



دانشگاه کاشان
University of Kashan

مجله محاسبات نرم

SOFT COMPUTING JOURNAL

تارنمای مجله: scj.kashanu.ac.ir



بکارگیری مدل تحلیل احساسات در سطح حروف مبتنی بر شبکه عصبی روی نظرات فارسی ثبت شده در شبکه‌های اجتماعی و فروشگاه‌های اینترنتی

امید خلف بیگی^۱، دانشجوی کارشناسی ارشد، سید علیرضا بشیری موسوی^{۲*}، استادیار، سینا قارلقی^۳، دانشجوی کارشناسی گروه مهندسی برق و کامپیوتر، دانشگاه خوارزمی، تهران، ایران.
^۳ گروه مهندسی برق و کامپیوتر، مرکز آموزش عالی فنی و مهندسی بوئین زهرا، قزوین، ایران.

چکیده

اطلاعات مقاله

تاریخچه مقاله:

دریافت ۲۴ شهریور ماه ۱۴۰۱
پذیرش ۱۶ اردیبهشت ماه ۱۴۰۲

کلمات کلیدی:

پردازش زبان طبیعی
تشخیص احساسات
مدل مبتنی بر محتوا
شبکه عصبی عمیق
بسترهای اینترنتی

امروزه با توجه به تمایل روزافزون مردم برای خرید اجناس از طریق فروشگاه‌های اینترنتی و شبکه‌های مجازی، شاهد افزایش داده‌های بدون ساختار مانند متن در سطح اینترنت هستیم. لذا پردازش متون و توسعه الگوریتم‌های کارآمد جهت استخراج دانش، توجه پژوهشگران حوزه علوم داده را در بسترهای مذکور به خود جلب کرده است. از رویکردهای پردازش متن می‌توان به دسته‌بندی جملات به گروه‌های احساسی متفاوت با استفاده از الگوریتم‌ها و روش‌های گوناگون اشاره کرد. در پژوهش حاضر، چارچوبی برای دسته‌بندی نظرات، مبتنی بر احساسات کاربران توسعه داده شده است که از پردازش در سطح حروف بهره می‌برد. از این رو در چارچوب پیشنهادی، از معماری تعبیه شده در مدل‌های زبانی استفاده شده است که لایه‌های چهارگانه تعبیه (جهت انتقال حروف به فضای برداری)، پیچش یک بعدی (جهت استخراج بردار ویژگی برای هر واژه)، نگاشت و شبکه عصبی بازگشتی را شامل می‌شود. در چارچوب پیشنهادی، ابتدا با بکارگیری لایه تعبیه در سطح حروف، برداری ثابت برای آنها تعیین می‌شود. سپس، مبتنی بر عملگرهای پیچش یک بعدی که به صورت موازی بکارگیری شده‌اند، ارتباط معنایی و منطقی بین حروف تشکیل‌دهنده هر واژه به دست آمده و بردار ۱۲۸ بعدی برای هر لغت، حاصل می‌شود. پس از دستیابی به بردارهای واژگان، با استفاده از دو معماری شبکه‌های عصبی بازگشتی، ارتباط بین واژگان کشف شده و احساس مرتبط با دیدگاه، تعیین می‌شود. نتایج حاصل از بکارگیری مدل پیشنهادی بر روی مجموعه نظرات مبتنی بر سنج‌های Accuracy و F-score، به ترتیب ۷۹/۸۷٪ و ۷۹/۹۰٪ می‌باشد.

© ۱۴۰۲ نویسندگان. مقاله با دسترسی آزاد تحت مجوز CC-BY

۱. مقدمه

پس از همه‌گیری ویروس کرونا، تمام ابعاد زندگی انسان‌ها تحت

تأثیر قرار گرفته است. از مهمترین جنبه‌های شایان ذکر، می‌توان به فعالیت‌های تجاری و کسب و کارهای خرد اشاره کرد که با پیشرفت تکنولوژی، گستره وسیعی از وظایف نیروی انسانی، تقلیل یافته و به صورت از راه دور و خودکار انجام می‌شوند. همچنین، شیوه فروش و ارائه خدمات از رویکرد سنتی فاصله گرفته و مبتنی بر فناوری‌های دیجیتالی، به سیر تحولی خود ادامه

* نوع مقاله: پژوهشی

نویسنده مسئول

پست(های) الکترونیک: omidkhalafbeigi903@gmail.com (خلف بیگی)

abashirimosavi@bzte.ac.ir (بشیری موسوی)

sina.gharloghi99@gmail.com (قارلقی)

نحوه ارجاع به مقاله: خلف بیگی، امید، بشیری موسوی، سید علیرضا، قارلقی، سینا، «بکارگیری مدل تحلیل احساسات در سطح حروف مبتنی بر شبکه عصبی روی نظرات فارسی ثبت شده در شبکه‌های اجتماعی و فروشگاه‌های اینترنتی»، مجله محاسبات نرم، جلد ۱۱، شماره ۲، ص ۱۱۸-۱۳۳، پاییز و زمستان ۱۴۰۱.

می‌دهد. برای مثال می‌توان به تجارت‌های دیجیتال محور از قبیل وبگاه‌های شخصی و فروشگاه‌هایی مانند دیجی‌کالا و شبکه‌های اجتماعی مانند اینستاگرام اشاره کرد. با توسعه بسترهای مذکور، زمینه لازم جهت تبادل نظر کاربران و تولید حجم انبوهی از داده‌های بدون ساختار فراهم شده است که طیف گسترده‌ای از اطلاعات همچون میزان فروش محصولات، بازخورد مشتریان و تبادل نظر کاربران درباره وقایع سیاسی و مذهبی را دربرمی‌گیرد. همزمان با فعالیت کاربران در اینترنت به زبان‌های پرکاربرد دنیا همچون انگلیسی، فرانسوی، آلمانی و عربی، میزان استفاده فارسی زبانان مانند ایرانیان، تبعه‌های افغانی و پاکستانی از این محیط نیز افزایش پیدا کرده است که نتیجه چنین پیشامدی، رشد چشمگیر دیدگاه‌ها به زبان فارسی می‌باشد. با افزایش میزان دیدگاه‌های فارسی، فرصت‌های تحقیقاتی و تجاری گسترده‌ای برای محققان فارسی زبان ایجاد شده است. از سوی دیگر، تعداد پژوهش‌های صورت گرفته در زمینه پردازش متون فارسی در مقایسه با سایر زبان‌های مذکور اندک بوده که دلیل آن را می‌توان تعداد محدود داده‌ها و عدم انتشار عمومی آنها توسط شرکت‌ها یا گروه‌های فعال فارسی زبان در این حوزه، دانست.

پردازش زبان‌های طبیعی، موضوعی بین رشته‌ای بوده که به مطالعه توانایی ماشین در پردازش، درک زبان انسان‌ها و در صورت نیاز، تعامل با آنها می‌پردازد. از زیرشاخه‌های این حوزه می‌توان به دسته‌بندی جملات بر اساس معیارهایی همانند احساسات [۱] و رویدادهای متفاوت [۲] به کمک الگوریتم‌های گوناگون اشاره کرد. با بررسی پژوهش‌های پیشین، در می‌یابیم که اکثر مطالعات به بررسی جملات یا اسناد بر اساس واژه پرداخته و تعداد اندکی از آنها، مبنای پردازش خود را بر اساس حروف قرار داده‌اند. رویکرد فوق در زبان‌هایی همچون زبان فارسی که دارای ساختاری غنی هستند، باعث ایجاد محدودیت می‌شود. محدودیت مذکور را می‌توان در دگرگونی مفهوم یک واژه با تغییر یا حذف حروف دانست (مانند واژگان «خار» و «خوار» یا «شست» و «شست»). به همین دلیل در هنگام پردازش جملات بر اساس واژه، چنانچه انتظار بهره‌مندی الگوریتم از توانایی پردازش تمام واژگان را داشته باشیم، باید

مجموعه داده بسیار عظیم و غنی‌ای را جمع‌آوری کرد که علاوه بر واژه‌های مورد استفاده توسط افراد یک گروه عام یا مشخص، واژه‌های خطادار را نیز شامل شود که گردآوری چنین مجموعه‌ای در عمل غیرممکن است. از دیگر راه‌حل‌ها می‌توان به پیش‌پردازش و پاکسازی داده‌ها به صورت دقیق یا پردازش جملات بر اساس حروف اشاره کرد که مساله یاد شده را تا حدود زیادی حل می‌کنند. شایان توجه است که پژوهش‌های موجود در حوزه تشخیص احساسات، از مجموعه داده‌ها با یک محوریت مشخص همانند دیدگاه‌های جمع‌آوری شده از وبگاه‌های فروش تلفن همراه استفاده می‌کنند. وجود محدودیت مذکور، منجر به کمبود داده‌های آموزشی و عملکرد ضعیف الگوریتم در پردازش دیدگاه‌هایی با موضوع متفاوت خواهد شد. در پژوهش حاضر، چارچوبی مبتنی بر معماری طراحی شده در مرجع [۳] (ELMo¹) پیشنهاد شده است که توانایی پردازش جملات بر اساس حروف را دارا است و به وسیله لایه‌های چهارگانه تعبیه حروف، پیچش یک بعدی، نگاشت لایه قبل و شبکه عصبی بازگشتی تجهیز شده است. همچنین، در صورت وجود تعداد داده‌های کافی برای آموزش مدل پیشنهادی، می‌توان از بردارهای هر واژه که بر اساس جایگاه، مفهوم آنها و نوع وظیفه (تشخیص احساسات) در سطح جمله استخراج شده‌اند، استفاده نمود و عملکرد بکارگیری آنها در مساله‌های مشابه را در برابر سایر روش‌های تعبیه واژگان همچون Word2Vec [۴]، BERT [۵] یا ELMo [۳] که بردار هر واژه در آنها بر اساس وظیفه مدل‌سازی زبانی کشف می‌شود، مقایسه کرد. شایان ذکر است که در پژوهش حاضر، از چارچوب پیشنهادی تحت عنوان Convolutional-based Bidirectional Recurrent Character-Level Mapping Network (CBRCMN) یاد می‌شود.

در ادامه مقاله به صورت زیر سازماندهی شده است. در بخش ۲، پژوهش‌های پیشین شرح داده شده است. سپس، در بخش ۳، مجموعه داده‌های مورد استفاده در این پژوهش ذکر شده و در بخش ۴، مولفه‌های چارچوب پیشنهادی توصیف شده است. بخش ۵ به معرفی چارچوب پیشنهادی پرداخته و نتایج آن در

¹ Embedding from Language Models

بخش ۶ ارزیابی شده است. در نهایت بخش ۷ به نتیجه گیری و پژوهش های آتی می پردازد.

