نوع مقاله: پژوهشی

چارچوبی برای تحلیل احساسات داده‌های متنی و شکلک‌ها در شبکه‌های اجتماعی

مرتضی شریفی[۱]، کارشناسی ارشد، حجت اله حمیدی[۲]، دانشیار

۱- گروه فناوری اطلاعات- دانشگاه صنعتی خواجه نصیرالدین طوسی - تهران - ایران - m.sharifi@email.kntu.ac.ir
۲- گروه فناوری اطلاعات- دانشگاه صنعتی خواجه نصیرالدین طوسی - تهران - ایران - h_hamidi@kntu.ac.ir

چکیده: امروزه افراد به‌منظور افزایش احساس خود در متن یا خلاصه کردن عبارات ،از شکلک‌ها استفاده می‌کنند، تکنیک‌های قبلی یادگیری ماشین فقط شامل طبقه‌بندی متن، شکلک یا تصاویر، به‌تنهایی بوده است، جایی که شکلک‌ها با متن اکثر اوقات نادیده گرفته شده است. تحقیقات نشان می‌دهد یادگیری عمیق به‌ندرت در تحلیل احساسات بر روی ترکیب داده‌های متن و شکلک اعمال‌شده است. بنابراین، این مقاله متن و شکلک‌ها را به‌طور جداگانه و به‌صورت ترکیبی برای یافتن احساسات، تجزیه‌وتحلیل کرده است. بدین‌صورت که ابتدا یک تجزیه‌وتحلیل مقایسه‌ای از شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت انجام می‌گردد، که جهت افزایش دقت تعبیه کلمات، بردارهای تعبیه کلمات از قبل آموزش‌دیده word2vec، glove، BERT و RoBERTa مورداستفاده قرار گرفته است. سپس یک مدل ترکیبی مبتنی بر توجه ارائه می‌گردد که از شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت دوطرفه استفاده می‌کند. نتایج نشان می‌دهد که درنظر گرفتن شکلک‌ها به عنوان ویژگی، سبب افزایش دقت مدل شده است، همچنین مدل پیشنهادی بر اساس چهار معیار ارزیابی دقت، صحت، فراخوانی و امتیاز f1 از مدل‌های پایه و کارهای قبلی بررسی شده، بهتر عمل کرده، که دقت مدل پیشنهادی بر روی مجموعه داده بدون شکلک ۸۵/۸۰ بوده و برای مجموعه داده شامل شکلک ۹۰/۱۸ است.

واژه‌های کلیدی: تجزیه‌وتحلیل احساسات، عقیده کاوی، پردازش زبان طبیعی، یادگیری عمیق.

A framework for sentiment analysis of textual data and emoticons on social networks

Morteza Sharifi, MSc Student[1], Hojatollah Hamidi, Associate Professor[2]

1- Information Technology Engineering Group, K.N.Toosi University of Technology, Tehran, Iran, Email:m.sharifi@email.kntu.ac.ir
2- Information Technology Engineering Group, K.N.Toosi University of Technology, Tehran, Iran, Email: h_hamidi@kntu.ac.ir

Abstract: Nowadays, people use emoticons in the text to increase their feelings or summarize expressions. Earlier machine learning techniques only involve the classification of text, emoticons, or images solely, whereas emoticons with text have most of the time been neglected. Research shows deep learning was rarely applied in sentiment analysis on text and emoticon data combination. Therefore, this article has analyzed the text and emoticons separately and in combination to find the emotions. First a comparative analysis of convolutional neural networks and long-short term memory networks is performed, to increase the accuracy of word embedding, pre-trained word embedding vectors word2vec, glove, BERT and RoBERTa were used. Then, a new hybrid attention-based model is presented, which uses convolutional neural networks and Bidirectional long-short term memory networks. The results show that considering emoticons as a feature increases the accuracy of the model, also the proposed model has performed better than the basic models and previously reviewed works based on the four evaluation criteria of accuracy, precision, recall, and f1 score that accuracy of the proposed model on the dataset without emoticons is 85.80 and on the dataset including emoticons is 90.18.

Keywords: Sentiment analysis, Opinion mining, Natural language processing, Deep learning.

تاریخ ارسال مقاله:۱۴۰۱/۰۶/۱۷
تاریخ اصلاح مقاله: ۱۴۰۲/۰۲/۱۶
تاریخ پذیرش مقاله: ۱۴۰۲/۰۴/۲۱
نام نویسنده مسئول: حجت اله حمیدی
نشانی نویسنده مسئول: ایران - تهران – دانشگاه صنعتی خواجه نصیرالدین طوسی – گروه آموزشی فناوری اطلاعات.

