

بکارگیری مدل تحلیل احساسات در سطح حروف مبتنی بر شبکه‌ی 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

^۱ گروه مهندسی برق و کامپیوتر - دانشگاه خوارزمی - تهران - ایران

^۲ مرکز آموزش عالی فنی و مهندسی بوئین زهرا - گروه مهندسی برق و کامپیوتر - بوئین زهرا - قزوین - ایران

abashirimosavi@bzte.ac.ir

^۳ مرکز آموزش عالی فنی و مهندسی بوئین زهرا - گروه مهندسی برق و کامپیوتر - بوئین زهرا - قزوین - ایران

چکیده: امروزه با توجه به تمایل روز افزون مردم برای خرید اجناس از طریق فروشگاه‌های اینترنتی و شبکه‌های مجازی، شاهد افزایش داده‌های بدون ساختار مانند متن در سطح اینترنت هستیم. از این رو، پردازش متون و توسعه‌ی الگوریتم‌های کارآمد جهت استخراج دانش، توجه پژوهشگران حوزه‌ی علوم داده در بسترهای مذکور را به خود جلب کرده است. از رویکردهای پردازش متن می‌توان به موضوع 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 از مدل‌های زبانی استفاده شده است که لایه‌های چهارگانه‌ی تعبیه (جهت انتقال حروف به فضای برداری)، پیچش یک بُعدی (جهت استخراج بردار ویژگی برای هر واژه)، نگاشت و شبکه‌ی عصبی بازگشتی را شامل می‌شود. در چارچوب پیشنهادی، ابتدا با بکارگیری لایه‌ی تعبیه در سطح حروف، برداری ثابت برای آن‌ها تعیین شده است. سپس، مبتنی بر عملگرهای پیچش یک بُعدی که به صورت موازی بکارگیری شده‌اند، ارتباط معنایی و منطقی بین حروف تشکیل‌دهنده‌ی هر واژه بدست آمده و بردار ۱۲۸ بُعدی 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 مجموعه‌ی نظرات مبتنی بر سنجه‌های Accuracy و F-score، به ترتیب ۷۹٫۸۷٪ و ۷۹٫۹٪ می‌باشد.

واژه‌های کلیدی: پردازش زبان طبیعی، تشخیص احساسات، مدل مبتنی بر محتوا، شبکه‌ی عصبی عمیق، پلتفرم‌های اینترنتی.

* سید علیرضا بشیری موسوی (نویسنده‌ی مسئول)، abashirimosavi@bzte.ac.ir

Applying Character-Level Neural Network-Based Sentiment Analysis Model on Persian Comments of the Social Media-Online Store Platforms

Omid Khalaf Beigi¹, Master's Degree Student, Seyed Alireza Bashiri Mosavi^{2*}, Assistant Professor, Sina Gharloghi³, Bachelor Student

¹ Department of Electrical and Computer Engineering, Kharazmi University, Tehran, Iran.

² Department of Electrical and Computer Engineering, Buein Zahra Technical University, Buein Zahra, Qazvin, Iran, abashirimosavi@bzte.ac.ir.

³ Department of Electrical and Computer Engineering, Buein Zahra Technical University, Buein Zahra, Qazvin, Iran.

Abstract: Nowadays, due to people being more willing to shop online through online stores and social media, we are facing the growth of unstructured data like texts on the internet. Hence, text processing and the development of optimal algorithms for extracting knowledge have drawn scholar's attention to this field. One of the aspects of the text processing field is classifying texts in the form of classes of various sentiments using different algorithms. We propose a novel framework to classify the comments based on the user's sentiment performed in the character-level scenario. Hence, the proposed framework is mounted on the architecture of embedding from the language model triggered by the quad-layer, namely embedding, one-dimensional convolution, the map, and the recurrent neural network. In the proposed framework, first, by using the embedding layer at the level of the character, a constant vector is assigned to them. Next, the semantic and logical relation between the characters per word for surviving word-specific 128-dimensional vectors is extracted by exerting the parallel-oriented one-dimensional convolution operators. After obtaining vectors, based on two recurrent neural network architectures, the relationship between the discovered words and the comment-specific sentiment is determined. The obtained results show that the proposed framework has an Accuracy of 79.87% and a F-score of 79.9% for comments class labeling.

Keywords: Natural Language Processing, Sentiment Analysis, Context Based Model, Deep Neural Network, Internet Platforms.

* Seyed Alireza Bashiri Mosavi, abashirimosavi@bzte.ac.ir

۱. مقدمه

پس از همه‌گیری ویروس کرونا، تمام ابعاد زندگی انسان‌ها تحت تاثیر قرار گرفته است. از مهمترین جنبه‌های شایان ذکر، می‌توان به فعالیت‌های تجاری و کسب و کارهای خرد اشاره کرد که با پیشرفت تکنولوژی، گستره‌ی وسیعی از وظایف نیروی انسانی، تقلیل یافته و به صورت از راه دور و خودکار انجام می‌شوند. همچنین، شیوه‌ی فروش و ارائه‌ی خدمات از رویکرد سنتی فاصله گرفته و مبتنی بر تکنولوژی‌های دیجیتال، به سیر تحولی خود ادامه می‌دهد. برای مثال می‌توان به تجارت‌های دیجیتال محور از قبیل وبسایت‌های شخصی و فروشگاه‌های همانند دیجی کالا و شبکه‌های اجتماعی همانند اینستاگرام اشاره کرد. با توسعه‌ی بسترهای مذکور، زمینه‌ی لازم جهت تبادل نظر کاربران و تولید حجم انبوهی از داده‌های بدون ساختار فراهم شده است که طیف وسیعی از اطلاعات همچون میزان فروش محصولات، بازخورد مشتریان و تبادل نظر کاربران در مورد واقعه‌های سیاسی و مذهبی را در بر می‌گیرد.

همزمان با فعالیت کاربران در اینترنت به زبان‌های پرکاربرد دنیا همچون انگلیسی، فرانسوی، آلمانی و عربی، میزان استفاده‌ی فارسی زبانان مانند ایرانیان، تبعه‌های افغانی و پاکستانی از این محیط نیز افزایش پیدا کرده است که نتیجه‌ی 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 دیگر سو، تعداد پژوهش‌های صورت گرفته در زمینه‌ی 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 میان رشته‌ای بوده که به مطالعه‌ی توانایی ماشین در پردازش، درک زبان انسان‌ها و در

صورت نیاز، تعامل با آن‌ها می‌پردازد. از زیرشاخه‌های این حوزه می‌توان به دسته‌بندی جملات براساس 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 براساس واژه پرداخته‌اند و تعداد اندکی از آن‌ها، مبنای پردازش خود را براساس حروف قرار داده‌اند. رویکرد فوق در زبان‌هایی همچون زبان فارسی که از ساختاری غنی برخوردار 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 شصت). به همین دلیل در هنگام پردازش جملات براساس واژه، چنانچه انتظار بهره‌مندی الگوریتم از توانایی پردازش تمام واژگان را داشته باشیم، باید مجموعه داده‌ی بسیار عظیم و غنی‌ای را جمع‌آوری کرد که علاوه بر واژه‌های مورد استفاده توسط افراد یک گروه عام یا مشخص، واژه‌های خطا دار را نیز شامل شود که گردآوری چنین مجموعه‌ای عملاً غیرممکن است. از دیگر راه حل‌ها می‌توان به پیش پردازش و پاکسازی داده‌ها به صورت دقیق یا پردازش جملات براساس حروف اشاره کرد که مسئله‌ی یاد شده را تا حدود زیادی حل می‌کنند. شایان توجه است که پژوهش‌های موجود در حوزه‌ی تشخیص احساسات، از مجموعه‌ی داده‌ها با یک محوریت مشخص همانند دیدگاه‌های جمع‌آوری شده از وبسایت‌های 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 توسط [۳] (ELMo) پیشنهاد شده است که توانایی پردازش جملات براساس حروف را داراست و توسط لایه‌های چهارگانه

رویکردی ترکیبی شامل Uni-gram، Bi-Gram و Tri-Gram با میزان دقت ۸۸,۳۶٪ است.

در پژوهش [۷] به دسته‌بندی متون خبری (گردآوری شده از وبسایت همشهری) مبتنی بر حروف پرداخته شده است. مجموعه داده‌ی 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 ابتدا مجموعه‌ی متون را با استفاده از روش One-Hot Encoding (در سطح حرف) به دنباله‌ای از بردارهای عددی تبدیل کرده، سپس با استفاده از دو لایه پیچش یک بعدی، دو لایه Max Pooling و یک لایه خروجی، به دسته‌بندی آن‌ها پرداخته‌اند. همچنین در پژوهش مذکور، از الگوریتم‌های دیگری مانند Term Frequency (تبدیل جملات به مقادیر عددی)، Naïve Bayes و Support Vector Machine (جهت کلاس‌بندی) نیز بهره‌برداری شده است که بهترین عملکرد مربوط به رویکرد شبکه‌ی عصبی با میزان دقت ۴۹,۵۹٪ می‌باشد. دو پژوهش بررسی شده، به پردازش جملات در سطح واژگان یا حروف می‌پردازند.

فارغ از رویکرد پردازشی موجود در [۶] و [۷] (پردازش در سطح واژگان)، محققان در [۸] به دسته‌بندی احساسی جملات براساس رویکرد واژگان-حروف پرداخته‌اند. مجموعه داده‌ی مورد استفاده در این پژوهش، شامل ۱۵۱۰۰۰ دیدگاه جمع‌آوری شده از وبسایت دیجی کالا می‌باشد که در ۵ کلاس متفاوت دسته‌بندی شده است. محققان پس از آماده‌سازی داده‌ها، واژگان را با استفاده از تکنیک Word2Vec و حروف را با استفاده از One-Hot Encoding به بردارهای عددی تبدیل کرده، سپس با استفاده از الگوریتم‌های شبکه‌ی عصبی پیچشی و بازگشتی، آن‌ها را دسته‌بندی نموده‌اند. بهترین عملکرد در سطح واژه، مربوط به ترکیب Word2Vec، لایه‌ی پیچشی و بازگشتی با میزان دقت

تعبیه‌ی حروف، پیچش یک بعدی، نگاشت لایه‌ی قبل و شبکه‌ی عصبی بازگشتی تجهیز شده است. همچنین، در صورت وجود تعداد داده‌های کافی برای آموزش مدل پیشنهادی، می‌توان از بردارهای هر واژه که براساس جایگاه، مفهوم آن‌ها و نوع وظیفه (تشخیص احساسات) در سطح جمله استخراج شده‌اند، استفاده نمود و عملکرد بکارگیری آن‌ها در مسائل مشابه را در برابر سایر تکنیک‌های تعبیه‌ی واژگان همچون Word2Vec [۴]، BERT [۵] یا ELMo که بردار هر واژه در آن‌ها براساس وظیفه‌ی مدل‌سازی زبانی کشف می‌شود، مقایسه کرد. شایان ذکر است در پژوهش حاضر، از چارچوب پیشنهادی تحت عنوان Convolutional-based Bidirectional Recurrent Character-Level Mapping Network (CBRCMN) یاد می‌شود.