۲. پژوهش های پیشین

از تحقیقات انجام شده در حوزه دسته بندی جملات، می توان به مرجع [۶] اشاره کرد که در آن، محققان ابتدا به پیش پردازش و حذف داده های خطا دار پرداخته، سپس از روش های N-Gram، POS (Part-of-Speech) Tagging و PerSent (تعیین قطبیت واژگان) جهت استخراج ویژگی در سطح واژگان و از الگوریتم ماشین بردار پشتیبان برای دسته بندی آنها، بهره برده اند. لازم به ذکر است، محققان برای بررسی رویکردهای انتخاب ویژگی^۱، از مجموعه دیدگاه های ثبت شده در وبگاه caffeinema (بین سال های ۲۰۱۴ و ۲۰۱۶) استفاده کرده اند. مجموعه مذکور شامل ۵۰۰ دیدگاه مثبت و ۵۰۰ دیدگاه منفی می باشد که بهترین عملکرد مربوط به رویکردی ترکیبی شامل Bi- Uni-gram و Gram و Tri-Gram با میزان دقت ۸۸/۳۶٪ است.

در پژوهش [۷]، به دسته بندی متون خبری (گردآوری شده از وبگاه همشهری) مبتنی بر حروف پرداخته شده است. مجموعه داده مورد استفاده، متشکل است از ۲۲۰۰۰ جمله در ۲۰ موضوع متفاوت (مانند ورزشی، سیاسی و اجتماعی) که ۲۰۰۰۰ نمونه از آن جهت آموزش و ۲۰۰۰ نمونه جهت آزمایش تعیین شده است. محققان در ابتدا مجموعه متون را با استفاده از روش One-Hot Encoding (در سطح حرف) به دنباله ای از بردارهای عددی تبدیل کرده، سپس با استفاده از دو لایه پیچش یک بعدی، دو لایه Max Pooling و یک لایه خروجی، به دسته بندی آنها پرداخته اند. همچنین در پژوهش مذکور، از الگوریتم های دیگری مانند Term Frequency (تبدیل جملات به مقادیر عددی)، Naïve Bayes و Support Vector Machine (به منظور کلاس بندی) نیز بهره برداری شده است که بهترین عملکرد مربوط به رویکرد شبکه عصبی با میزان دقت ۴۹/۵۹٪ می باشد. دو پژوهش بررسی شده، به پردازش جملات در سطح واژگان یا حروف می پردازند.

فارغ از رویکرد پردازشی موجود در مراجع [۶] و [۷] (پردازش در سطح واژگان)، محققان در مرجع [۸] به دسته بندی احساسی جملات بر اساس رویکرد واژگان-حروف پرداخته اند. مجموعه داده مورد استفاده در این پژوهش، شامل ۱۵۱۰۰۰ دیدگاه جمع آوری شده از وبگاه دیجی کالا می باشد که در ۵ کلاس متفاوت دسته بندی شده است. محققان پس از آماده سازی داده ها، واژگان را با استفاده از روش Word2Vec و حروف را با استفاده از One-Hot Encoding به بردارهای عددی تبدیل کرده، سپس با استفاده از الگوریتم های شبکه عصبی پیچشی و بازگشتی، آنها را دسته بندی نموده اند. بهترین عملکرد در سطح واژه، مربوط به ترکیب Word2Vec، لایه پیچشی و بازگشتی با میزان دقت ۹۵٪ (بر روی مجموعه داده Binary) بوده و بیشترین میزان دقت در سطح حرف توسط ترکیب لایه پیچشی و بازگشتی با میزان دقت ۹۴٪ حاصل شده است.

برای بهبود عملکرد چارچوب های پیشنهادی در پژوهش های مذکور که از مشکل کمبود تعداد داده های آموزشی رنج می برند، محققان در پژوهش [۹]، با استفاده از روش های وب کاوی موفق به گردآوری ۳ میلیون بازخورد ثبت شده در وبگاه دیجی کالا شده اند. هر دیدگاه توسط خریدار صاحب نظر، عنوان «پیشنهاد می کنم» یا «پیشنهاد نمی کنم» را دریافت کرده که از آن به عنوان برچسب مرتبط با دیدگاه بهره برداری شده است. در پژوهش ذکر شده با هدف مقایسه نتایج مدل های پیش بینی مختلف و تاثیر پیش پردازش در عملکرد آنها، علاوه بر الگوریتم شبکه های عصبی (معماری های پیچشی و بازگشتی)، بکارگیری سایر روش های نظارتی چون Naïve Bayes و Logistic Regression در دستور کار قرار گرفته است. در رویکرد شبکه عصبی از روش FastText و در سایر مدل های نظارتی از روش TF-IDF جهت تبدیل واژگان به بردارهای عددی استفاده شده است. از بین الگوریتم های مورد مطالعه، بهترین عملکرد مربوط به شبکه پیچشی با میانگین F-score برابر با ۹۵/۶٪ و Area Under Curve (AUC) برابر با ۹۹/۶٪ است. در ضمن در پژوهش ذکر شده، تعیین برچسب داده ها توسط روش Pseudo-Labeling صورت پذیرفته است.

^۱ Feature Selection

پردازش آنها توسط رمزنگار BERT استفاده شده است. مدل BERT بکارگیری شده توسط محققین، نسخه از پیش آموزش دیده متشکل از ۱۲ سر^۵ مجهز به مکانیزم خود-توجه^۶ و ۱۲ لایه رمزنگار می‌باشد که پس از محاسبه خروجی توسط آن، نمونه مورد نظر توسط یک لایه تمام-متصل^۷ به کلاس‌های مشخص، دسته‌بندی می‌شود. لازم به ذکر است پژوهشگران در این تحقیق، از پنج مجموعه داده که دارای تعداد نمونه و کلاس‌های متمایز (دسته‌بندی‌های احساسی و موضوعات خبری) بودند، استفاده کرده‌اند.

از دیگر پژوهش‌های موجود در حوزه تشخیص احساسات به زبان فارسی می‌توان به مرجع [۱۳] اشاره نمود. در تحقیق فوق، محققان با استفاده از معماری مبتنی بر شبکه‌های عصبی پیچشی و بازگشتی، به دسته‌بندی احساسی توئیت‌ها با محوریت نظر مردم در ارتباط با واکسن کووید-۱۹ و رویکرد آنها در مقابل واکسن‌های وارداتی و داخلی پرداخته‌اند. پژوهشگران به منظور تحقق هدف ذکر شده، حدود ۸۰۰۰۰۰ توئیت فارسی را از سطح وبگاه توئیت جمع‌آوری کرده، سپس با استفاده از مدل پیشنهادی آنها را به سه کلاس مثبت، منفی و خنثی دسته‌بندی نموده‌اند. علی‌رغم پژوهش‌های [۸]، [۱۰]، [۱۲] و [۱۳] که محققان از روش‌های مبتنی بر شبکه عصبی بهره برده‌اند، در مرجع [۱۴] چارچوبی مبتنی بر ماشین بردار پشتیبان جهت دسته‌بندی احساسی دیدگاه‌ها (با موضوع قیمت سهام) و تکنیک مبتنی بر لغت‌نامه با هدف تبدیل داده‌های متنی به بردارهای عددی متشکل از وزن مشخص برای هر واژه، معرفی نموده‌اند. لازم به ذکر است در پژوهش فوق، اطلاعات قیمت سهام که با دیدگاه کاربران مرتبط است، با هدف بهبود بیشتر در عملکرد چارچوب پیشنهادی (حدود ۸٪)، برای ایجاد لغت‌نامه مورد استفاده قرار گرفته است. در تحقیق مذکور، بهترین عملکرد، با استفاده از مقدار ۶۰۰۰ برای طول بردار ویژگی به میزان F-score برابر با ۵۷/۹٪ حاصل شده است.

در مرجع [۱۰]، محققان به طراحی سیستم تشخیص احساسات در نظرات وبگاه caffecinema با رویکرد یادگیری عمیق پرداخته‌اند. در این پژوهش، ۱۰۲۱ دیدگاه با برجسب مثبت و ۹۸۹ دیدگاه با برجسب منفی جمع‌آوری، پیش‌پردازش و آماده‌سازی شده و سپس با استفاده از شبکه‌های پیچشی، بازگشتی و رمزنگار خودکار^۱ دسته‌بندی شده‌اند. معماری Stacked-BiLSTM (شبکه LSTM دوطرفه متشکل از دو لایه) نسبت به بقیه مدل‌ها با میزان دقت ۹۵/۶۱٪ و F-score ۹۶٪، عملکرد بهتری را به همراه داشته است. همچنین در پژوهش مذکور، نسخه از پیش‌آموزش دیده الگوریتم FastText جهت تبدیل واژگان به بردار ۳۰۰ بُعدی، مورد استفاده قرار گرفته است.

از پژوهش‌های انجام شده توسط محققین آسیایی می‌توان به مرجع [۱۱] اشاره نمود که در آن یک مدل یادگیری گروهی چندسطحی^۲ با تمرکز بر وظیفه دسته‌بندی احساسی دیدگاه‌های گردآوری شده از شبکه اجتماعی توئیت، معرفی شده است. مدل فوق از سه رویکرد یادگیری متمایز، با نام‌های Bagging، Multinomial Naive Bayes و Boosting با هدف خنثی‌سازی مشکلات الگوریتم‌هایی همانند درخت تصمیم^۳ تشکیل شده است. همچنین، خروجی مدل پیشنهادی، مبتنی بر رای اکثریت الگوریتم‌های مورد استفاده می‌باشد. لازم به ذکر است محققان، مجموعه داده مورد نظر خود را دیدگاه‌هایی با موضوع رضایت از رئیس پلیس کشور اندونزی (در سه کلاس منفی، خنثی و مثبت) و روش بکار گرفته شده جهت تبدیل دیدگاه‌های ورودی به ماتریس‌های عددی، بردار ساز شمارش^۴ معرفی نموده‌اند.