۱- مقدمه

تجزیه و تحلیل احساسات[۱]در نظارت بر شبکه‌های اجتماعی بسیار ارزشمند است زیرا به شما امکان می‌دهد بینشی در موضوعات خاص بدست آورید [۱]. رسانه‌های اجتماعی مانند توییتر، فیس‌بوک و یوتیوب در حال حاضر از محبوب‌ترین مکان‌هایی هستند که مشتریان می‌توانند محصولات و خدمات جدید را در بازارهای مختلف مورد بحث و بررسی قرار دهند [۲]. به‌طور سنتی، تجزیه و تحلیل احساسات بر روی متن انجام شده است، اما مقدار زیادی از داده‌ها اکنون به‌عنوان نظرات، شامل شکلک‌ها هم هستند. با بررسی این داده‌ها، می‌توان احساسات عمومی نسبت به یک موضوع خاص را تحلیل، بررسی و کشف کرد [۳]. مدل تجزیه و تحلیل احساسات، بسته به اینکه شکلک‌ها[۲] را در نظر می‌گیرد یا آن‌ها را نادیده می‌گیرد، حالت‌های مختلفی را برای هر توییت، تولید می‌کند [۴]. در سال‌های اخیر، نشان داده‌شده است که استفاده از شبکه‌های عصبی عمیق در طبقه‌بندی متن مؤثر بوده است[۵]، همچنین تحقیقات اخیر در مورد کشف قطبیت شامل تکنیک‌های شبکه عصبی کانولوشنی[۳]، حافظه طولانی کوتاه‌مدت[۴] و جاسازی کلمات[۵] است [۶]. شبکه عصبی کانولوشنی در ابتدا برای حوزه بینایی ماشین ساخته شد. پس‌ازآن، شبکه‌های عصبی کانولوشنی برای اهداف پردازش زبان طبیعی مورداستفاده قرار گرفتند و برای طبقه‌بندی متن و تحلیل احساسات به نتایج عالی دست‌یافتند [۷ ، ۸]. حافظه طولانی کوتاه‌مدت، یک مدل یادگیری عمیق مؤثر در تجزیه و تحلیل دنباله‌های طولانی، برای طبقه‌بندی احساسات در زمینه پردازش زبان طبیعی مورداستفاده قرارگرفته است [۹]. روش‌های یادگیری عمیق از جاسازی کلمات برای یادگیری روابط معنایی کلمات به‌منظور بهبود عملکرد مدل استفاده می‌کنند [۱۰]. در این مقاله اهمیت قرار دادن شکلک‌ها در پیام‌های توییتر بررسی می‌شود و یک مدل ترکیبی مبتنی بر توجه شبکه عصبی کانولوشنی و شبکه حافظه طولانی کوتاه‌مدت پیاده‌سازی می‌گردد.

تحلیل احساسات زیرمجموعه پردازش زبان طبیعی[۶] است. تحقیقات بسیار کمی در مورد شکلک‌ها برای یافتن احساسات انجام‌شده است [۱۱]. زمینه برای گسترش بیشتر در حوزه تحلیل احساسات با متن و شکلک وجود دارد. اهداف این تحقیق عبارت‌اند از:

۱) بررسی تأثیر شکلک‌ها در تجزیه‌وتحلیل احساسات داده‌های شبکه‌های اجتماعی
۲) بررسی دقت طبقه‌بندی تجزیه‌وتحلیل احساسات با استفاده از الگوریتم‌های یادگیری عمیق
۳) بررسی تأثیر بردارهای تعبیه کلمات از قبل آموزش‌دیده مختلف در افزایش دقت مدل
۴) ارائه یک مدل ترکیبی مبتنی بر توجه با استفاده از الگوریتم‌های یادگیری عمیق برای تجزیه‌وتحلیل احساسات

در ادامه، مقاله به شکل زیر سازماندهی شده است: در بخش دوم، مروری بر تحقیقات پیشین بیان شده است، در بخش سوم، سطوح تجزیه و تحلیل احساس بررسی شده، در بخش چهارم، پیاده‌سازی روش پیشنهادی ارایه شده است، در بخش پنجم، نتایج مدل‌ها، در بخش ششم، ارزیابی نتایج و در بخش پایانی، نتیجه‌گیری و پیشنهاد برای تحقیقات آتی، بیان شده است.