ادامه‌ی مقاله به شرح زیر سازماندهی شده است:

۲. پژوهش‌های پیشین، ۳. مجموعه‌ی داده‌ها، ۴. مولفه‌های چارچوب پیشنهادی، ۵. معرفی چارچوب پیشنهادی، ۶. نتایج پژوهش و ۷. نتیجه‌گیری و پژوهش‌های آتی.

۲. پژوهش‌های پیشین

از تحقیقات انجام شده در حوزه‌ی دسته‌بندی جملات، می‌توان به [۶] اشاره کرد. در این پژوهش، محققان ابتدا به پیش‌پردازش و حذف داده‌های خطا دار پرداخته، سپس از روش‌های N-Gram، Part-of-Speech (POS) Tagging و PerSent (تعیین قطبیت واژگان) جهت استخراج ویژگی در سطح واژگان و از الگوریتم ماشین بردار پشتیبان برای دسته‌بندی آن‌ها، بهره برده‌اند. لازم به ذکر است، پژوهشگران جهت بررسی رویکردهای انتخاب ویژگی^۱، از مجموعه دیدگاه‌های ثبت شده در وبسایت caffeinema (بین سال‌های ۲۰۱۴ و ۲۰۱۶) استفاده کرده‌اند. مجموعه‌ی مذکور شامل ۵۰۰ دیدگاه مثبت و ۵۰۰ دیدگاه منفی می‌باشد که بهترین عملکرد مربوط به

۹۵٪ (بر روی مجموعه داده‌ی Binary) بوده و بیشترین میزان دقت در سطح حرف توسط ترکیب لایه‌ی پیچشی و بازگشتی با میزان دقت ۹۴٪ حاصل شده است.

برای بهبود عملکرد چارچوب‌های پیشنهادی در پژوهش‌های مذکور که از مشکل کمبود تعداد داده‌های آموزشی رنج می‌برند، محققان در پژوهش [۹] با استفاده از تکنیک‌های وب‌کاوی موفق به گردآوری ۳ میلیون بازخورد ثبت شده در وب‌سایت دیجی‌کالا شده‌اند. هر دیدگاه توسط خریدار صاحب نظر، عنوان "پیشنهاد می‌کنم" یا "پیشنهاد نمی‌کنم" را دریافت کرده که از آن به عنوان برچسب مرتبط با دیدگاه بهره‌برداری شده است. در پژوهش ذکر شده با هدف مقایسه‌ی نتایج مدل‌های پیش‌بینی مختلف و تاثیر پیش‌پردازش در عملکرد آن‌ها، علاوه بر الگوریتم شبکه‌های عصبی (معماری‌های پیچشی و بازگشتی)، به‌کارگیری سایر روش‌های نظارتی همچون Naïve Bayes و Logistic Regression در دستور کار قرار گرفته است. در رویکرد شبکه‌ی عصبی از تکنیک FastText و در سایر مدل‌های نظارتی از تکنیک TF-IDF جهت تبدیل واژگان به بردارهای عددی استفاده شده است. از بین الگوریتم‌های مورد مطالعه، بهترین عملکرد مربوط به شبکه‌ی پیچشی با میانگین F-score ۹۵٫۶٪ و Area Under Curve (AUC) ۹۹٫۶٪ است. لازم به ذکر است در پژوهش مذکور، تعیین برچسب داده‌ها توسط روش Pseudo-Labeling صورت پذیرفته است.

در [۱۰]، پژوهشگران به طراحی سیستم تشخیص احساسات در نظرات وب‌سایت caffecinema با رویکرد یادگیری عمیق پرداخته‌اند. در این پژوهش، ۱۰۲۱ دیدگاه با برچسب مثبت و ۹۸۹ دیدگاه با برچسب منفی جمع‌آوری، پیش‌پردازش و آماده‌سازی شده، سپس با استفاده از شبکه‌های پیچشی، بازگشتی

و رمزنگار خودکار^۱ دسته‌بندی شده‌اند. معماری Stacked-BiLSTM (شبکه‌ی LSTM دو طرفه متشکل از دو لایه) نسبت به بقیه‌ی مدل‌ها با میزان دقت ۹۵٫۶۱٪ و F-score ۹۶٪، عملکرد بهتری را به همراه داشته است. همچنین در پژوهش مذکور، نسخه‌ی از پیش‌آموزش دیده‌ی الگوریتم FastText جهت تبدیل واژگان به بردار ۳۰۰ بُعدی، مورد استفاده قرار گرفته است.

از پژوهش‌های انجام شده توسط محققین آسیایی می‌توان به [۱۱] اشاره نمود. در تحقیق مذکور، یک مدل یادگیری گروهی چند سطحی^۲ با تمرکز بر وظیفه‌ی^۳ دسته‌بندی احساسی دیدگاه‌های گردآوری شده از شبکه‌ی اجتماعی توئیتر^۴، معرفی شده است. مدل فوق از سه رویکرد یادگیری متمایز، به نام‌های Boosting، Multinomial Naïve Bayes و Bagging هدف خنثی‌سازی مشکلات الگوریتم‌هایی مانند درخت تصمیم^۵ تشکیل شده است. همچنین، خروجی مدل پیشنهادی، مبتنی بر رای اکثریت^۶ الگوریتم‌های مورد استفاده می‌باشد. لازم به ذکر است، پژوهشگران، مجموعه داده‌ی مورد نظر خود را از دیدگاه‌هایی با موضوع رضایت از رئیس پلیس کشور اندونزی (در سه کلاس منفی، خنثی و مثبت) و تکنیک بکارگرفته شده جهت تبدیل دیدگاه‌های ورودی به ماتریس‌های عددی، بردار ساز شمارش^۷ معرفی نموده‌اند.

در [۱۲] محققان به معرفی مدلی بر مبنای تنظیم دقیق^۸ BERT با هدف دسته‌بندی پیام‌های طولانی (به زبان عربی) که از شبکه‌ی توئیتر 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

3- Task

5- Decision Tree

7- Count Vectorizer

2- Multi-Level Ensemble Learning

4- Twitter

6- Majority Voting

8- Fine-Tuning

پردازش آنها توسط رمزنگار BERT استفاده شده است. مدل

BERT

پذیرفته شده در مجله محاسبات نرم

1- Auto-Encoder

3- Task

5- Decision Tree

7- Count Vectorizer

2- Multi-Level Ensemble Learning

4- Twitter

6- Majority Voting

8- Fine-Tuning

استفاده از مقدار ۶۰۰۰ برای طول بردار ویژگی به میزان F-score ۵۷,۹٪ حاصل شده است.

۳. مجموعه‌ی داده‌ها

در این بخش، مجموعه داده‌های مورد استفاده، شیوه‌ی پیش پردازش و چگونگی آماده‌سازی آن‌ها، مورد بررسی قرار گرفته است.

۱,۳. جمع آوری مجموعه‌ی داده‌ها

در این پژوهش، ۶ مجموعه‌ی دیدگاه با موضوعات متفاوت که از وبسایت‌های اسنپ‌فود^۴ و اینستاگرام^۵ و دیجی‌کالا^۶ جمع‌آوری شده‌اند، وارد CBRCMN می‌شوند. سرویس‌های مذکور با هدف تسهیل ارائه‌ی خدمات اینترنتی راه‌اندازی شده‌اند که حوزه‌ی فعالیت هر یک عبارت است از: (۱)؛ دیجی‌کالا بستری جهت خرید و فروش محصولات در دسته‌بندی‌های گوناگون، (۲)؛ اسنپ‌فود سرویسی جهت عرضه‌ی مواد غذایی و (۳)؛ اینستاگرام شبکه‌ای اجتماعی برای انتشار مطالبی با موضوعات متفاوت و تبادل نظر درباره‌ی آن‌ها.

جدول (۱) - تعداد نمونه‌های موجود در مجموعه‌ی داده‌ها

شماره	مجموعه‌ی داده	منفی	خنثی	مثبت
1	Miras Opinion [۱۵]	14,882	29,471	49,515
2	Snapfood [۱۶]	35,002	0	35,000
3	SentiPers [۱۸، ۱۷]	2,190	4000	7,856
4	Instagram Revs	3,237	2,495	2,780
5	Digikala Revs1 [۱۹]	460	419	2,382
6	Digikala Revs2	290	250	500
	Total	56,061	36,635	98,033

بکارگیری شده توسط محققین، نسخه‌ی از پیش آموزش دیده‌ی متشکل از ۱۲ سر^۱ مجهز به مکانیزم خود-توجه^۲ و ۱۲ لایه‌ی رمزنگار می‌باشد که پس از محاسبه‌ی خروجی توسط آن، نمونه‌ی مورد نظر توسط یک لایه‌ی تمام-متصل^۳ به کلاس‌های مشخص، دسته‌بندی می‌شود. لازم به ذکر است پژوهشگران در تحقیق مذکور، از پنج مجموعه داده با تعداد نمونه و کلاس‌های متمایز (دسته‌بندی‌های احساسی و موضوعات خبری) بهره‌برداری کرده‌اند.

از دیگر پژوهش‌های موجود در حوزه‌ی تشخیص احساسات به زبان فارسی می‌توان به [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 برای تحقق هدف ذکر شده، حدود ۸۰۰۰۰۰ توئیت فارسی را از سطح وبسایت توئیتر جمع‌آوری کرده، سپس با استفاده از مدل پیشنهادی آن‌ها را به سه کلاس مثبت، منفی و خنثی دسته‌بندی نموده‌اند.