در مرجع [۱۲]، محققان به معرفی مدلی بر مبنای تنظیم دقیق BERT با هدف دسته‌بندی پیام‌های طولانی (به زبان عربی) که از شبکه توئیت گردآوری شده‌اند، معرفی نموده‌اند. در پژوهش فوق، از سه سطح تعبیه با عناوین تعبیه در سطح نشانه، قسمت و جایگاه، با هدف دستیابی به بردار مناسب برای واژگان و

¹ Auto-Encoder

² Multi-Level Ensemble Learning

³ Decision Tree

⁴ Count Vectorizer

⁵ Head (Multi-Head Attention)

⁶ Self-Attention

⁷ Fully-Connected Layer

۳. مجموعه داده‌ها

در این بخش، مجموعه داده‌های مورد استفاده این پژوهش، شیوه پیش‌پردازش و چگونگی آماده‌سازی آنها، مورد بررسی قرار گرفته است.

۳.۱. جمع‌آوری مجموعه داده‌ها

در این پژوهش، ۶ مجموعه دیدگاه با موضوعات متفاوت که از وبگاه‌های اسنپ‌فود، اینستاگرام و دیجی کالا جمع‌آوری شده‌اند، وارد CBRCMN می‌شوند. سرویس‌های مذکور با هدف تسهیل ارائه خدمات اینترنتی راه‌اندازی شده‌اند که حوزه فعالیت هر یک عبارتند از:

- (۱) دیجی کالا بستری جهت خرید و فروش محصولات در دسته‌بندی‌های گوناگون،
- (۲) اسنپ‌فود سرویسی جهت عرضه مواد غذایی،
- (۳) اینستاگرام شبکه‌ای اجتماعی برای انتشار مطالبی با موضوعات متفاوت و تبادل نظر درباره آنها.

پس از جمع‌آوری مجموعه داده‌های شرح داده شده در جداول (۱) و (۲)، در مجموع ۱۹۰۷۲۹ دیدگاه به دست می‌آید که پس از پاکسازی و انتخاب دیدگاه‌های با تعداد واژگان بیشتر از ۲ و کمتر از ۱۰۰، مجموعه جدیدی با تعداد ۱۸۴۴۳۸ نمونه حاصل می‌شود. شکل (۱) چگونگی توزیع مجموعه داده نهایی را در قالب برچسب‌های سه‌گانه برچسب مثبت (۱)، برچسب خشی (۰) و برچسب منفی (-۱)، نشان می‌دهد. درخور ذکر است، از مجموعه نهایی ۱۰۰۰۰ نمونه به صورت تصادفی برای آزمایش و ۱۷۴۴۳۸ نمونه جهت آموزش انتخاب شده است. در ادامه، فرآیند پیش‌پردازش و چگونگی آماده‌سازی داده‌ها در بخش ۳،۲، به تفصیل بیان خواهد شد.

۳.۲. پیش‌پردازش و آماده‌سازی داده‌ها

برای پیش‌پردازش مجموعه دیدگاه‌ها، ابتدا فاصله‌های موجود در بین واژگان که بیش از یک بار تکرار شده‌اند را حذف کرده، سپس لغات غیرفارسی و علائم نگارشی حذف می‌شوند. پس از

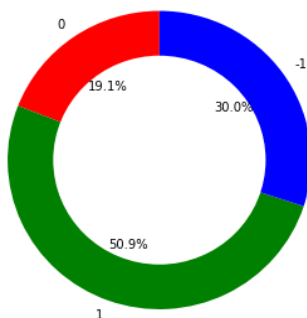
اجرای مراحل مذکور، نرمال‌سازی و ریشه‌یابی واژگان با هدف کاهش تعداد واژگان متمایز، در دستور کار قرار گرفته و جملات با تعداد لغات بیشتر از ۲ و کمتر از ۱۰۰ انتخاب می‌شوند. مقصود از نرمال‌سازی، تبدیل تمام فاصله‌ها به نیم‌فاصله‌ها (مانند فعل «می‌ماند» به «می‌ماند» و مقصود از ریشه‌یابی، جایگزینی هر فعل با ریشه خود می‌باشد. ذکر این نکته لازم است که به منظور آماده‌سازی داده‌ها، به هر یک از حروف موجود در واژه، عددی ویژه اختصاص داده می‌شود تا لغات و جملات به بردار و ماتریس‌های عددی تبدیل شوند. این مراحل در شکل (۲) نشان داده شده است.

جدول (۱): تعداد نمونه‌های موجود در مجموعه داده‌ها

شماره	مجموعه داده	مثبت	خشی	منفی
۱	Miras Opinion [15]	۴۹,۵۱۵	۲۹,۴۷۱	۱۴,۸۸۲
۲	Snapfood [16]	۳۵,۰۰۰	۰	۳۵,۰۰۲
۳	SentiPers [17], [18]	۷,۸۵۶	۴,۰۰۰	۲,۱۹۰
۴	Instagram Revs	۲,۷۸۰	۲,۴۹۵	۳,۲۳۷
۵	Digikala Revs1 [19]	۲,۳۸۲	۴۱۹	۴۶۰
۶	Digikala Revs2	۵۰۰	۲۵۰	۲۹۰
	جمع	۹۸,۰۳۳	۳۶,۶۳۵	۵۶,۰۶۱

جدول (۲): گزاره‌های هندسی مرتبط با مجموعه داده‌های شش‌گانه

مقدار	گزاره در مجموعه داده‌های ۱ تا ۶
۱	حداقل تعداد نشانه (Token) پیش از پاکسازی
۱۵۳۵	حداکثر تعداد نشانه پیش از پاکسازی
۳۱	میانگین تعداد نشانه پیش از پاکسازی
۲,۷۰۴,۳۹۲	تعداد کل واژگان پس از پاکسازی
۱۲,۵۷۴,۸۱۶	تعداد کل حروف پس از پاکسازی



شکل (۱): چگونگی توزیع دیدگاه‌ها مبتنی بر برچسب‌های سه‌گانه در

مجموعه داده‌های ۱ تا ۶

۵۰) اختصاص می‌دهد. لایه نگاشت^۲ نیز مانند تعبیه عمل کرده، منتها پس از ضرب در ماتریس وزن، با مقدار بایاس جمع شده و تابع فعالسازی^۳ مشخصی بر روی آن اعمال می‌شود.

$$e = (X * W) \quad (1)$$

در رابطه (۱)، مقصود از X ماتریس حروف، W ماتریس وزن، e ماتریس تعبیه و $*$ عملگر ضرب داخلی می‌باشد.

۴.۲. پیچش یک بعدی

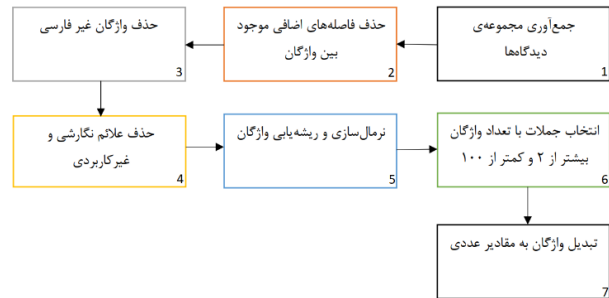
بر اساس فرآیند پیچش (رابطه (۲))، ویژگی‌های موجود در هر واژه با اعمال بردارهای فیلتر بر روی حروف آن، استخراج و به لایه‌های بعدی موجود در شبکه عصبی، به منظور تصمیم‌گیری منتقل می‌شوند [۲۱].

$$c[i] = f \left(\sum (X[i:i+w-1]^o K) + b \right) \quad (2)$$

در رابطه (۲)، X بردار ورودی (محاسبه شده توسط لایه تعبیه)، b مقدار بایاس، K بردار فیلتر، w اندازه هسته^۴، o عملگر ضرب عنصر به عنصر، f تابع فعال‌سازی و c بردار خروجی می‌باشد. در CBRCMN، سه گروه فیلتر با تعداد (۳۲، ۳۲، ۶۴) و با مقدار اندازه هسته‌های (۱، ۲، ۳) جهت استخراج ویژگی هر واژه بر اساس پردازش تک حرفی، دو حرفی و سه حرفی به کار رفته است. شایان ذکر است با اجرای عملیات تعبیه، پیچش و انتخاب بردار با بزرگترین مقادیر^۵، می‌توان برای هر واژه برداری با تعداد ابعاد ۱۲۸ به دست آورد.

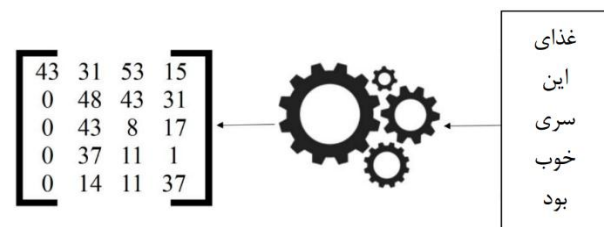
۴.۳. حافظه بلند-کوتاه مدت

معماری حافظه بلند-کوتاه مدت^۵ (رابطه‌های (۳) تا (۸)) [۲۲]، نوع خاصی از شبکه عصبی بازگشتی است که با هدف رفع مشکل گرادیان محوشدنی^۶ [۲۳]، در هنگام پردازش داده‌های ترتیبی طراحی شده است. مشکل مذکور بر اثر محاسبه مشتق زنجیره‌ای از شبکه بازگشتی با تعداد گام‌های زمانی طولانی



شکل (۲): مراحل پیش‌پردازش و آماده‌سازی داده‌ها

به عنوان مثال، در شکل (۳) به یک نمونه از دیدگاه‌های موجود در مجموعه نظرات اسنپ‌فود، یک ماتریس عددی اختصاص داده شده است. در این راستا، کتابخانه‌های Keras، Hazm و re مورد استفاده قرار گرفته‌اند.



شکل (۳): مثالی از تخصیص ماتریس اعداد به یک دیدگاه

در بخش ۴، به جزئیات چارچوب پیشنهادی و مولفه‌های مربوط به آن خواهیم پرداخت.

۴. مولفه‌های چارچوب پیشنهادی

در بخش حاضر، روابط حاکم بر مولفه‌های موجود در چارچوب پیشنهادی مورد مطالعه قرار گرفته است. سپس، معرفی و بررسی نمای کلی چارچوب پیشنهادی در بخش ۵ صورت می‌پذیرد.