۲- مروری بر تحقیقات پیشین

فاکتور تازگی، همراه با کاربردهای گسترده و در دسترس بودن داده‌های رسانه‌های اجتماعی محققان زیادی را به حوزه تجزیه و تحلیل احساسات جذب کرده است. تحقیقات جدید تجزیه‌وتحلیل احساسات اکنون در حوزه‌ها و زبان‌های مختلفی انجام‌شده است [۱۲-۱۳]. وان و گائو از یک روش جمعی طبقه‌بندی احساسات با کمک روش‌های مختلف طبقه‌بندی مانند بیز ساده، ماشین بردار پشتیبان، درخت تصمیم‌گیری و جنگل تصادفی استفاده کردند، آن‌ها شش تکنیک طبقه‌بندی را آزمایش کردند و با توجه به رویکرد جمعی آموزش داده‌شده و تست شدند[۱۴]. پاک و پروبک داده‌های میکروبلاگ را طبقه‌بندی کرده و احساسات مثبت، منفی و خنثی را تعیین کردند. سهم اصلی جمع‌آوری خودکار مجموعه داده‌ها و تحلیل احساسات است، تکنیک پیشنهادی آن‌ها کارآمد بوده و عملکرد بهتری نسبت به روش‌های پیشنهادی قبلی دارد [۱۵]. ویزلاو ولنی داده‌های توییتر را با استفاده از روش‌های یادگیری ماشین طبقه‌بندی کرد. تعداد محدودی از شکلک‌های استاندارد از داده‌های توییتر شناسایی شد و احساسات مربوطه را بدست آورد [۱۶]. کولومبیس و همکاران، ۲۰۸۰ توییت به زبان هلندی را که حاوی شکلک هستند ارزیابی کردند. نشان دادند شرح سطح پاراگراف برای احساسات دلالت شده توسط شکلک‌ها به طرز قابل‌توجهی از دقت طبقه‌بندی احساسات پیشی می‌گیرد که نشان می‌دهد مشارکت شکلک‌ها بر احساسات منتقل‌شده توسط نشانه‌های متنی غلبه می‌کند [۱۷]. تعبیه کلمه به‌عنوان نمایش کلمات جهت تحلیل متن، معمولاً در قالب یک بردار که معنای کلمه را رمزگذاری می‌کند، تعریف می‌شود. به‌این‌ترتیب انتظار می‌رود کلماتی که در یک فضای برداری نزدیک‌تر ظاهر می‌شوند معانی مشابهی داشته باشند. مدل‌هایی مانند word2vec و Glove معنای لغوی کلمات را در نظر می‌گیرند [۱۸]. موقعیت‌هایی وجود دارد که در آن‌ها جنبه ترتیبی متن مهم جلوه می‌کند. از این نظر، جاسازی‌های کلمه، تکنیک‌های تفسیر متنی هستند که تلاش می‌کنند زمینه[۷] و موضوع کلمات را در قالب بردارهایی با طول ثابت نشان دهند [۱۹]. Word2vec یکی از پرکاربردترین الگوریتم‌های جاسازی کلمه است که نتایج پیشرفته‌ای را در کاربردهای متعدد پردازش زبان طبیعی ارائه کرده است[۲۰]. لو و همکاران از Word2Vec برای نمایش کلمات مجموعه داده به‌عنوان ورودی خودرمزگذار بازگشتی جهت تحلیل احساسات در سطح سند[۸] استفاده کردند. نتایج نشان داد که مدل پیشنهادی نسبت به برخی از روش‌ها بانظارت مانند ماشین بردار پشتیبان[۹] و بیز ساده[۱۰] عملکرد بهتری دارد [۲۱]. GloVe یک مدل رگرسیون لگاریتمی دوسویه است که تلاش می‌کند با ترکیب مزایای روش‌های مبتنی بر شمارش و مبتنی بر پیش‌بینی، بر محدودیت‌های روش‌های وان-هات[۱۱]، فروانی کلمه- معکوس فراوانی سند[۱۲] و Word2Vec غلبه کند [۲۲]. شارما و همکاران مدلی را با GloVe و شبکه