علی‌رغم پژوهش‌های [۸، ۱۰، ۱۲، ۱۳] که محققان از تکنیک‌های مبتنی بر شبکه‌ی عصبی بهره برده‌اند، در [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 می‌باشد، با هدف بهبودی بیشتر در عملکرد چارچوب پیشنهادی (حدود ۸٪)، برای ایجاد لغت‌نامه مورد استفاده قرار گرفته است. در تحقیق مذکور، بهترین عملکرد، با

1- Head (Multi-Head Attention)

2- Self-Attention

3- Fully-Connected Layer

4- Snapfood

5- Instagram

6- Digikala

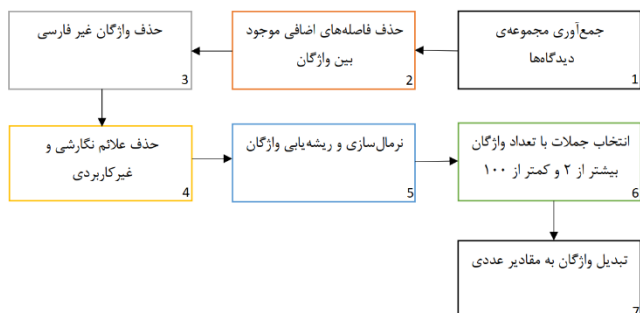
جدول (۲) - گزاره‌های هندسی مرتبط با مجموعه داده‌های شش گانه

۱	حداقل تعداد نشانه (Token) در مجموعه‌ی داده‌های ۱ تا ۶ (پیش از پاکسازی)
۱۵۳۵	حداکثر تعداد نشانه در مجموعه‌ی داده‌های ۱ تا ۶ (پیش از پاکسازی)
۳۱	میانگین تعداد نشانه در مجموعه‌ی داده‌های ۱ تا ۶ (پیش از پاکسازی)
۲,۷۰۴,۳۹۲	تعداد کل واژگان در مجموعه‌ی داده‌های ۱ تا ۶ (پس از پاکسازی)
۱۲,۵۷۴,۸۸۶	تعداد کل حروف در مجموعه‌ی داده‌های ۱ تا ۶ (پس از پاکسازی)

در ادامه، فرآیند پیش پردازش و چگونگی آماده‌سازی داده‌ها

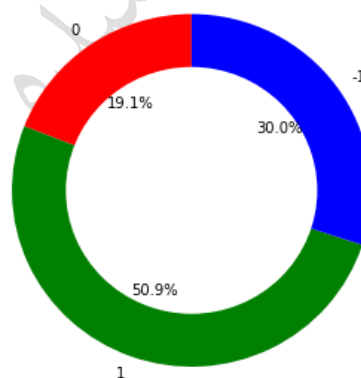
در بخش ۲,۳، به تفصیل بیان شده است.

۲,۳. پیش‌پردازش و آماده‌سازی داده‌ها



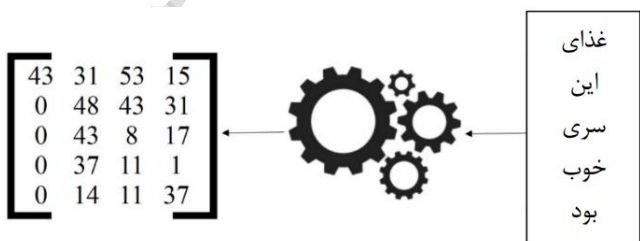
شکل (۲) - مراحل پیش پردازش و آماده‌سازی داده‌ها

برای پیش پردازش مجموعه‌ی 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 می‌گیرد 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 ریشه‌ی خود می‌باشد. ذکر این نکته لازم است که جهت 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.



شکل (۱) - چگونگی توزیع دیدگاه‌ها مبتنی بر چسب‌های سه گانه در مجموعه داده‌ی ۱ تا ۶

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، مجموعه‌ی جدیدی با تعداد ۱۸۴۴۳۸ نمونه حاصل می‌شود. شکل (۱) چگونگی توزیع مجموعه‌ی داده‌ی نهایی را در قالب برچسب‌های سه گانه (برچسب مثبت (1)، برچسب خنثی (0) و برچسب منفی (-1))، نشان می‌دهد. درخور ذکر است، از مجموعه‌ی نهایی ۱۰۰۰۰ نمونه به صورت تصادفی برای آزمایش و ۱۷۴۴۳۸ نمونه جهت آموزش انتخاب شده است.



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

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

در بخش ۴ به جزئیات چارچوب پیشنهادی و مولفه‌های 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.

۱.۴. تعبیه^۱

اولین مولفه از چارچوب پیشنهادی که تحت عنوان لایه‌ی تعبیه (مشابه [۲۰]) شناخته می‌شود، وظیفه‌ی انتقال حروف به مختصات برداری را بر عهده دارد. لایه‌ی ذکر شده با انجام ضرب داخلی بین ماتریس ورودی و وزن‌ها (معادله‌ی (۱))، به هر یک از حروف؛ برداری مشخص با طول ثابت (در مسئله‌ی حاضر مقدار ۵۰) اختصاص می‌دهد. لایه‌ی نگاشت^۲ نیز همانند تعبیه عمل کرده، منتها پس از ضرب در ماتریس وزن، با مقدار بایاس^۳ جمع شده و تابع فعالسازی^۴ مشخصی بر روی آن اعمال می‌شود.

(۱)

$$e = (X * W)$$

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

2- Mapping Layer

4- Activation Function

6- Kernel

8- Long-Short Term Memory

10- Time Steps

۲.۴. پیچش یک بعدی^۵

براساس فرآیند پیچش (معادله‌ی (۲))، ویژگی‌های موجود در هر واژه براساس اعمال بردارهای فیلتر بر روی حروف آن، استخراج و به لایه‌های بعدی موجود در شبکه‌ی عصبی، جهت تصمیم‌گیری منتقل می‌شوند [۲۱].

(۲)

$$c[i] = f(\sum(X[i:i+w-1] \circ K) + b)$$

در معادله‌ی (۲)، مقصود از X بردار ورودی (محاسبه شده توسط لایه‌ی تعبیه)، b مقدار بایاس، K بردار فیلتر، w اندازه‌ی هسته^۶، 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 اندازه‌ی هسته‌های (۱، ۲، ۳) جهت استخراج ویژگی هر واژه براساس 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^۷، می‌توان برای هر واژه یک بردار با تعداد ابعاد ۱۲۸ بدست آورد.

۳.۴. حافظه بلند-کوتاه مدت^۸

معماری حافظه‌ی بلند-کوتاه مدت (معادلات (۳) تا (۸)) [۲۲]، نوع خاصی از شبکه‌ی عصبی بازگشتی است که با هدف رفع مشکل گرادیان محو شدنی^۹ [۲۳] در هنگام پردازش داده‌های ترتیبی طراحی شده است. مشکل مذکور بر اثر محاسبه‌ی مشتق زنجیره‌ای از شبکه‌ی بازگشتی با تعداد گام‌های زمانی^{۱۰} طولانی (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

1- Embedding

3- Bias

5- One Dimensional Convolution

7- Global Max Pooling

9- Vanishing Gradient

اطلاعات در پردازش داده‌های متوالی می‌باشد. همچنین از LSTM، در حل مسائل دیگر همچون Named Entity Recognition [۲۴] و تشخیص گفتار [۲۵] نیز استفاده می‌شود.

پذیرفته شده در مجله محاسبات نرم

1- Embedding

3- Bias

5- One Dimensional Convolution

7- Global Max Pooling

9- Vanishing Gradient

2- Mapping Layer

4- Activation Function

6- Kernel

8- Long-Short Term Memory

10- Time Steps

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

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

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

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

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

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

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

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

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

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

در معادلات بالا، مقصود از b بردار بایاس، W ماتریس وزن، X ماتریس ورودی، \circ عملگر ضرب عنصر به عنصر، h_{t-1} وضعیت پنهان^۱ در لحظه $t-1$ ، δ تابع فعالسازی Sigmoid و \tanh تابع فعالسازی Hyperbolic Tangent می باشد.

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 ماتریس های وزن ضرب داخلی و با مقدار بایاس جمع شده، سپس نتایج براساس رابطه (۷) (Cell State) مورد پردازش قرار گرفته و خروجی در لحظه t طبق رابطه (۸) حاصل می شود.

۴.۴. واحد دروازه ای بازگشتی^۲

معماری LSTM علی رغم رفع مشکل گرادیان محو شدنی، با چالش تعداد پارامترهای زیاد همراه است. وجود ایراد مذکور، باعث ایجاد محدودیت در بحث حافظه جهت ذخیره سازی معماری های عظیم و مدت زمان آموزش آنها خواهد شد. جهت رفع دو مشکل ذکر شده، معماری دیگری تحت عنوان واحد بازگشتی دروازه ای (معادلات (۹) تا (۱۲)) [۲۶] ابداع شده است. شایان ذکر است، معماری مطرح شده در برخی مسائل همچون مدل سازی داده های صوتی [۲۷] نسبت به LSTM عملکرد بهتری را از خود به نمایش گذاشته است.

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 ای بازنشانی^۳، z_t دروازه ای بروزرسانی^۴، \tilde{c}_t دروازه ای جدید، δ تابع فعالسازی Sigmoid و \tanh تابع فعالسازی Hyperbolic Tangent می باشد.

شبکه های بازگشتی از قابلیت تحت عنوان اتصال دوطرفه نیز می توانند بهره مند شوند که امکان پردازش دو طرفه ای داده ها را برای آنها فراهم می کند. شایان توجه است، اتصال مذکور در برخی مسائل همچون پردازش متن مورد استفاده قرار می گیرد. روابط حاکم بر معماری های LSTM و GRU در ساختار مذکور نیز وجود دارند، ولی هنگام محاسبات، مقادیر از دو طرف مورد پردازش قرار می گیرند. برای مثال، در بحث تشخیص احساس موجود در جمله، لغات از اول به آخر و بالعکس، به شبکه وارد می شوند. نمایی از نحوه پردازش جمله توسط شبکه ای بازگشتی دو طرفه، در شکل (۴) قابل ملاحظه است.