۴.۱. تعبیه

اولین مولفه از چارچوب پیشنهادی که تحت عنوان لایه تعبیه^۱ (مشابه مرجع [۲۰]) شناخته می‌شود، وظیفه انتقال حروف به مختصات برداری را بر عهده دارد. لایه ذکر شده با انجام ضرب داخلی بین ماتریس ورودی و وزن‌ها (رابطه (۱))، به هر یک از حروف، برداری مشخص با طول ثابت (در مساله حاضر مقدار

¹ Embedding

² Mapping Layer

³ Activation Function

⁴ Global Max Pooling

⁵ Long-Short Term Memory

⁶ Vanishing Gradient

ذکر است، معماری مطرح شده در برخی مساله‌ها همچون مدل‌سازی داده‌های صوتی [۲۷]، نسبت به LSTM عملکرد بهتری را از خود به نمایش گذاشته است.

$$r_t = \delta(W_r[c_{t-1}, X_t] + b_{r1} + b_{r2}) \quad (9)$$

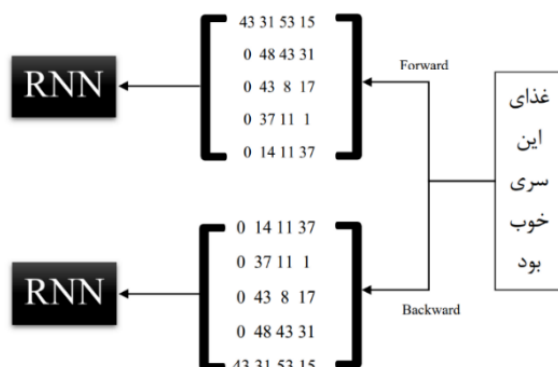
$$z_t = \delta(W_z[c_{t-1}, X_t] + b_{z1} + b_{z2}) \quad (10)$$

$$\tilde{c}_t = \tanh(W_{c1}X_t + b_{c1} + r_t \circ (W_{c2}c_{t-1} + b_{c2})) \quad (11)$$

$$c_t = z_t \circ c_{t-1} + (1 - z_t) \circ \tilde{c}_t \quad (12)$$

در روابط (۹) تا (۱۲)، مقصود از b بردار بایاس، W ماتریس وزن، X ماتریس ورودی، \circ عملگر ضرب عنصر به عنصر، c_{t-1} وضعیت پنهان در لحظه $t-1$ ، r_t دروازه بازنشانی^۲، z_t دروازه بروزسانی^۳، \tilde{c}_t دروازه جدید، δ تابع فعال‌سازی Sigmoid و \tanh تابع فعال‌سازی Hyperbolic Tangent می‌باشد.

شبکه‌های بازگشتی از قابلیت تحت عنوان اتصال دو طرفه نیز می‌توانند بهره‌مند شوند که امکان پردازش دو طرفه داده‌ها را برای آنها فراهم می‌کند. شایان توجه است که اتصال ذکر شده در برخی مساله‌ها همچون پردازش متن مورد استفاده قرار می‌گیرد. روابط حاکم بر معماری‌های LSTM و GRU در ساختار مذکور نیز وجود دارند، ولی هنگام محاسبات، مقادیر از دو طرف مورد پردازش قرار می‌گیرند. برای مثال، در بحث تشخیص احساس موجود در جمله، لغات از اول به آخر و بالعکس، به شبکه وارد می‌شوند. نمایی از نحوه پردازش جمله توسط شبکه بازگشتی دو طرفه، در شکل (۴) قابل ملاحظه است.



شکل (۴): نمایی از شبکه عصبی بازگشتی دوطرفه

(مانند جملات با تعداد واژگان زیاد) و سپس میل گرادیان به مقدار صفر، پدیدار می‌شود. همچنین دلیل رفع مشکل ذکر شده در معماری LSTM، ساختار دروازه‌ای آن و امکان کنترل جریان اطلاعات در پردازش داده‌های متوالی است. در ضمن از LSTM، در حل مساله‌های دیگر همانند Named Entity Recognition [۲۴] و تشخیص گفتار [۲۵]، نیز استفاده می‌شود.

$$i_t = \delta(W_i[h_{t-1}, X_t] + b_{i1} + b_{i2}) \quad (3)$$

$$f_t = \delta(W_f[h_{t-1}, X_t] + b_{f1} + b_{f2}) \quad (4)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, X_t] + b_{c1} + b_{c2}) \quad (5)$$

$$o_t = \delta(W_o[h_{t-1}, X_t] + b_{o1} + b_{o2}) \quad (6)$$

$$c_t = (f_t \circ c_{t-1} + i_t \circ \tilde{c}_t) \quad (7)$$

$$h_t = o_t \circ \tanh(c_t) \quad (8)$$

در روابط بالا، مقصود از b بردار بایاس، W ماتریس وزن، X ماتریس ورودی، \circ عملگر ضرب عنصر به عنصر، h_{t-1} وضعیت پنهان در لحظه $t-1$ ، δ تابع فعال‌سازی Sigmoid و \tanh تابع فعال‌سازی Hyperbolic Tangent می‌باشد.

در معماری ذکر شده، ورودی در لحظه t به صورت مستقل طبق روابط (۳) (Input Gate)، (۴) (Forget Gate)، (۵) (Cell State) و (۶) (Output Gate) با ماتریس‌های وزن ضرب داخلی و با مقدار بایاس جمع شده، سپس نتایج بر اساس رابطه (۷) یعنی (Cell State) مورد پردازش قرار گرفته و خروجی در لحظه t طبق رابطه (۸) حاصل می‌شود.

۴.۴. واحد دروازه‌ای بازگشتی

معماری LSTM علی‌رغم رفع مشکل گرادیان محوشدنی، با چالش تعداد پارامترهای زیاد همراه است. وجود ایراد ذکر شده، باعث ایجاد محدودیت در حافظه برای ذخیره‌سازی معماری‌های عظیم و مدت زمان آموزش آنها خواهد شد. به منظور رفع دو مشکل ذکر شده، معماری دیگری تحت عنوان واحد بازگشتی دروازه‌ای^۱ (رابطه‌های (۹) تا (۱۲)) [۲۶] ابداع شده است. شایان

^۲ Reset Gate

^۳ Update Gate

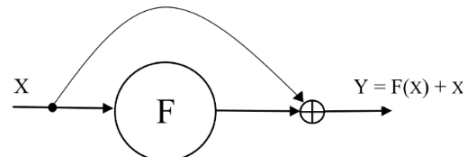
^۱ Gated Recurrent Unit

۴.۵. اتصال باقی مانده

داده شده است. در روش CBRCMN، ابتدا دیدگاه مورد نظر پیش پردازش و آماده سازی شده و مطابق پیمانہ Embed_Conv (شکل (۷)) به واژه موجود در جمله، یک بردار ۱۲۸ بُعدی (مستقل از مفهوم) اختصاص داده می شود. در پیمانہ مذکور، ابتدا حروف تشکیل دهنده واژه توسط لایه تعبیه به یک فضای ۵۰ بُعدی منتقل شده، سپس ۳ دسته فیلتر با تعداد و اندازه های هسته مشخص بر روی آنها (استخراج ویژگی های موجود هر واژه بر اساس ۳ نوع نگرش متفاوت به حروف) اعمال می شود. پس از اجرای عملیات مذکور، بزرگترین مقدار موجود در ویژگی های استخراج شده برای هر لغت بر اساس تعداد فیلتر انتخاب می شود. به عنوان مثال، پس از اعمال ۳۲ گروه فیلتر بر روی حروف یک واژه، بزرگترین مقدار هر گروه انتخاب شده و برداری به طول ۳۲ بُعد برای واژه مورد نظر حاصل می شود. پس از کسب بردار برای هر واژه، ماتریس جدیدی با ابعاد ذکر شده در شکل (۷) تولید می شود. لازم به ذکر است که در نگارش دستورات موجود در شکل (۷)، از قواعد زبان برنامه نویسی پایتون استفاده شده است. پس از انجام مراحل مذکور و تولید ماتریس واژگان، بردار لغات طی فرآیند نگاشت به فضای دیگری با ابعاد ۱۲۸ منتقل و تابع فعال سازی ReLU بر روی آنها اعمال شده و سپس توسط شبکه عصبی بازگشتی (LSTM یا GRU) پردازش می شود. پس از کشف ارتباطات موجود بین واژگان و استخراج بردار مبتنی بر مفهوم توسط شبکه بازگشتی، نتایج پردازش به لایه بعد با تعداد ۳ گره منتقل شده و احساس مربوط به دیدگاه ورودی تعیین می شود. همچنین، در صورت وجود داده های کافی و بهره مندی الگوریتم از قدرت تعمیم بالا، می توان از بردارهای محاسبه شده توسط لایه های نگاشت، به عنوان بردارهای از پیش آموزش دیده (بر اساس وظیفه تشخیص احساسات) در مساله های دیگر استفاده کرد.

طبق توضیحات مطرح شده در بخش ۴.۵، دلیل بهره برداری از اتصال باقی مانده در چارچوب پیشنهادی، جلوگیری از پدیده گرادیان محوشدنی در لایه های ابتدایی است. در دو شکل (۸) و (۹) می توان تاثیر اتصال مذکور را در تغییرات گرادیان لایه ها مشاهده کرد.

