شناس ایرانی و استاد گروه زبان‌شناسی دانشگاه تهران است. او در سال ۱۳۶۶، از پایان‌نامه کارشناسی ارشد خود دفاع کرد و در سال ۱۳۷۴ دوره دکتری زبان‌شناسی را در دانشگاه ایلینوی در اوربانا شامپاین، آمریکا فارغ‌التحصیل شد.

وحید رنجبر کارشناسی و کارشناسی ارشد خود را در رشته مهندسی فناوری اطلاعات به‌ترتیب در سال‌های ۱۳۹۰ و ۱۳۹۲ به پایان رسانید. وی در سال ۱۳۹۷ دکتری تخصصی خود را در رشته فناوری اطلاعات از دانشگاه تهران دریافت کرد. وی هم‌اکنون به عنوان عضو هیأت علمی دانشکده مهندسی کامپیوتر دانشگاه یزد مشغول به فعالیت است. زمینه‌های پژوهشی مورد علاقه ایشان تحلیل شبکه‌های اجتماعی، یادگیری ماشین و کلان داده می‌باشد.

[26] B. Roshanfekr, S. Khadivi, and M. Rahmati, "Sentiment analysis using deep learning on Persian texts," in *Iranian Conf. on Electrical Engineering, ICEE'17,* pp. 1503-1508, Tehran, Iran, 2-4 May 2017.

[27] K. Wang, X. Wang, L. Lin, M. Wang, and W. Zuo, "3D human activity recognition with reconfigurable convolutional neural networks," in *Proc. of the 22nd ACM Int Conf. on Multimedia,* pp. 97-106, Orland, FL, USA, 18-19 Jun. 2014.

[28] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. Advances in Neural Information Processing Systems*, *NIPS'13,* pp. 3111-3119, Lake Tahoe, CA, USA, 5-10 Dec. 2013.

[29] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, "A neural probabilistic language model," *J. of Machine Learning Research*, vol. 3, no. 6, pp. 1137-1155, Feb. 2003.

[۳۰] پ. حسینی، ع. احمدیان رامکی، ح. ملکی، م. انواری و س. ا. میرروشن‌دل، "پیکره فارسی تحلیل احساس سنتی‌پرس،" *مجموعه مقالات سومین همایش ملی زبان‌شناسی رایانشی*، ، ۸ صص.، تهران، ایران، آبان ۱۳۹۳.

[31] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, *Enriching Word Vectors with Subword Iinformation*, arXiv preprint arXiv:1607.04606, 2016.

[32] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, "Theano: a CPU and GPU math compiler in python," in *Proc. 9th Python in Science Conf.*, 7 pp., Austin, Texas, 28 Jun.-3 Jul. 2010.

[33] F. Pedregosa*et al.*, "Scikit-learn: machine learning in python," *J. of Machine Learning Research*, vol. 12, pp. 2825-2830, Oct. 2011.