1- Hidden State

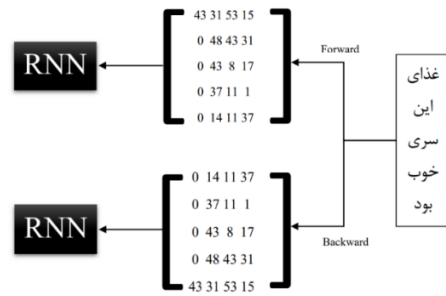
3- Reset Gate

5- Residual Connection

2- Gated Recurrent Unit

4- Update Gate

6- Feed-Forward



شکل (۴) - نمایی از شبکه‌ی عصبی بازگشتی دو طرفه

۵,۴. اتصال باقی‌مانده^۰

با توجه افزایش تعداد پارامترهای موجود در مسائل، ممکن است شبکه‌های عصبی پیش‌خور^۰ با تعداد محدودی از لایه‌های پنهان،

مجله محاسبات نرم

1- Hidden State

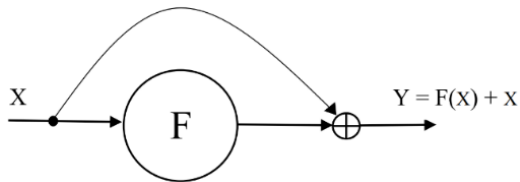
2- Gated Recurrent Unit

3- Reset Gate

4- Update Gate

5- Residual Connection

6- Feed-Forward



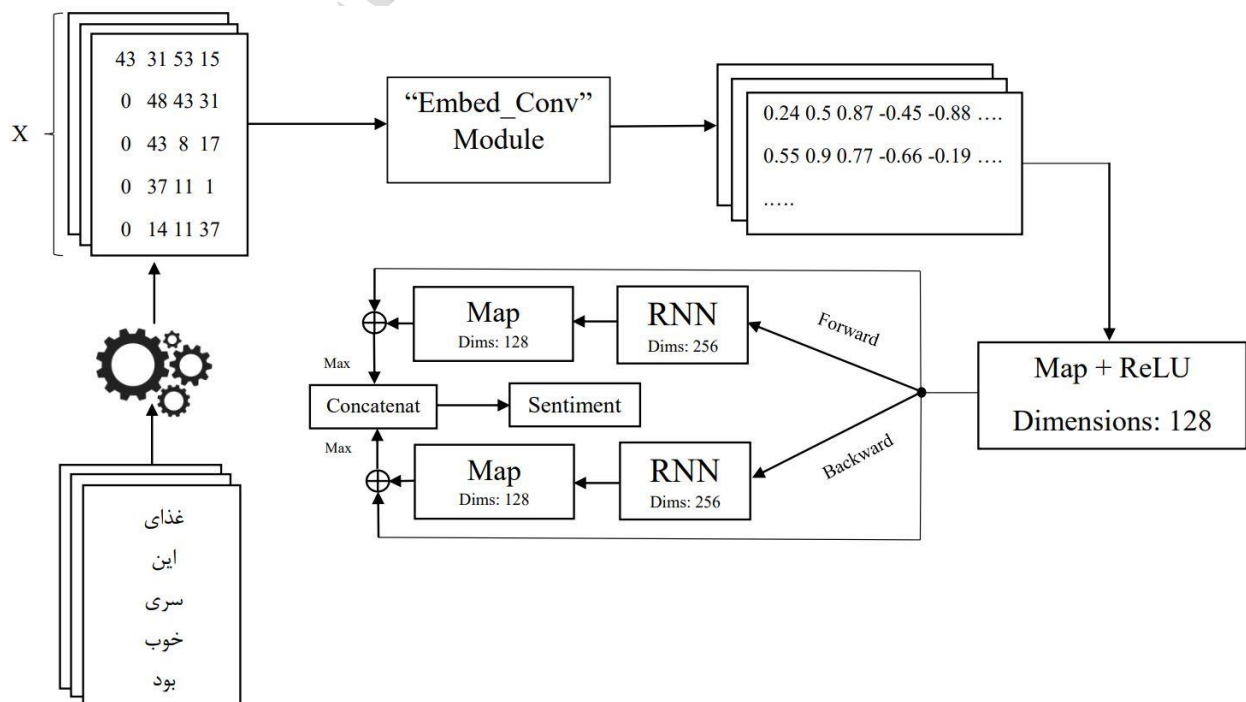
شکل (۵) - دیاگرام و معادله‌ی اتصال باقی‌مانده

در شکل (۵)، مقدار ورودی، X تابع فعال‌سازی و Y خروجی شبکه پس از جمع با ورودی F می‌باشد.

۵. معرفی چارچوب پیشنهادی

در پژوهش حاضر، دسته‌بندی جملات براساس پردازش حروف توسط لایه‌ی تعبیه، پیچش یک بعدی، شبکه‌ی بازگشتی دوطرفه و نگاشت انجام می‌شود. دیاگرام CBRCMN را در شکل (۶) می‌توان ملاحظه کرد. در CBRCMN، ابتدا دیدگاه مورد نظر پیش‌پردازش و آماده‌سازی شده، سپس طبق پیمان‌ه‌ی ۱ Embed_Conv (شکل (۷)) به واژه‌ی موجود در جمله، یک بردار ۱۲۸ بُعدی (مستقل از مفهوم) اختصاص داده می‌شود. در پیمان‌ه‌ی مذکور، ابتدا حروف تشکیل‌دهنده‌ی واژه توسط لایه‌ی تعبیه به یک فضای ۵۰ بُعدی منتقل شده، سپس ۳ دسته فیلتر با تعداد و اندازه‌های هسته‌ی مشخص بر روی آن‌ها (استخراج

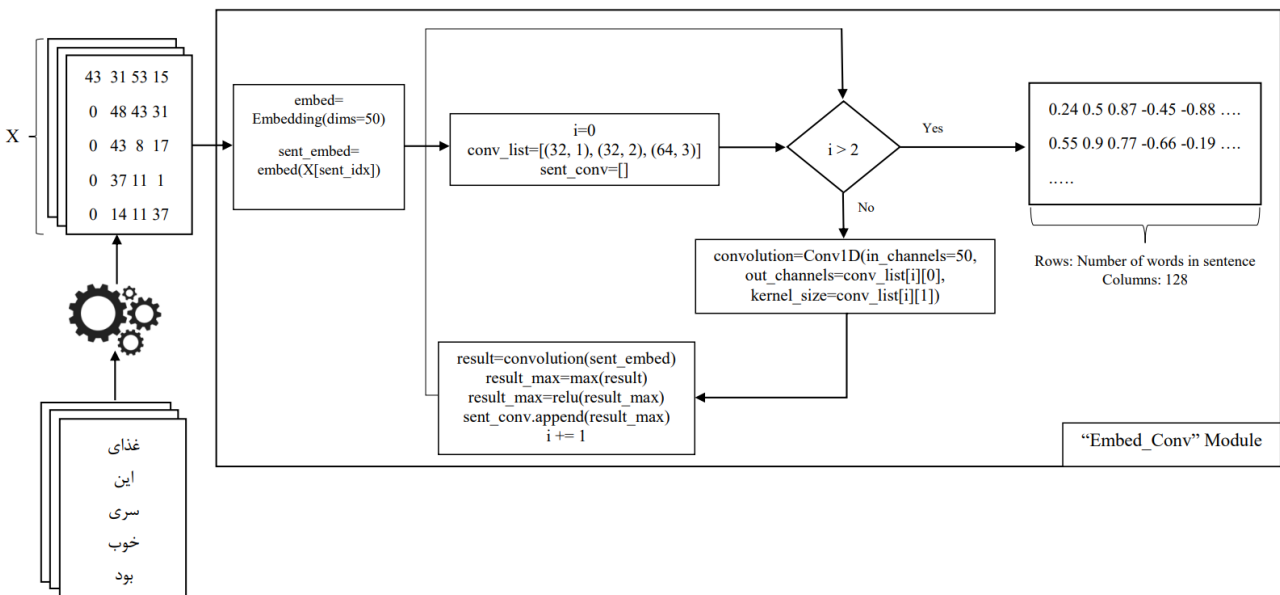
توانایی مدل‌سازی آن‌ها را نداشته باشند. همچنین، تمام داده‌ها از ماهیت دنباله‌ای برخوردار نیستند و نمی‌توان از معماری بازگشتی برای آن‌ها استفاده کرد. در نتیجه، برای پردازش چنین مسائلی به معماری پیش‌خور نیاز است. با این حال، در معماری ذکر شده با افزایش تعداد لایه‌های شبکه، مشکل گرادیان محو شدن برای لایه‌های ابتدایی ظاهر می‌شود. به عبارت دیگر، مقدار حاصل از محاسبه‌ی مشتق زنجیره‌ای تابع خطا نسبت به پارامتر لایه‌های اولیه، به صفر میل کرده و پارامترهای الگوریتم به درستی بروزرسانی نمی‌شوند. برای حل این مشکل، می‌توان از اتصال تحت عنوان "اتصال باقی‌مانده" (نام دیگر Skip Connection یا اتصال پرشی) بهره برد. روش نامبرده، با فراهم‌سازی یک مسیر برای جریان مشتق در جهت لایه‌های اولیه، از محو شدن مقدار گرادیان (میل به صفر) جلوگیری می‌کند. در مسئله‌ی حاضر، از اتصال ذکر شده جهت بروزرسانی پارامترهای موجود در لایه‌های اولیه‌ی شبکه، بهره‌برداری شده است. لازم به ذکر است، مقایسه‌ی نتایج حاصل از بکارگیری و عدم بکارگیری چنین اتصالی، در بخش ۵ مورد بررسی قرار گرفته است. همچنین، دیاگرام و معادله‌ی اتصال مذکور، در شکل (۵) قابل مشاهده است.



شکل (۶) - دیاگرام معماری پیشنهادی (CBRCMN)

دیدگاه ورودی تعیین می‌شود. همچنین، در صورت وجود

داده‌های کافی



شکل (۷) پیمانه‌ی Embed_Conv

حروف) اعمال می‌شود. پس از اجرای عمل

و بهره‌مندی الگوریتم از قدرت تعمیم بالا، می‌توان از بردارهای محاسبه شده توسط لایه‌های نگاشت، به عنوان بردارهای از پیش آموزش دیده (بر اساس وظیفه‌ی تشخیص احساسات) در مسائل دیگر استفاده کرد.