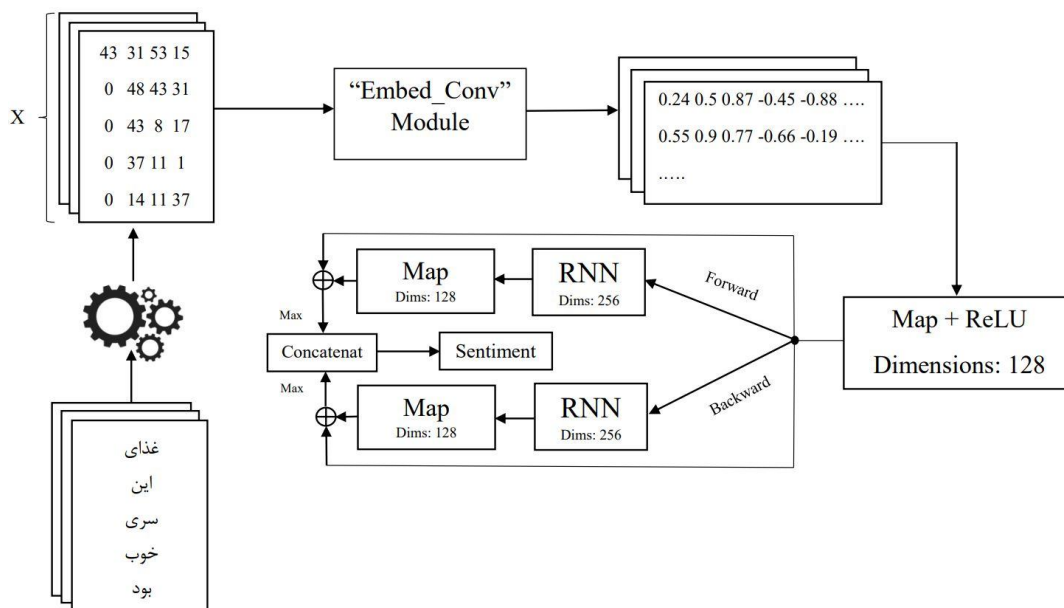
با توجه افزایش تعداد پارامترهای موجود در مساله ها، ممکن است شبکه های عصبی پیش خور (Feed-Forward) با تعداد محدودی از لایه های پنهان، امکان مدل سازی آنها را نداشته باشند. همچنین، تمام داده ها از ماهیت دنباله ای برخوردار نیستند و نمی توان از معماری بازگشتی برای آنها استفاده کرد. در نتیجه، برای پردازش چنین مساله هایی به معماری پیش خور نیاز است. با این حال، در این معماری با افزایش تعداد لایه های شبکه، مشکل گرادیان محوشدنی برای لایه های ابتدایی ظاهر می شود. به عبارت دیگر، مقدار حاصل از محاسبه مشتق زنجیره ای تابع خطا نسبت به پارامتر لایه های اولیه، به صفر میل کرده و پارامترهای الگوریتم به درستی بروزرسانی نمی شوند. برای حل این مشکل، می توان از اتصالی تحت عنوان «اتصال باقی مانده» (نام دیگر Skip Connection یا اتصال پرشی) بهره برد. روش نامبرده، با فراهم سازی یک مسیر برای جریان مشتق در جهت لایه های اولیه، از محو شدن مقدار گرادیان (میل به صفر) جلوگیری می کند. در مساله حاضر، از اتصال ذکر شده به منظور بروزرسانی پارامترهای موجود در لایه های اولیه شبکه، استفاده شده است. لازم به ذکر است که مقایسه نتایج حاصل از بکارگیری و عدم بکارگیری چنین اتصالی، در بخش ۵ مورد بررسی قرار گرفته است. همچنین، دیاگرام و معادله اتصال ذکر شده، در شکل (۵) قابل مشاهده است. در این شکل، X مقدار ورودی، F تابع فعال سازی و Y خروجی شبکه پس از جمع با ورودی F می باشد.



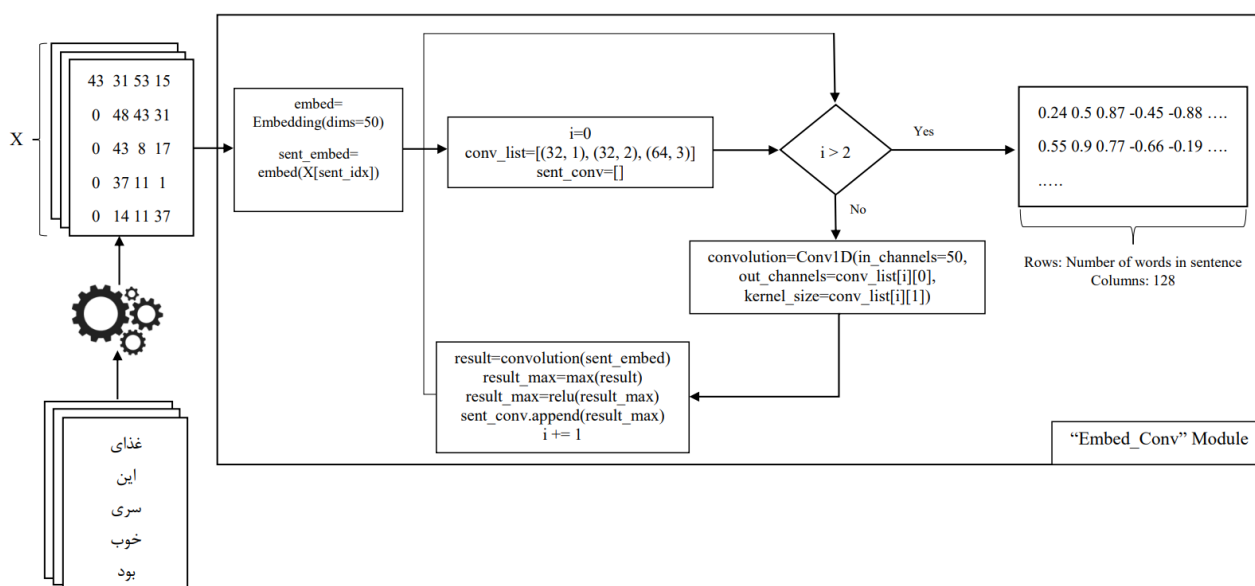
شکل (۵): دیاگرام و معادله اتصال باقی مانده

۵. معرفی چارچوب پیشنهادی

در پژوهش حاضر، دسته بندی جملات بر اساس پردازش حروف توسط لایه تعبیه، پیش از یک بعدی، شبکه بازگشتی دو طرفه و نگاشت انجام می شود. دیاگرام CBRCMN در شکل (۶) نشان



شکل (۶): دیاگرام معماری پیشنهادی (CBRCMN)



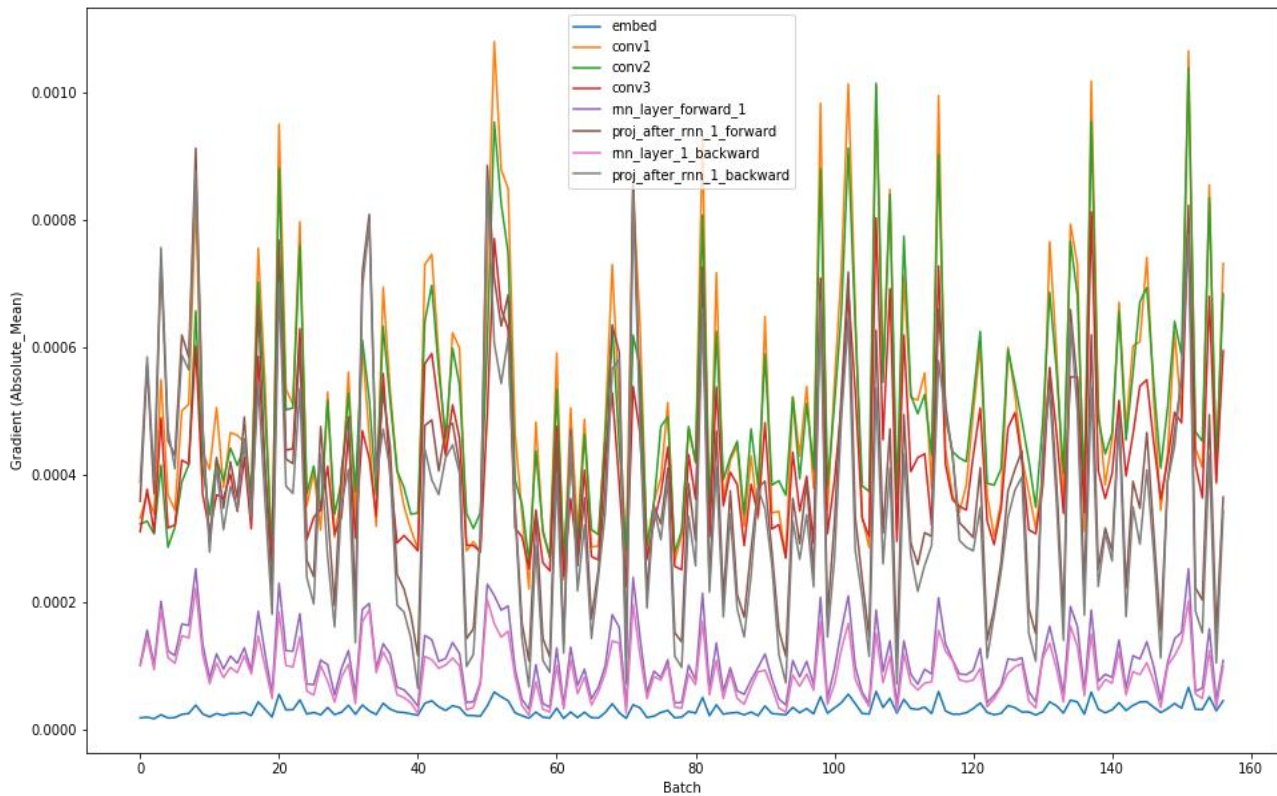
شکل (۷): پیمانه‌ی EMBED_CONV

طرز چشمگیری تغییر پیدا کرده‌اند و نیز تغییرات لایه embed (رنگ آبی) با مقداری جزئی همراه است. در بخش بعد، نتایج مدل پیشنهادی و مثال عددی از شیوه عملکرد آن، بیان می‌شود.

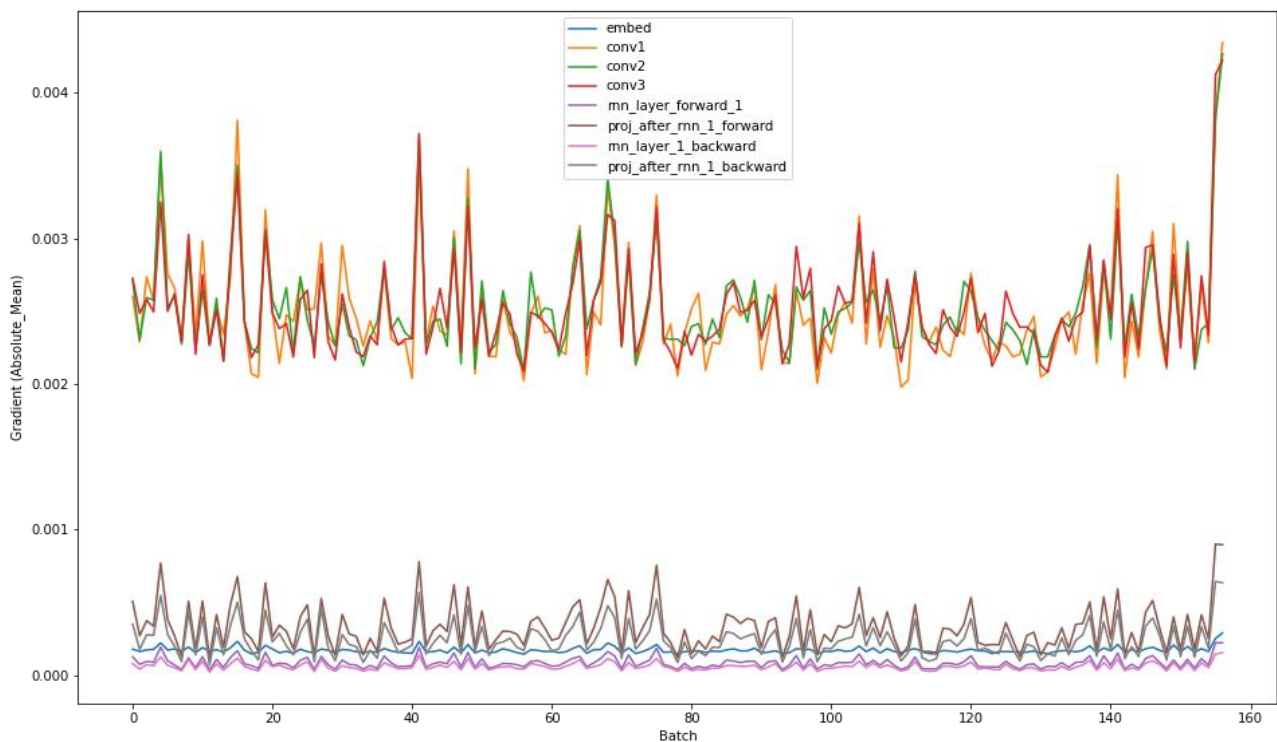
۶. نتایج پژوهش

در این بخش، ابتدا نتایج مدل پیشنهادی و مقایسه عملکرد آن با مدل‌های مبتنی بر واژه و سپس نمونه‌ای عددی از عملکرد روش CBRCMN مورد بررسی قرار خواهد گرفت.