عصبی بازگشتی جهت تحلیل احساسات ارائه کرده‌اند، که این مدل علاوه بر طبقه‌بندی دودویی برای طبقه‌بندی چند کلاسه هم می‌تواند مورداستفاده قرار بگیرد [۲۳]. ورشنی و همکاران در [۲۴] ، یک تحلیل مقایسه‌ای از شبکه‌های LSTM، حافظه طولانی کوتاه-مدت دوطرفه[۱۳] و نمایش رمزگذار دوطرفه از ترانسفورمرها[۱۴] بر روی دو مجموعه داده انجام دادند. نتایج نشان داد که BERT ازنظر دقت، صحت، بازخوانی و امتیاز f1 از الگوریتم‌های دیگر عملکرد بهتری دارد. در این مطالعه یک مدل ترکیبی برای تجزیه‌وتحلیل احساسات پیشنهاد می‌گردد که دقت را افزایش می‌دهد. روش‌های زیادی برای ساخت مدل‌های ترکیبی وجود دارد، برای مثال دنگ و همکاران در [۲۵] عملکرد ترکیب ماشین بردار پشتیبان، CNN و LSTM را با استفاده از تکنیک‌های جاسازی، Word2vec و BERT، روی هشت مجموعه داده متنی از توییت‌ها آزمایش کردند. پس‌ازآن، مدل ترکیبی تولیدشده را با مدل‌های تک مقایسه کردند. نتایج نشان داد که قابلیت اطمینان مدل‌های ترکیبی در بین تمام مدل‌های آزمایش‌شده برای تحلیل قطبیت احساسات بهتر عمل می‌کند و ترکیب مدل‌های یادگیری عمیق با تکنیک SVM نتایج بهتری نسبت به استفاده از یک مدل فردی برای انجام تحلیل احساسات به همراه دارد. در مطالعه‌ای دیگر نویسندگان در [۲۶] یک مدل عمیق شبکه عصبی بازگشتی با شبکه عصبی کانولوشنی دوجهته مبتنی بر توجه برای تحلیل احساسات در سطح سند پیشنهاد می‌کنند. مدل پیشنهادی با استفاده از دولایه LSTM و GRU دوطرفه مستقل، هر دو زمینه گذشته و آینده را با در نظر گرفتن جریان اطلاعات زمانی در هر دو جهت استخراج می‌کند. همچنین مکانیسم توجه بر روی خروجی لایه‌های دوطرفه برای تأکید بیشتر یا کمتر بر کلمات مختلف اعمال می‌شود. نتایج نشان می‌دهد مدل پیشنهادی نسبت به کارهای بررسی شده عملکرد بهتری دارد و به نتایج پیشرفته‌ای در طبقه‌بندی توییت‌های طولانی و کوتاه دست می‌یابد. نویسندگان در [۲۷] یک مدل مبتنی بر توجه را با استفاده از شبکه‌های عصبی کانولوشنی و حافظه طولانی کوتاه‌مدت پیشنهاد می‌کنند که از وزن دهی TF-IDF و رویکرد از پیش-آموزش‌دیده GloVe برای استخراج اطلاعات معنی‌دار استفاده می‌کند. علاوه بر این، یک مکانیسم توجه در انتهای لایه‌های CNN برای جلب‌توجه بیشتر یا کمتر به کلمات مختلف به کار گرفته شده است. کارایی مدل بر روی چهار مجموعه داده استاندارد انگلیسی، با ماتریس‌های عملکرد مختلف ارزیابی شده و کارایی با مدل‌ها و رویکردهای پایه موجود مقایسه شده است. نتایج آزمایش نشان می‌دهد که روش پیشنهادی به‌طور قابل‌توجهی بهتر از مدل‌های پیشرفته عمل می‌کند. آیتوگ اونان یک معماری شبکه عصبی بازگشتی کانولوشنی دوطرفه را پیشنهاد می‌کند که از دولایه دوطرفه LSTM و GRU استفاده می‌کند تا با اتصال دولایه پنهان از جهت‌های مخالف به یک زمینه، هم زمینه‌های گذشته و هم آینده را استخراج کند. مکانیسم ارتقای گروهی بر روی ویژگی‌های استخراج‌شده توسط لایه‌های دو جهته به کار گرفته شده است، که ویژگی‌ها را به کلاس‌های متعدد تقسیم می‌کند و اهمیت ویژگی‌های مهم را در هر گروه افزایش می‌دهد درحالی‌که موارد کم‌اهمیت‌تر را تضعیف می‌کند. نتایج تجربی نشان می‌دهد که مکانیسم ارتقای گروهی می‌تواند نتایج امیدوارکننده‌ای را در مورد تخصیص مقادیر وزنی مناسب به ویژگی‌های زمینه به همراه داشته باشد و معماری پیشنهادی می‌تواند برای تحلیل احساسات بهتر عمل کند [۲۸]. محمد امان‌الله و همکاران یک الگوریتم و روش برای تجزیه‌وتحلیل احساسات با استفاده از متن و شکلک ارائه دادند. در این کار، هر دو حالت داده به‌صورت ترکیبی و جداگانه با استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق تحلیل کردند. نشان دادند که در تحلیل احساسات الگوریتم‌های یادگیری عمیق، شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت بهتر از الگوریتم‌های یادگیری ماشین عمل می‌کنند [۲۹]، بنابراین، در این تحقیق متن توییتر از مجموعه داده Sentiment140 ، [۲۸- ۲۵]، با استفاده از تکنیک‌های یادگیری عمیق شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت به همراه بردارهای تعبیه کلمات از قبل آموزش‌دیده در نظر گرفته شد تا بررسی گردد که آیا استفاده از شکلک‌ها می‌تواند دقت پیش‌بینی را بهبود دهد، درنهایت یک مدل ترکیبی مبتنی بر توجه شبکه‌های عصبی کانولوشنی با شبکه‌های حافظه طولانی کوتاه‌مدت پیشنهاد می‌گردد که عملکرد بهتری را نسبت به کارهای بررسی شده ارائه می‌کند. جدول ۱، تعدادی از مطالعات مهم در زمینه تحلیل احساسات را بیان می‌کند.