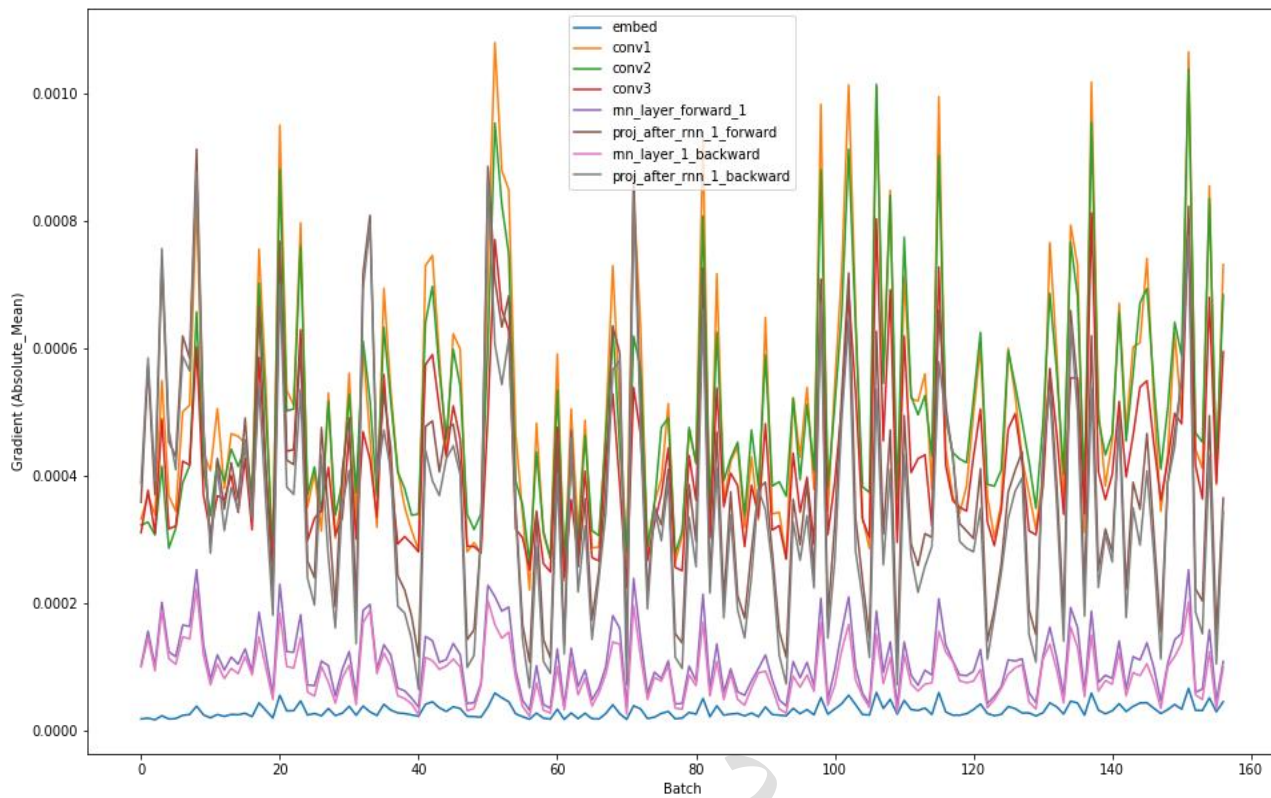
طبق توضیحات مطرح شده در بخش ۵،۴، دلیل بهره‌برداری از اتصال باقی‌مانده در چارچوب پیشنهادی، جلوگیری از پدیده‌ی گرادیان محو شدن در لایه‌های ابتدایی است. در دو شکل ۸ و ۹ می‌توان تاثیر اتصال مذکور را در تغییرات گرادیان لایه‌ها مشاهده کرد. نمودارهای ترسیم شده در تصاویر (۸) و (۹)، قدر مطلق میانگین تغییرات گرادیان در هر لحظه از آموزش (تقریباً به تعداد ۱۵۶ دسته‌ی آموزشی) را برای وزن‌دهی به لایه‌های چارچوب، نمایش می‌دهند. بر اساس شکل‌های (۸) و (۹)، 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. درحالی که با بکارگیری اتصال مذکور، مقادیر مشتق مربوط به لایه‌های

بزرگترین مقدار موجود در ویژگی‌های استخراج شده برای هر لغت براساس تعداد فیلتر انتخاب می‌شود. 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 واژه‌ی مورد نظر حاصل می‌شود. پس از کسب بردار برای هر واژه، ماتریس جدیدی با ابعاد ذکر شده در شکل (۷) تولید می‌شود. لازم به ذکر است در نگارش دستورات موجود در شکل (۷)، از قواعد زبان برنامه‌نویسی پایتون^۱ استفاده شده است.

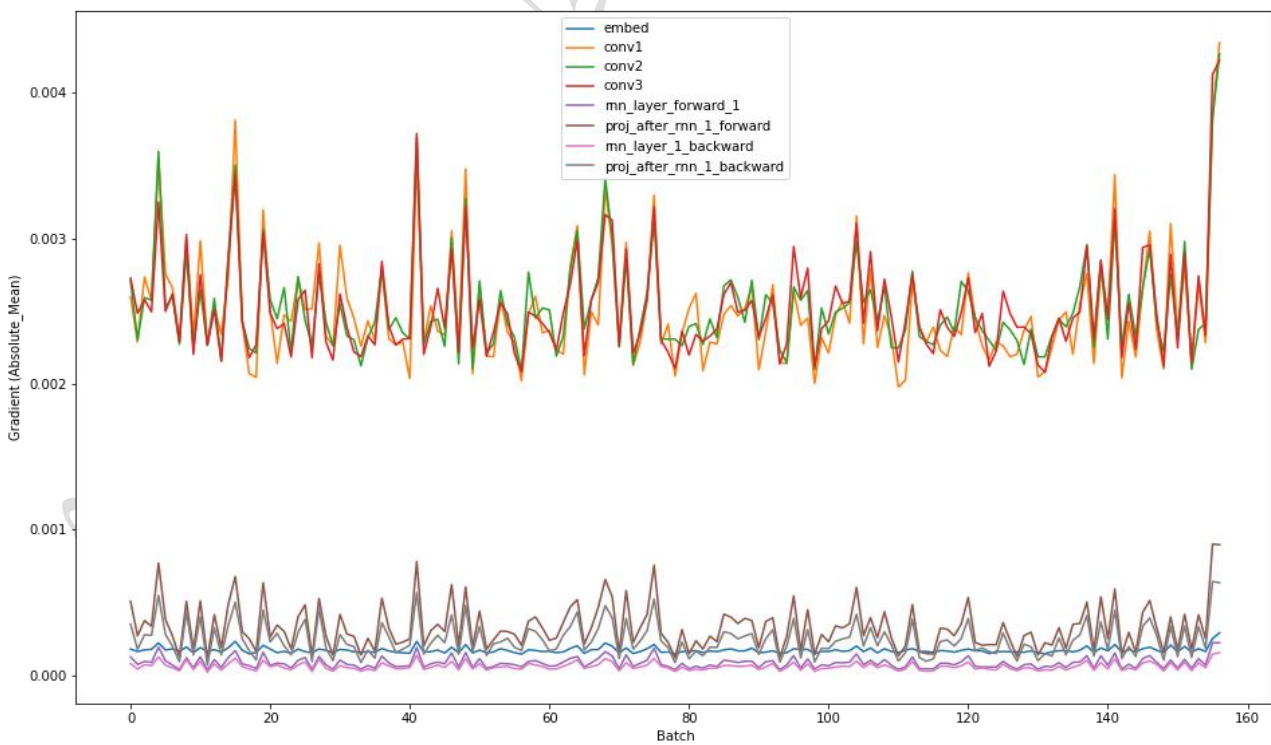
پس از انجام مراحل مذکور و تولید ماتریس واژگان، بردار لغات طی فرآیند نگاشت به فضای دیگری با ابعاد ۱۲۸ منتقل و تابع فعالسازی ReLU بر روی آن‌ها اعمال شده، سپس توسط شبکه‌ی عصبی بازگشتی (LSTM یا GRU) مورد پردازش قرار می‌گیرند. پس از کشف ارتباطات موجود بین واژگان و استخراج بردار مبتنی بر مفهوم توسط شبکه‌ی بازگشتی، نتایج پردازش به لایه‌ی بعد با تعداد ۳ گره منتقل شده و احساس مربوط به

conv1، conv2 و conv3، به طرز چشمگیری تغییر پیدا کرده‌اند. همچنین، تغییرات لایه‌ی embed (رنگ آبی) با مقداری جزئی هم‌راه است.

پذیرفته شده در مجله محاسبات نرم



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



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

مولفه‌ی دوم نمایان‌گر تعداد گره‌های موجود در لایه‌های بازگشتی می‌باشد.

چارچوب مورد نظر طبق جزئیات موجود در جدول (۳)، بر روی مجموعه دیدگاه‌های شرح داده شده (بخش ۳-۱)، آموزش یافته و جزئیات آموزش و عملکرد رویکردهای مورد نظر؛ در جدول (۴)، شکل‌های (۱۱)، (۱۲) و (۱۳) قابل مشاهده 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

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 لایه‌ی نگاشت ۱۲۸ بُعدی، میزان دقت ۷۹٫۸۷٪ و F-score 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 MaxPool1d، یک لایه جهت انتخاب بیشترین مقادیر و یک لایه خروجی) با میزان دقت ۷۶٫۵۶٪ و F-score ۷۶٫۷۴٪

در بخش ۶ به شرح نتایج مدل پیشنهادی و مثال عددی از شیوه‌ی عملکرد آن، پرداخته شده است.

۶. نتایج پژوهش

در زیربخش ۱، نتایج مدل پیشنهادی و مقایسه‌ی عملکرد آن با مدل‌های مبتنی بر واژه و در زیربخش ۲، نمونه‌ای عددی از عملکرد CBRCMN مورد بررسی قرار گرفته است.