نمودارهای رسم شده در دو شکل (۸) و (۹)، قدر مطلق میانگین تغییرات گرادیان در هر لحظه از آموزش (تقریباً به تعداد ۱۵۶ دسته آموزشی) را برای وزندهی به لایه‌های چارچوب، نمایش می‌دهند. بر اساس نتایج موجود در این شکل‌ها، در صورت عدم بهره‌گیری از اتصال باقی‌مانده، گرادیان مربوط به لایه‌های اولیه (رنگ‌های نارنجی، سبز و قرمز) ناچیز بوده و شبکه به درستی آموزش داده نمی‌شود. درحالی‌که با بکارگیری اتصال ذکر شده، مقادیر مشتق مربوط به لایه‌های conv1، conv2 و conv3، به



شکل (۸): نمودار تغییرات گرادیان بدون استفاده از اتصال



شکل (۹): نمودار تغییرات گرادیان هنگام استفاده از اتصال

۶.۱. نتایج و مقایسه با مدل‌های مبتنی بر واژه

در پژوهش حاضر، با هدف پیاده‌سازی و آموزش CBRCMN، ابزار PyTorch و الگوریتم بهینه‌سازی Adam [۲۸] مورد استفاده قرار گرفته است. جزئیات آموزش و سخت‌افزار بکارگیری شده، در جدول (۳) قابل مشاهده می‌باشد.

جدول (۳): جزئیات آموزش چارچوب

مقدار	مولفه
Adam	Optimization Algorithm
1e-4	Learning Rate
0.9	Beta-1
0.999	Beta-2
1e-8	Eps
10	Epochs
32	Batch Size
Python v3.7	Programming Language
PyTorch	Deep Learning Library
Intel® Xeon® 2.2 GHZ	CPU
Tesla T4 (15 GB)	GPU
12.68 GB	RAM

همچنین، با هدف کسب جامعیت در نتایج کسب شده، مدل‌های مبتنی بر واژه [۱۴] (شکل (۱۰)) و ساختارهای متفاوت (گره‌های موجود در لایه‌های نگاشت و بازگشتی) در چارچوب پیشنهادی، بکارگیری شده‌اند که نتایج هر یک در جدول (۴) قابل ملاحظه می‌باشد. در این جدول، ساختارهای چارچوب با ترکیب (map_dim, rnn_dim) معرفی شده‌اند که مولفه اول، معرف تعداد گره‌های موجود در لایه‌های نگاشت و مولفه دوم بیانگر تعداد گره‌های موجود در لایه‌های بازگشتی می‌باشد.

چارچوب مورد نظر طبق جزئیات موجود در جدول (۳)، روی مجموعه دیدگاه‌های شرح داده شده (بخش ۳.۱)، آموزش یافته و جزئیات آموزش و عملکرد رویکردهای مورد نظر، در جدول (۴) و شکل‌های (۱۱)، (۱۲) و (۱۳) قابل مشاهده است.

مطابق جدول (۴)، استفاده از معماری GRU با ساختار دو طرفه (مرسوم به CBRCMN_BiGRU_Medium و برخورداری هر شبکه از ۲۵۶ گره) همراه با لایه نگاشت ۱۲۸ بُعدی، میزان دقت

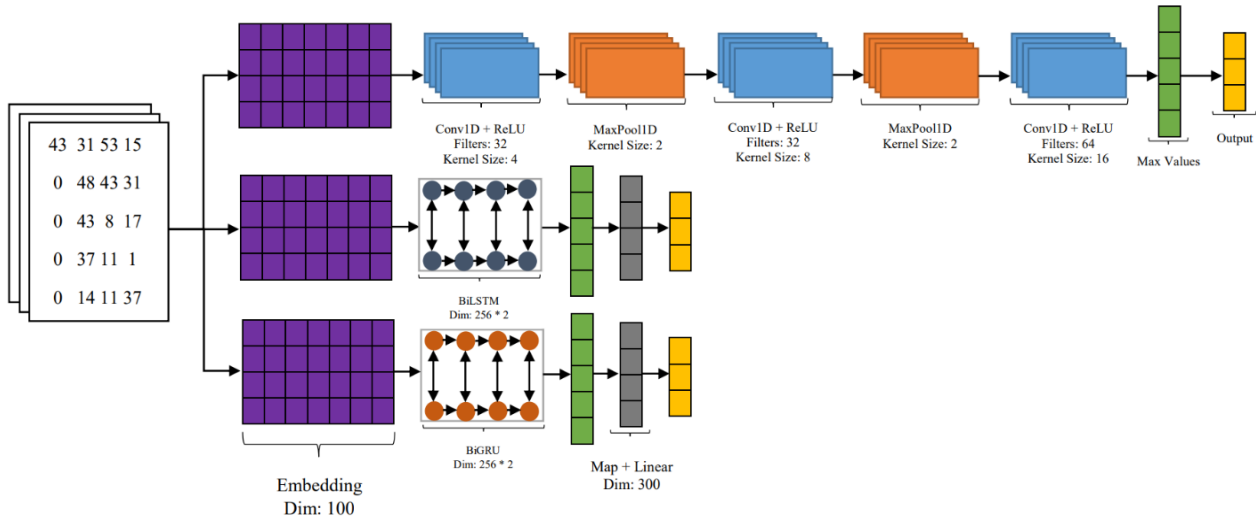
۷۹/۸۷٪ و F-score برابر با ۷۹/۹٪ بر روی مجموعه داده‌های آزمایشی را با خود به همراه داشته است. در پژوهش حاضر، از معماری‌های معرفی شده در بخش ۴ جهت ساخت مدل‌های مبتنی بر واژه نیز استفاده شده است که بهترین عملکرد مربوط به معماری CNN (سه لایه پیچش، دو لایه MaxPoolId، یک لایه جهت انتخاب بیشترین مقادیر و یک لایه خروجی) با میزان دقت ۷۶/۵۶٪ و F-score برابر با ۷۶/۷۴٪ می‌باشد. معماری رویکردهای مبتنی بر واژه، در شکل (۱۰) مشهود است.

جدول (۴): نتایج رویکردها و معماری‌ها بر روی داده‌های آزمایشی

Model	Base	Accuracy	F-score	Scale
CBRCMN (64, 128)	BiLSTM	78.68%	78.76%	Small
CBRCMN (64, 128)	BiGRU	79.11%	78.93%	Small
CBRCMN (128, 256)	BiLSTM	79.4%	79.43%	Medium
CBRCMN (128, 256)	BiGRU	79.87%	79.9%	Medium
CBRCMN (256, 512)	BiLSTM	78.95%	79.1%	Large
CBRCMN (256, 512)	BiGRU	78.5%	78.52%	Large
Word-Based	BiLSTM	74.97%	74.98%	Medium
Word-Based	BiGRU	74.15%	74.17%	Medium
Word-Based	CNN	76.56%	76.74%	Medium

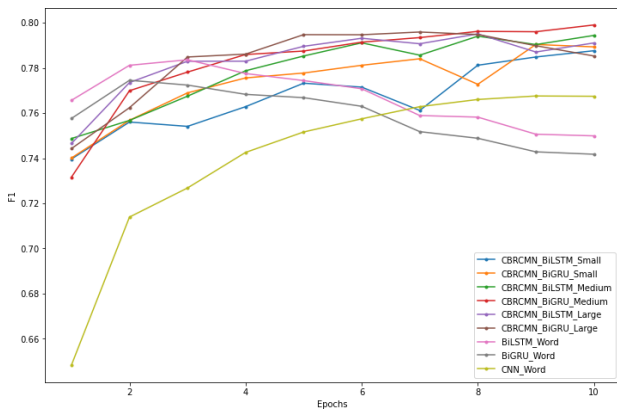
مدل‌های مبتنی بر واژه، بر مبنای دو رویکرد پردازشی، طرح‌ریزی شده‌اند. در رویکرد اول با بکارگیری عملگر پیچش و بردارهای فیلتر، ویژگی‌های سطح بالا از بردارهای واژگان موجود در جملات، استخراج شده و پس از تعیین بزرگترین مقادیر، با استفاده از یک لایه شبکه عصبی به سه گروه احساسی دسته‌بندی می‌شوند.

در رویکرد دوم، با استفاده از دو معماری بازگشتی ارتباط بین واژگان تعیین و پس از کسب بزرگترین مقادیر، به کلاس‌های مورد نظر دسته‌بندی می‌شوند. لازم به ذکر است که لایه تعبیه، در سه رویکرد فوق مشابه است. همچنین، عملکرد رویکردهای بکارگیری شده، در شکل‌های (۱۱)، (۱۲) و (۱۳) قابل مشاهده می‌باشد.



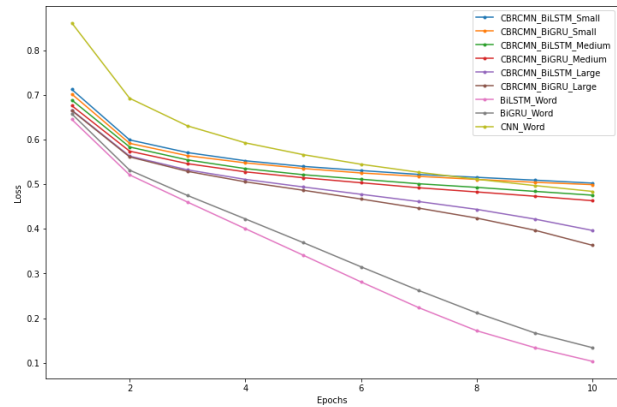
شکل (۱۰): معماری رویکردهای مبتنی بر واژه

واژه پس از ۱۰ دوره آموزش، به کمترین میزان خطا در داده‌های آموزشی دست یافته‌اند.

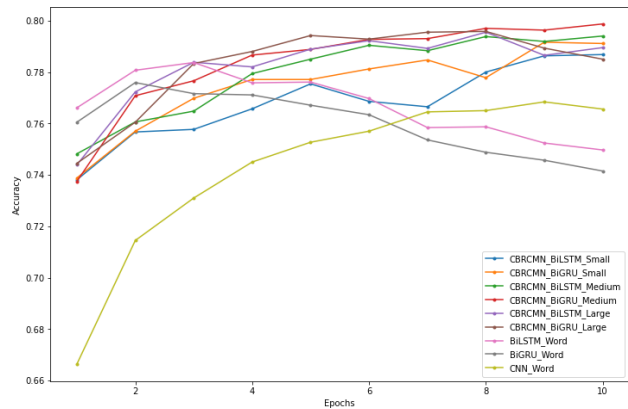


شکل (۱۳): مقدار F-SCORE (نمونه‌های آزمایش)

شایان توجه است که اکتساب کمترین میزان خطا در داده‌های آموزشی، بیانگر ارجحیت معماری در بحث عملکرد بر روی داده‌های آزمایشی نمی‌باشد، زیرا با توجه به شکل‌های (۱۲) و (۱۳)، مدل‌های مبتنی بر حروف بهترین میزان تعمیم‌دهی و عملکرد را با خود به همراه داشته‌اند. به عبارت دیگر، معماری Medium CBRCMN به بیشترین میزان دقت (رابطه (۱۳)) در داده‌های آزمایشی و BiGRU_Word به کمترین میزان عملکرد دست یافته است. همچنین در اثر عدم وجود توازن در داده‌های بکارگیری شده، از معیار F-score (رابطه (۱۴)) با هدف بررسی دقیق عملکرد مدل‌ها، استفاده شده است و نتایج معیار مذکور در شکل (۱۳) قابل ملاحظه می‌باشد.



شکل (۱۱): میزان خطا (نمونه‌های آموزش)



شکل (۱۲): مقدار دقت (نمونه‌های آزمایش)

مطابق شکل (۱۱)، علی‌رغم پیاده‌سازی‌های روش CBRCMN، مدل‌های مبتنی بر واژه که به شبکه بازگشتی تجهیز شده‌اند (BiLSTM_Word, BiGRU_Word) از سرعت بیشتری در امر آموزش بهره‌مند می‌باشند. به عبارت دیگر، مدل‌های مبتنی بر

حروف تعیین می‌شود. به عبارت دیگر، با بکارگیری پیمانه تحت اشاره، ویژگی‌ها و وابستگی‌های موجود بین حروف کشف می‌شود. پس از استخراج وابستگی‌ها، بزرگترین مقادیر در سطح کانال تعیین شده و بردارهای ۳۲، ۳۲ و ۶۴ بعدی برای هر واژه تعیین می‌شود. لازم به ذکر است که با هدف پیشگیری از ایراد گرادیان محوشدنی، تابع فعال‌سازی ReLU روی مقادیر اعمال شده و بردارها به یکدیگر الحاق می‌شوند (کسب بردار ویژگی ۱۲۸ بعدی). همچنین با هدف تعیین تعداد ابعاد مشخص و استفاده از اتصال باقی‌مانده، از یک لایه نگاشت با تابع ReLU بکارگیری شده است. گفتنی است شیوه کارکرد پیمانه بیان شده در برخورد با نظرات دیگر، مشابه شکل (۱۴) می‌باشد.

دلیل بهبود عملکرد فرآیند تشخیص احساسات توسط چارچوب پیشنهادی را می‌توان در استفاده از شبکه‌های بازگشتی و پیمانه Embed_Conv جستجو نمود. علی‌رغم مدل‌های مبتنی بر واژه، در CBRCMN پس از چهار گام محاسباتی (عملگرهای تعبیه و پیچش)، برای هر واژه بردار عددی تعیین می‌شود. به عبارت دیگر، در Embed_Conv، عمل پردازش ورودی و استخراج ویژگی در سطح حروف صورت پذیرفته است که منجر به کاهش تعداد پارامترها (جدول (۵))، کشف وابستگی‌های موجود بین حروف و بهره‌مندی از توانایی پردازش تمام واژگان می‌شود. با هدف دستیابی به کدهای پیاده‌سازی، اسناد مربوطه در مخزن گیت‌هاب^۲ قرار داده شده است. در ضمن با هدف کسب تعداد پارامترهای هر مدل، کتابخانه torchinfo بکارگیری شده است.

جدول (۵): تعداد پارامتر معماری‌های نامبرده

تعداد پارامترها	مدل
241,889	CBRCMN_BiLSTM_Small
192,225	CBRCMN_BiGRU_Small
891,681	CBRCMN_BiLSTM_Medium
694,049	CBRCMN_BiGRU_Medium
3,469,217	CBRCMN_BiLSTM_Large
2,680,737	CBRCMN_BiGRU_Large
9,897,187	BiLSTM_Word
9,713,891	BiGRU_Word
9,063,283	CNN_Word

^۲ https://github.com/omidkhalafbeigi/CBRCMN_WordBased

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

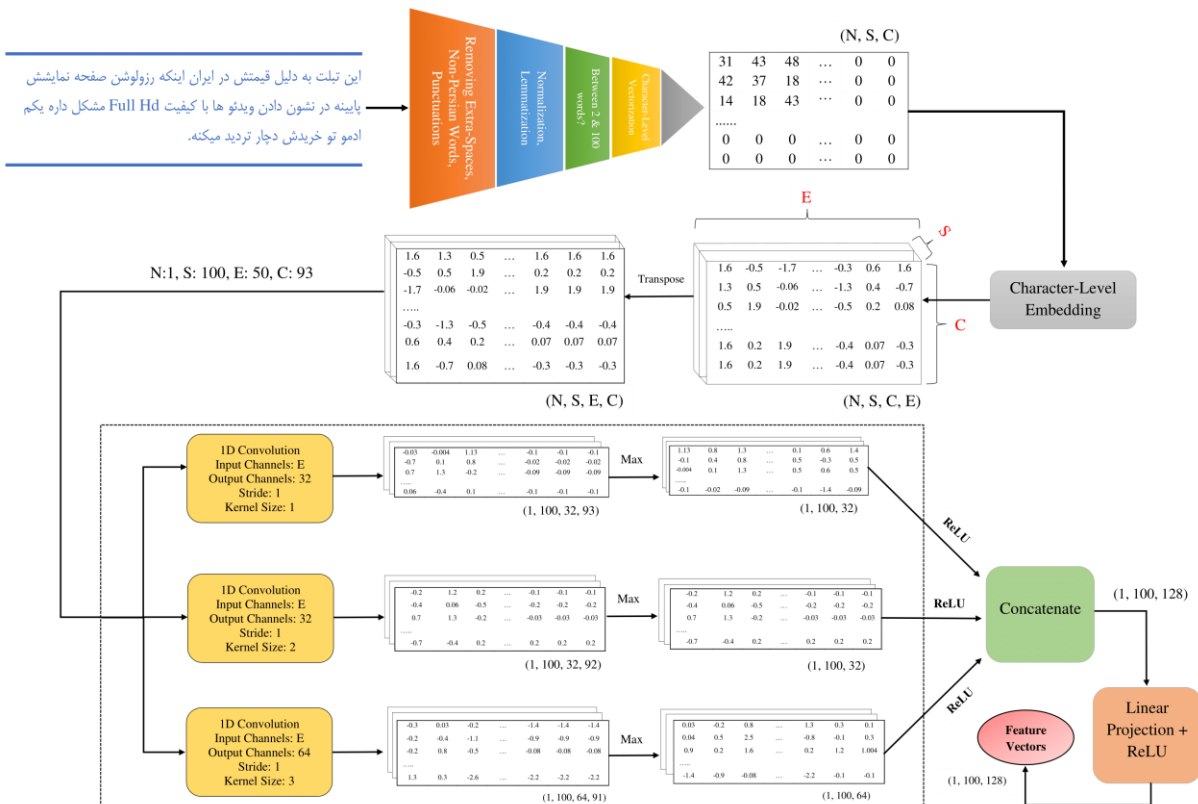
$$F-score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (14)$$

در روابط (۱۳) و (۱۴)، مقصود از TP همان True Positive همان TN همان True Negative، FP همان False Positive و FN همان False Negative می‌باشد. همچنین، با توجه به نتایج مندرج در شکل‌های (۱۲) و (۱۳) مشاهده می‌شود در دومین و سومین دوره آموزش، BiLSTM_Word و BiGRU_Word به بیشترین میزان دقت و F-score خود رسیده‌اند و پس از طی ۱۰ دوره آموزشی به سوی بیش برآزش^۱ میل می‌کنند. لازم به ذکر است که با هدف درک بهتر از عملکرد مولفه‌های استفاده شده در CBRCMN، مطالعه پژوهش‌های [۲۹]-[۳۱] به خوانندگان توصیه می‌شود.