۱.۶. نتایج و مقایسه با مدل‌های مبتنی بر واژه

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

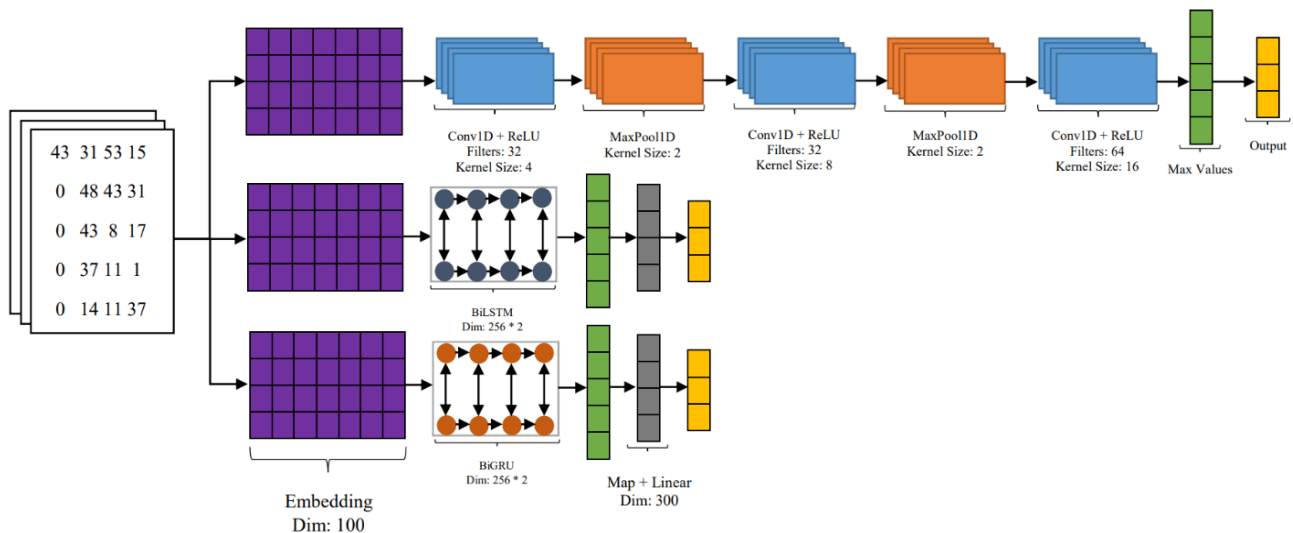
جدول (۳) - جزئیات آموزش چارچوب

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

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 جدول (۴)، ساختارهای چارچوب با ترکیب (map_dim, rnn_dim) معرفی شده‌اند که مولفه‌ی اول، معرف تعداد گره‌های موجود در لایه‌های نگاشت و

می‌باشد. معماری رویکردهای مبتنی بر واژه، در شکل (۱۰) مشهود است.

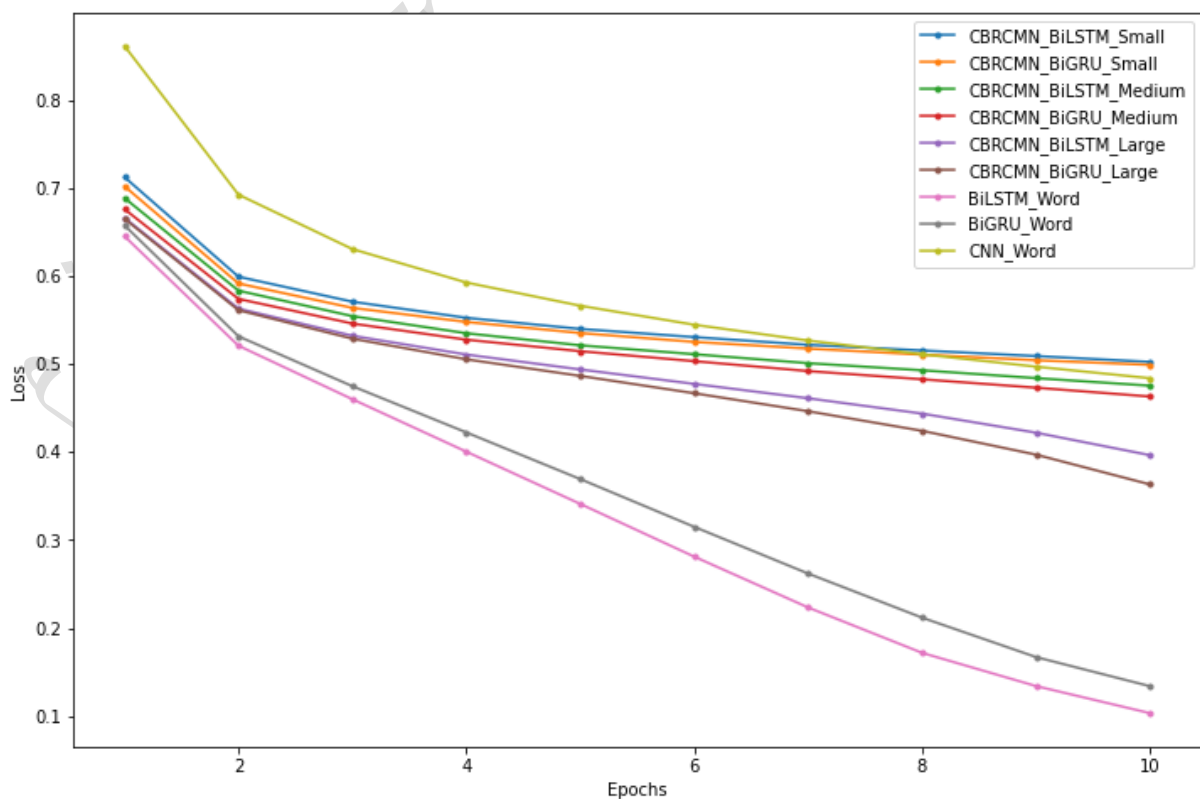
پذیرفته شده در مجله محاسبات نرم



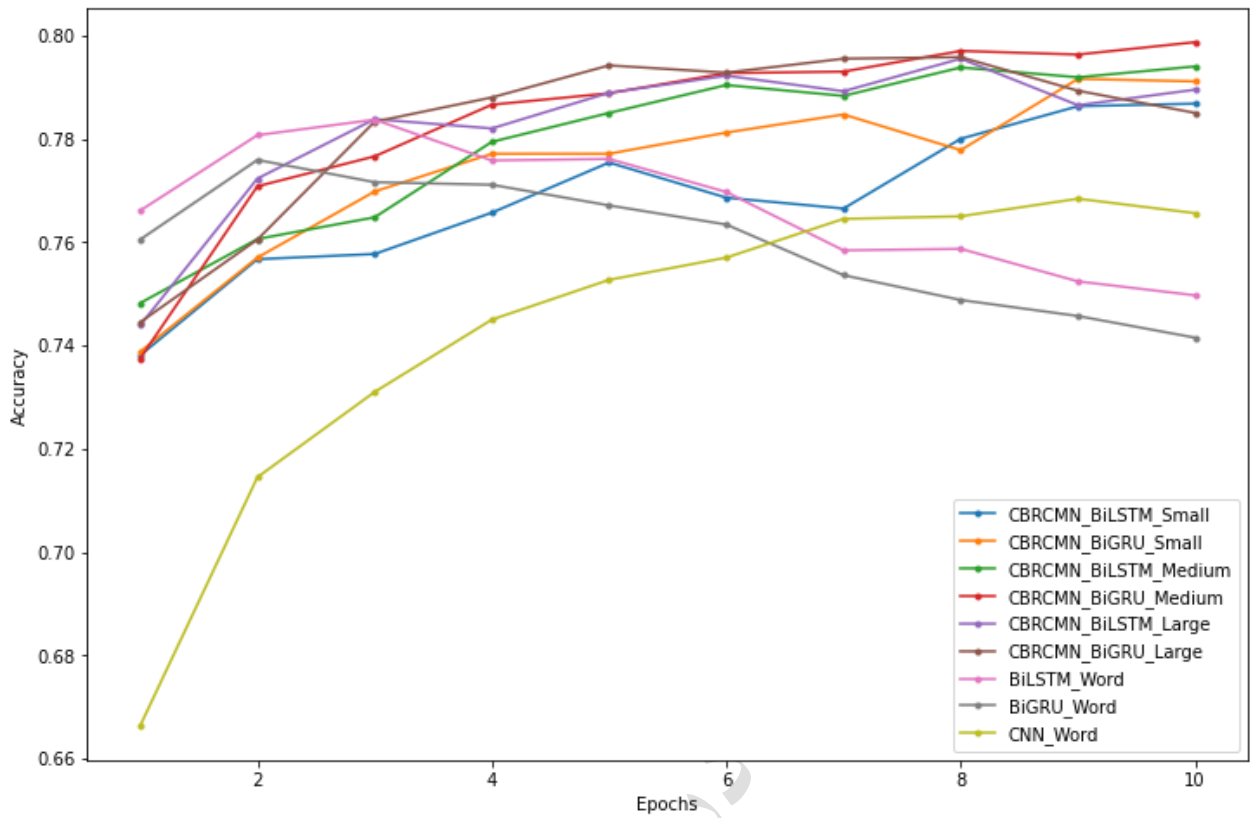
شکل (۱۰) - معماری رویکردهای مبتنی بر واژه

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

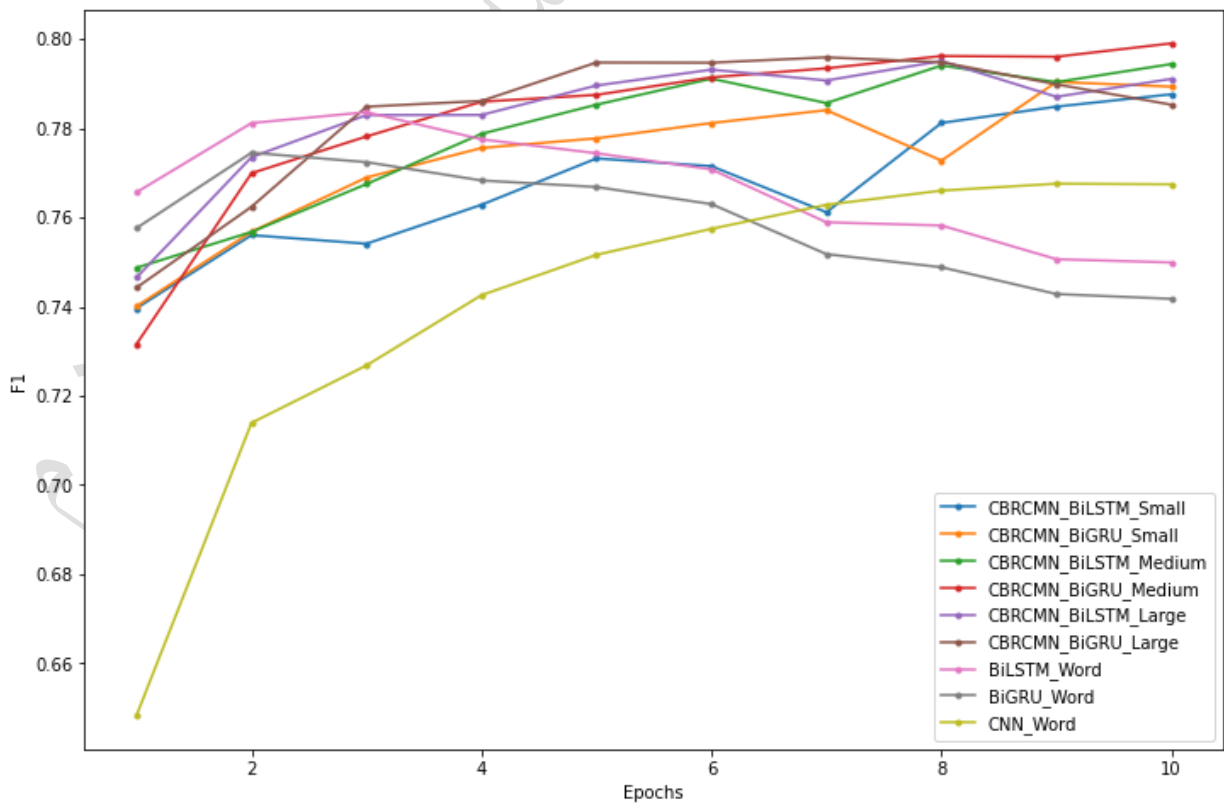
مدل‌های مبتنی بر واژه، بر مبنای دو رویکرد پردازشی، طرح‌ریزی شده‌اند. در رویکرد اول با بکارگیری عملگر پیچش و بردارهای فیلتر، ویژگی‌های سطح بالا از بردارهای واژگان موجود در جملات، استخراج شده و پس از تعیین بزرگترین مقادیر، با استفاده از یک لایه‌ی شبکه‌ی عصبی به سه گروه احساسی دسته‌بندی می‌شوند.



شکل (۱۱) - میزان خطا (نمونه‌های آموزش)



شکل (۱۲) - مقدار دقت (نمونه‌های آزمایش)



شکل (۱۳) - مقدار F-score (نمونه‌های آزمایش)

سوی بیش‌برازش^۱ میل می‌کنند. لازم به ذکر است با هدف درک بهتر از عملکرد مولفه‌های بکارگیری شده در CBRCMN، مطالعه‌ی پژوهش‌های [۲۹، ۳۰، ۳۱] به محققین توصیه می‌شود.

۲.۶. نمونه‌ی عددی از پیمانه‌ی Embed_Conv

به جهت ایجاد شفافیت در شیوه‌ی عملکرد پیمانه‌ی Embed_Conv، یک مثال کاربردی از استخراج ویژگی‌های ۱۲۸ بُعدی موجود در هر واژه (بر پایه‌ی پردازش حروف)، در شکل (۱۴) شرح داده شده است. طبق فرآیند فوق ابتدا نسخه‌ی پیش‌پردازش و پاکسازی شده‌ی جمله، حاصل شده سپس به تمام حروف تشکیل‌دهنده‌ی واژگان، مقادیر صحیح و مثبت تخصیص داده شده و ماتریسی از واژگان و حروف ایجاد می‌شود. در ماتریس ذکر شده، هر سطر و ستون بیان‌گر واژگان و حروف آن‌ها می‌باشد. پس از تعیین ماتریس اعداد و افزودن مقادیر صفر به بردارها با هدف ایجاد طول ثابت (عمل Padding)، عمل تعبیه در سطح حروف بر روی ماتریس اعمال شده و ماتریس چهاربُعدی 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 نامبرده و انتقال سطر و ستون‌ها (ترانواده)، سه سطح لایه‌ی پیش‌با اندازه هسته‌های ۱، ۲ و ۳ با تعداد کانال ۳۲، ۳۲ و ۶۴ بر روی آن اعمال شده و وابستگی‌های محلی موجود در بین حروف تعیین می‌شود. به عبارت دیگر، با بکارگیری پیمانه‌ی 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 بُعدی برای هر واژه تعیین می‌شود. لازم به ذکر است با هدف پیشگیری از ایراد گرادیان محوشدنی، تابع فعالسازی ReLU بر روی مقادیر اعمال شده و بردارها به یکدیگر الحاق می‌شوند (کسب بردار ویژگی ۱۲۸

مطابق شکل (۱۱)، علی‌رغم پیاده‌سازی‌های CBRCMN، مدل‌های مبتنی بر واژه که به شبکه‌ی بازگشتی تجهیز شده‌اند (BiLSTM_Word، BiGRU_Word) از سرعت بیشتری در امر آموزش بهره‌مند می‌باشند. به عبارت دیگر، مدل‌های مبتنی بر واژه پس از ۱۰ دوره آموزش، به کمترین میزان خطا در داده‌های آموزشی دست یافته‌اند. شایان توجه است 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 (رابطه‌ی (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 (رابطه‌ی (۱۴)) با هدف بررسی دقیق عملکرد مدل‌ها، بهره‌برداری شده است و نتایج معیار مذکور در شکل (۱۳) قابل ملاحظه می‌باشد.

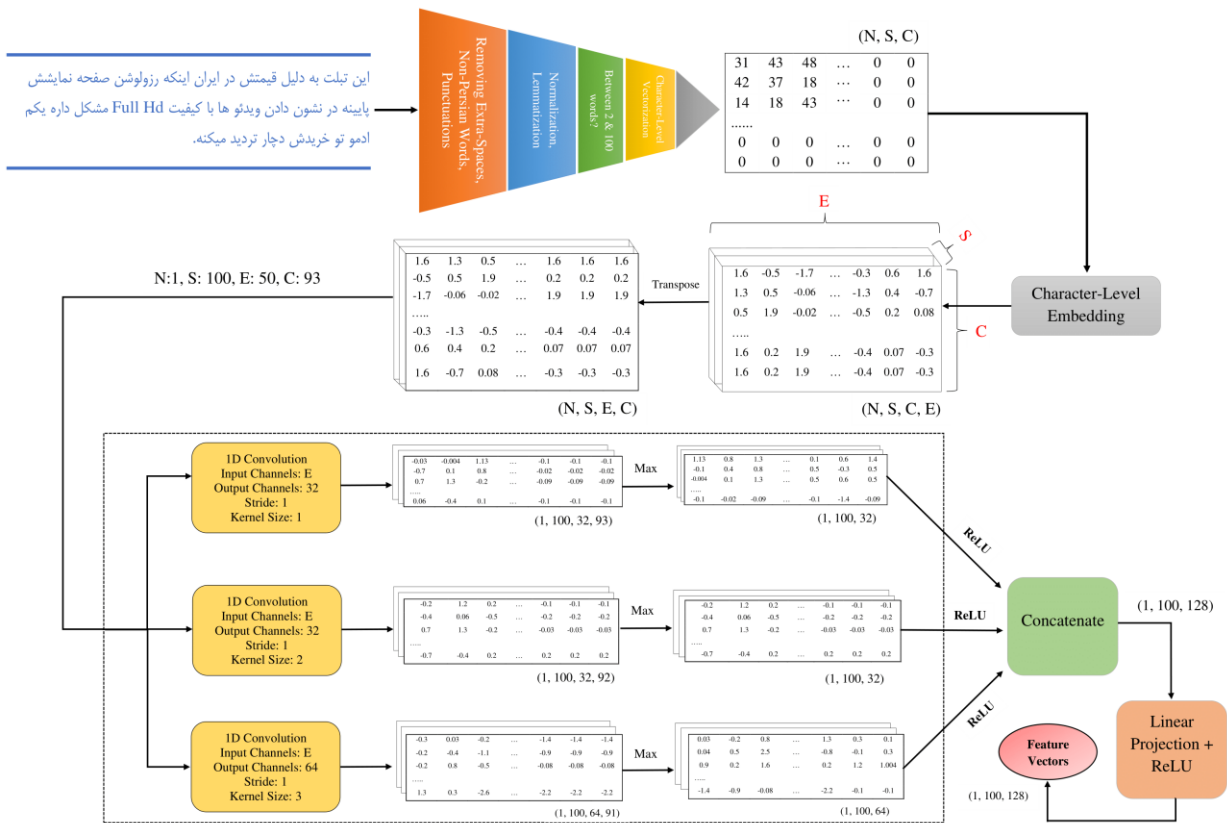
(۱۳)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

در روابط (۱۳) و (۱۴)، مقصود از TP: True Positive، TN: True Negative، FN: False Negative و FP: False Positive می‌باشد. همچنین، با توجه به نتایج مندرج در شکل‌های (۱۲) و (۱۳) مشاهده می‌شود در دومین و سومین دوره‌ی آموزش، BiLSTM_Word و BiGRU_Word به بیشترین میزان دقت و F-score خود رسیده‌اند و پس از طی ۱۰ دوره‌ی آموزشی به

بعدی). همچنین با هدف تعیین تعداد ابعاد مشخص و استفاده از اتصال باقی مانده، از یک لایه‌ی نگاشت با تابع ReLU استفاده بکارگیری شده است. گفتنی است شیوه‌ی کارکرد پیمان‌هی بیان شده در برخورد با نظرات دیگر، مشابه شکل (۱۴) می‌باشد.

پذیرفته شده در مجله محاسبات نرم



شکل (۱۴) - نمونه‌ی عددی از کارکرد پیمانه‌ی Embed_Conv

جدول (۵) - تعداد پارامتر معماری‌های نامبرده

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

با توجه به جدول (۵)، ملاحظه می‌شود پیاده‌سازی‌های معماری CBRCMN در مقایسه با مدل‌های مبتنی بر واژه، از تعداد پارامترهای پایین‌تری بهره‌مند می‌باشند. نکته‌ی اشاره شده، منجر به کاهش حجم مورد نیاز برای ذخیره‌سازی مدل برای مسائل بزرگ‌تر، می‌شود.