۶.۲. نمونه عددی از پیمانه Embed_Conv

به جهت ایجاد شفافیت در شیوه عملکرد پیمانه Embed_Conv، یک مثال کاربردی از استخراج ویژگی‌های ۱۲۸ بعدی موجود در هر واژه (بر پایه پردازش حروف)، در شکل (۱۴) شرح داده شده است. طبق فرآیند فوق ابتدا نسخه پیش‌پردازش و پاکسازی شده جمله، حاصل شده و سپس به تمام حروف تشکیل دهنده واژگان، مقادیر صحیح و مثبت تخصیص داده شده و ماتریسی از واژگان و حروف ایجاد می‌شود. در ماتریس مذکور، هر سطر و ستون بیانگر واژگان و حروف آنها می‌باشند. پس از تعیین ماتریس اعداد و افزودن مقادیر صفر به بردارها با هدف ایجاد طول ثابت (عمل Padding)، عمل تعبیه در سطح حروف روی ماتریس اعمال شده و ماتریس چهار بعدی حاصل می‌شود. بعد اول از ماتریس مذکور بیانگر جملات (N)، بعد دوم واژگان (S)، بعد سوم حروف (C) و بعد چهارم تعداد ویژگی‌های استخراج شده از حروف (E) می‌باشد. پس از تعیین ماتریس ذکر شده و انتقال سطر و ستون‌ها (ترانهاده)، سه سطح لایه پیچش با اندازه هسته‌های ۱، ۲ و ۳ با تعداد کانال ۳۲، ۳۲ و ۶۴ بر روی آن اعمال شده و وابستگی‌های محلی موجود در بین

^۱ Over-fitting



شکل (۱۴): نمونه‌ی عددی از کارکرد پیمانه EMBED_CONV

پرداختیم. مدل پیشنهادی از یک لایه تعبیه واژگان، سه دسته پیچش یک بُعدی مستقل، سه لایه نگاشت و یک لایه بازگشتی دو طرفه (تجهیز شده به اتصال پرشی) تشکیل شده است که بیشترین میزان دقت (۷۹/۸۷٪) با تعداد گره‌های ۱۲۸ و ۲۵۶ (متعلق به لایه‌های نگاشت و بازگشتی)، حاصل می‌شود. لازم به ذکر است، در رویکرد مبتنی بر واژه نیز بهترین عملکرد مربوط به معماری شبکه پیچشی یک بُعدی با میزان دقت ۷۶/۵۶٪ می‌باشد. معماری مذکور از سه لایه پیچش یک بُعدی (تعداد فیلترهای ۳۲، ۳۲ و ۶۴) و دو لایه MaxPool1D تشکیل شده است. در مطالعات آینده قصد داریم از چارچوب معرفی شده، برای پیشبرد مسائلی مانند خلاصه‌سازی و ترجمه متون استفاده و عملکرد آنها را بررسی نماییم.

تعارض منافع: نویسندگان اعلام می‌کنند که هیچ تعارض منافعی ندارند.

با توجه به جدول (۵)، ملاحظه می‌شود پیاده‌سازی‌های معماری CBRCMN در مقایسه با مدل‌های مبتنی بر واژه، دارای تعداد پارامترهای کمتری است. نکته اشاره شده، منجر به کاهش حجم مورد نیاز برای ذخیره‌سازی مدل برای مسائل بزرگ‌تر، می‌شود.

۷. نتیجه‌گیری و کارهای آتی

در این پژوهش به بررسی مطالعات انجام شده در زمینه تشخیص احساسات که زیرمجموعه‌ای از علم پردازش زبان‌های طبیعی می‌باشد، پرداخته شده است. برای این منظور، در ابتدا پایگاه داده‌ای حاوی مجموعه داده‌های شش‌گانه، برگرفته از سرویس‌های دیجی کالا، اسنپ‌فود و اینستاگرام گردآوری شده و با اعمال روش‌های پیش‌پردازش، زمینه لازم برای ورود آنها به چارچوب یادگیری ماشین پیشنهادی فراهم شده است. پس از شرح مجموعه داده‌ها، به معرفی و بررسی اجزای سازنده روش پیشنهادی (CBRCMN)، چالش‌ها و راه‌حل‌های موجود، معرفی مدل اصلی و عملکرد آن بر روی مجموعه داده‌های شش‌گانه

- [1] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," in Proc. Conf. Empir. Methods Nat. Lang. Process., (EMNLP), Philadelphia, PA, USA, 2002, pp. 79-86, doi: 10.3115/1118693.1118704.
- [2] A. Rao and N. Spasojevic, "Actionable and Political Text Classification Using Word Embeddings and Lstm," arXiv, 2016, doi: 10.48550/arXiv.1607.02501.
- [3] M.E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep Contextualized Word Representations," in Proc. Conf. North American Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol., (NAACL-HLT), New Orleans, Louisiana, USA, 2018, vol. 1, pp. 2227-2237, doi: 10.18653/v1/N18-1202.
- [4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in 1st Int. Conf. Learn. Represent. (ICLR), Scottsdale, Arizona, USA, 2013.
- [5] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in Proc. Conf. North American Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol., (NAACL-HLT), Minneapolis, MN, USA, 2019, vol. 1, pp. 4171-4186, doi: 10.18653/v1/N19-1423.
- [6] K. Dashtipour, M. Gogate, A. Adeel, A. Hussain, A. Alqarafi, and T. Durrani, "A comparative study of persian sentiment analysis based on different feature combinations," in Int. Conf. Commun. Signal Process. Syst. (CSPS), Harbin, China, 2017, pp. 2288-2294, doi: 10.1007/978-981-10-6571-2_279.
- [7] S. Ghasemi and A.H. Jadidinejad, "Persian text classification via character-level convolutional neural networks," in 8th Conf. AI Robot. 10th RoboCup Iranopen Int. Symp. (IRANOPEN), Qazvin, Iran, 2018, pp. 1-6, doi: 10.1109/RIOS.2018.8406623.
- [8] S. Zobeidi, M. Naderan, and S.E. Alavi, "Opinion mining in Persian language using a hybrid feature extraction approach based on convolutional neural network", *Multim. Tools Appl.*, vol. 78, no. 22, pp. 32357-32378, 2019, doi: 10.1007/s11042-019-07993-4.
- [9] S. Shumaly, M. Yazdinejad, and Y. Guo, "Persian sentiment analysis of an online store independent of pre-processing using convolutional neural network with fastText embeddings," *PeerJ Comput. Sci.*, vol. 7, p. e422, 2021, doi: 10.7717/peerj-cs.422.
- [10] K. Dashtipour, M. Gogate, A. Adeel, H. Larjani, and A. Hussain, "Sentiment Analysis of Persian Movie Reviews Using Deep Learning," *Entropy*, vol. 23, no. 5, p. 596, 2021, doi: 10.3390/e23050596.
- [11] W.F. Satrya, R. Aprilliyani, and E.H. Yossy, "Sentiment analysis of Indonesian police chief using multi-level ensemble model," *Proc. Comput. Sci.*, vol. 216, pp. 620-629, 2023, doi: 10.1016/j.procs.2022.12.177.
- [12] M. Alruily, A.M. Fazal, A.M. Mostafa, and M. Ezz, "Automated Arabic Long-Tweet Classification Using Transfer Learning with BERT," *Appl. Sci.*, vol. 13, no. 6, p. 3482, 2023, doi: 10.3390/app13063482.
- [13] Z.B. Nezhad and M.A. Deihimi, "Twitter sentiment analysis from Iran about COVID 19 vaccine," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 16, no. 1, p. 102367, 2022, doi: 10.1016/j.dsx.2021.102367.
- [14] M. Ahangari and A. Sebt, "A Hybrid Approach to Sentiment Analysis of Iranian Stock Market User's Opinions," *Int. J. Eng.*, vol. 36, no. 3, pp. 573-584, 2023, doi: 10.5829/ije.2023.36.03c.18.
- [15] [15] Ashrafi Asli, S. A., Sabeti, B., Majdabadi, Z., Golazizian, P., Fahmi, R., Momenzadeh, O., "Optimizing Annotation Effort Using Active Learning Strategies: A Sentiment Analysis Case Study in Persian", in Proc. Twelfth Lang. Resour. Eval. Conf., Marseille, France, 2020, pp. 2855-2861.
- [16] M. Farahani, M. Gharachorloo, M. Farahani, and M. Manthouri, "ParsBERT: Transformer-based Model for Persian Language Understanding," *Neural Process Lett.*, vol. 53, no. 6, pp. 3831-3847, 2021, doi: 10.1007/s11063-021-10528-4.
- [17] P. Hosseini, A.A. Ramaki, H. Maleki, M. Anvari, and S.A. Mirroshandel, "SentiPers: a sentiment analysis corpus for Persian," arXiv, 2018, doi: 10.48550/arXiv.1801.07737.
- [18] J.P.R. Sharami, P.A. Sarabestani, and S.A. Mirroshandel, "Deepsentipers: Novel deep learning models trained over proposed augmented Persian sentiment corpus," arXiv, 2020, doi: 10.48550/arXiv.2004.05328.
- [19] M. Samizadeh, (15 Sep. 2022), Sentiment-Analysis-with-LSTM-in-Persian: v 1.0.0, [Online]. Available: <https://zenodo.org/records/6862064>.
- [20] S. Takase and S. Kobayashi, "All word embeddings from one embedding," *Ann. Conf. Neural Inf. Process. Syst.*, NeurIPS, virtual, 2020, vol. 33, pp. 3775-3785.
- [21] Y. Kim, Y. Jernite, D.A. Sontag, and A.M. Rush, "Character-aware neural language models," in Proc. 30th AAAI Conf. Artif. Intell., Phoenix, Arizona,

- USA, 2016, pp. 2741-2749, doi: 10.1609/aaai.v30i1.10362.
- [22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [23] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in 30th Int. Conf. Mach. Learn. (ICML), Atlanta, GA, USA, 2013, pp. 1310-1318.
- [24] J.P.C. Chiu and E. Nichols, "Named entity recognition with bidirectional LSTM-CNNs," *Trans. Assoc. Comput. Linguistics*, vol. 4, pp. 357-370, 2016, doi: 10.1162/tacl_a_00104.
- [25] A. Graves, A.R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, Vancouver, BC, Canada, 2013, pp. 6645-6649, doi: 10.1109/ICASSP.2013.6638947.
- [26] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," in *Proc. 8th Worksh. Syntax Semant. Struct. Stat. Transl.*, Doha, Qatar, 2014, pp. 103-111, doi: 10.3115/v1/W14-4012.
- [27] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *NIPS Worksh. Deep Learn.*, 2014.
- [28] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015.
- [29] M. Feizi-Derakhshi, Z. Mottaghinia, and M. Asgari-Chenaghlu, "Persian Text Classification Based on Deep Neural Networks," *Soft Comput. J.*, vol. 11, no. 1, pp. 120-139, 2022, doi: 10.22052/scj.2023.243182.1010 [In Persian].
- [30] F. Zare Mehrjardi, M. Yazdian-Dehkordi, and A. Latif, "Evaluating classical machine learning and deep-learning methods in sentiment analysis of Persian telegram message," *Soft Comput. J.*, vol. 11, no. 1, pp. 88-105, 2022, doi: 10.22052/scj.2023.246553.1077 [In Persian].
- [31] A. Khosravi and H. Abdolhosseini, "Personality in social networks using thematic modelling of user feedback," *Soft Comput. J.*, vol. 11, no. 2, pp. 50-61, 2023, doi: 10.22052/scj.2023.243197.1006 [In Persian].