دلیل بهبودی عملکرد فرآیند تشخیص احساسات توسط چارچوب پیشنهادی را می‌توان در استفاده از شبکه‌های بازگشتی و پیمانه‌ی Embed_Conv جستجو نمود. علی‌رغم مدل‌های مبتنی بر واژه، در CBRCMN پس از چهار گام محاسباتی (عملگرهای تعبیه و پیش‌چسب)، 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. لازم به ذکر است با هدف دستیابی محققان به کدهای پیاده‌سازی، اسناد مربوطه در مخزن گیت‌هاب^۱ قابل دستیابی می‌باشند.

با هدف کسب تعداد پارامترهای هر مدل، کتابخانه‌ی torchinfo بکارگیری شده است.

۷. نتیجه‌گیری و پژوهش‌های آتی

در این پژوهش به بررسی مطالعات انجام شده در حوزه‌ی 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 دیجی‌کالا، اسنپ‌فود و اینستاگرام گردآوری شده و با اعمال تکنیک‌های پیش‌پردازش، زمینه‌ی لازم برای ورود به چارچوب یادگیری ماشین پیشنهادی فراهم شده است. پس از شرح مجموعه‌ی داده‌ها، به معرفی و بررسی اجزای سازنده‌ی چارچوب پیشنهادی (CBRCMN)، چالش‌ها و راه‌حل‌های موجود، معرفی مدل اصلی و عملکرد چارچوب بر روی مجموعه داده‌های شش‌گانه پرداختیم. مدل پیشنهادی از یک لایه تعبیه‌ی واژگان، سه دسته پیش‌بینی یک بُعدی مستقل، سه لایه نگاشت و یک لایه‌ی بازگشتی دو طرفه (تجهیز شده به اتصال پُرشی) 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 شبکه‌ی پیش‌بینی یک بُعدی با میزان دقت ۷۶٫۵۶٪ می‌باشد. معماری مذکور از سه لایه پیش‌بینی یک بُعدی (تعداد فیلترهای ۳۲، ۳۲ و ۶۴) و دولایه MaxPool1D تشکیل شده است. در مطالعات آینده قصد داریم از چارچوب معرفی شده، برای پیش‌برد مسائلی مانند خلاصه‌سازی و ترجمه‌ی متون استفاده و عملکرد آن‌ها را بررسی نماییم.

مراجع

- [2] Rao, A., Spasojevic, N., "Actionable and Political Text Classification Using Word Embeddings and Lstm.", arXiv (2016). <https://doi.org/10.48550/arXiv.1607.02501>
- [3] Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L., "Deep Contextualized Word Representations.", NAACL 1: 2227-2237 (2018). <http://dx.doi.org/10.18653/v1/N18-1202>
- [4] Mikolov, T., Chen, K., Corrado, G., Dean, J., "Efficient estimation of word representations in vector space", arXiv (2013). <https://doi.org/10.48550/arXiv.1301.3781>
- [5] Devlin, J., Chang, M. W., Lee, K., Toutanova, K., "Bert: Pre-training of deep bidirectional transformers for language understanding", NAACL 1: 4171-4186 (2019). <http://dx.doi.org/10.18653/v1/N19-1423>
- [6] Dashtipour, K., Gogate, M., Adeel, A., Hussain, A., Alqarafi, A., Durrani, T., "A comparative study of persian sentiment analysis based on different feature combinations", International Conference in Communications, Signal Processing, and Systems: 2288-2294 (2019). https://doi.org/10.1007/978-981-10-6571-2_279
- [7] Ghasemi, S., Jadidinejad, A. H., "Persian text classification via character-level convolutional neural networks", 8th Conference of AI & Robotics and 10th RoboCup Iranopen International Symposium (IRANOPEN): 1-6 (2018). <https://doi.org/10.1109/RIOS.2018.8406623>
- [8] Zobeidi, S., Naderan, M., Alavi, S. E., "Opinion mining in Persian language using a hybrid feature extraction approach based on convolutional neural network", Multimedia Tools and Applications 78(22): 32357-32378 (2019). <https://doi.org/10.1007/s11042-019-07993-4>
- [9] Shumaly, S., Yazdinejad, M., Guo, Y., "Persian sentiment analysis of an online store independent of pre-processing using convolutional neural network with fastText embeddings", PeerJ Computer Science (2021). <https://doi.org/10.7717/peerj-cs.422>
- [10] Dashtipour, K., Gogate, M., Adeel, A., Larijani, H., Hussain, A., "Sentiment Analysis of Persian Movie Reviews Using Deep Learning", Entropy 23(5): 596 (2021). <https://doi.org/10.3390/e23050596>

- [1] Pang, B., Lee, L., Vaithyanathan, S., "Thumbs Up? Sentiment Classification Using Machine Learning Techniques.", EMNLP 10: 79-86 (2002). <http://dx.doi.org/10.3115/1118693.1118704>

- Information Processing Systems 33: 3775-3785 (2020).
<https://doi.org/10.48550/arXiv.2004.12073>
- [21] Kim, Y., Jernite, Y., Sontag, D. and Rush, A., March. "Character-aware neural language models", In Proceedings of the AAAI conference on artificial intelligence 30(1): (2016).
<https://dl.acm.org/doi/10.5555/3016100.3016285>
- [22] Hochreiter, S., Schmidhuber, J., "Long short-term memory", Neural Computation 9(8): 1735-1780 (1997).
<https://doi.org/10.1162/neco.1997.9.8.1735>
- [23] Pascanu, R., Mikolov, T., Bengio, Y., "On the difficulty of training recurrent neural networks", The 30th International Conference on Machine Learning 28(3): 1310-1318 (2013).
<https://doi.org/10.48550/arXiv.1211.5063>
- [24] Chiu, J. P. C., Nichols, E., "Named entity recognition with bidirectional LSTM-CNNs", Transactions of the Association for Computational Linguistics 4: 357-370 (2016).
https://doi.org/10.1162/tacl_a_00104
- [25] Graves, A., Mohamed, A. R., Hinton, G., "Speech recognition with deep recurrent neural networks", IEEE International Conference on Acoustics, Speech and Signal Processing: 6645-6649 (2013).
<https://doi.org/10.1109/ICASSP.2013.6638947>
- [26] Cho, K., Van Merriënboer, B., Bahdanau, D., Bengio, Y., "On the properties of neural machine translation: Encoder-decoder approaches", SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation: 103-111 (2014).
<http://dx.doi.org/10.3115/v1/W14-4012>
- [27] Chung, J., Gulcehre, C., Cho, K., Bengio, Y., "Empirical evaluation of gated recurrent neural networks on sequence modeling", NIPS 2014 Workshop on Deep Learning (2014).
<https://doi.org/10.48550/arXiv.1412.3555>
- [28] Kingma, D. P., Ba, J., "Adam: A method for stochastic optimization", 3rd International Conference on Learning Representations (2015).
<https://doi.org/10.48550/arXiv.1412.6980>
- [29] Feizi-Derakhshi, M.R., Mottaghinia, Z. and Asgari-Chenaghlu, M., "Persian Text Classification Based on Deep Neural Networks", Soft Computing Journal 11(1) (2022). 10.22052/SCJ.2023.243182.1010
- [11] Satrya, W.F., Aprilliyani, R. and Yossy, E.H., "Sentiment analysis of Indonesian police chief using multi-level ensemble model", Procedia Computer Science 216: 620-629 (2023).
<https://doi.org/10.1016/j.procs.2022.12.177>
- [12] Alruily, M., Manaf Fazal, A., Mostafa, A.M. and Ezz, M., "Automated Arabic Long-Tweet Classification Using Transfer Learning with BERT", Applied Sciences 13(6): 3482 (2023).
<https://doi.org/10.3390/app13063482>
- [13] Nezhad, Z.B. and Deihimi, M.A., "Twitter sentiment analysis from Iran about COVID 19 vaccine", Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 16(1): 102367 (2022).
<https://doi.org/10.1016/j.dsx.2021.102367>
- [14] Ahangari, M. and Sebti, A., "A Hybrid Approach to Sentiment Analysis of Iranian Stock Market User's Opinions", International Journal of Engineering 36(3): 573-584 (2023).
<https://doi.org/10.5829/ije.2023.36.03c.18>
- [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", The 12th Language Resources and Evaluation Conference: 2855-2861 (2020).
<https://aclanthology.org/2020.lrec-1.348>
- [16] Farahani, M., Gharachorloo, M., Farahani, M., Manthouri, M., "ParsBERT: Transformer-based Model for Persian Language Understanding", Neural Process Letters 53(6): 3831-3847 (2021).
<https://doi.org/10.1007/s11063-021-10528-4>
- [17] Hosseini, P., Ramaki, A. A., Maleki, H., Anvari, M., Mirroshandel, S. A., "SentiPers: a sentiment analysis corpus for Persian", arXiv (2018).
<https://doi.org/10.48550/arXiv.1801.07737>
- [18] Sharami, J. P. R., Sarabestani, P. A., Mirroshandel, S. A., "Deepsentipers: Novel deep learning models trained over proposed augmented persian sentiment corpus", arXiv (2020).
<https://doi.org/10.48550/arXiv.2004.05328>
- [19] Samizadeh, M., "Sentiment-Analysis-with-LSTM-in-Persian" (2018).
<https://doi.org/10.5281/zenodo.6862064>
- [20] Takase, S. and Kobayashi, S., "All word embeddings from one embedding", Advances in Neural

[30] Zare Mehrjardi, F., Yazdian-Dehkordi, M. and Latif, A., "Evaluating Machine learning and Deep-learning methods in Sentiment Analysis on Persian Telegram Message", *Soft Computing Journal* 11(1) (2022). 10.22052/SCJ.2023.246553.1077

[31] Khosravi, A. and Abdolhosseini, H., "Personality in social networks using thematic modeling of user feedback", *Soft Computing Journal* 11(1) (2023). 10.22052/SCJ.2023.243197.1006

پذیرفته شده در مجله محاسبات نرم