جدول ۱: تعدادی از مطالعات مهم در زمینه تجزیه و تحلیل احساسات

نویسنده و سال	الگوریتم و روش / محیط کار	هدف و یافته‌ها
ویسلاو و ولنی (۲۰۱۶)	طبقه‌بندی احساسات سطح سند، تجزیه‌وتحلیل نمادها/ twitter	تجزیه‌وتحلیل نمادها که از شکلک‌ها و کاراکترهای ایموجی استفاده می‌کند، می‌تواند به‌طور قابل‌توجهی دقت در شناخت انواع احساسات را افزایش دهد.
محمد احسان بصیری و همکاران (۲۰۲۰)	CNN BI-LSTM BI-GRU ATTENTION /MECHANISM Sentiment140 Airline dataset	ارائه یک مدل ترکیبی یادگیری عمیق برای تحلیل احساسات و نتایج نشان می‌دهد که روش پیشنهادی روی دیتاست های شامل توییت های طولانی بهتر از دیتاست با توییت های کوتاه عمل می‌کند.
محمد امان‌الله و همکاران (۲۰۲۰)	TF–IDF SVM NB RF CNN /LSTM twitter	این تحقیق نشان می‌دهد که هر زمان شکلک استفاده می‌شود، احساسات مرتبط با آن بر احساس منتقل‌شده توسط تجزیه‌وتحلیل داده‌های متنی غلبه می‌کند. همچنین، الگوریتم‌های یادگیری عمیق بهتر از الگوریتم‌های یادگیری ماشین هستند.
مرجان کامیاب و همکاران (۲۰۲۱)	TF-IDF CNN LSTM ATTENTION /MECHANISM US-airline- Sentiment140-MV SA4A	ارائه یک مدل مبتنی بر توجه جدید که از CNN با LSTM استفاده می‌کند. دقت مدل پیشنهادی نسبت به روش‌های پایه بهبود یافته است.
ماریا مورنو و همکاران (۲۰۲۱)	SVM CNN LSTM -/BERT Sentiment140 Tweets Airline	نتایج نشان می‌دهد که قابلیت اطمینان مدل‌های ترکیبی در بین تمام مدل‌های آزمایش‌شده برای تحلیل قطبیت احساسات بهتر عمل می‌کند.

	SemEval	

۳- سطوح تجزیه‌وتحلیل احساس

تحقیقات تجزیه‌وتحلیل احساسات به‌طور عمده در سه سطح انجام‌شده است:

سطح سند، سطح جمله[۱۵] و سطح جنبه[۱۶].

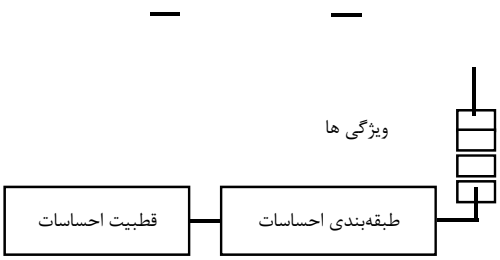
سطح سند: کار در سطح سند این است که طبقه‌بندی یک سند کامل دارای احساسات منفی است یا مثبت [۳۰]. این سطح از تجزیه‌وتحلیل به‌طور ضمنی فرض می‌کند که هر سند نظرات خود را در مورد یک موجودیت واحد (به‌عنوان‌مثال، یک محصول یا خدمت واحد) بیان می کند؛ بنابراین برای اسنادی که چندین نهاد را ارزیابی یا مقایسه می‌کنند که برای آن‌ها تجزیه‌وتحلیل دقیق‌تر لازم است، قابل‌استفاده نیست.

سطح جمله: سطح جمله تعیین این است که هر جمله نظر مثبت، منفی یا خنثی را بیان می‌کند. توجه داشته باشید که «نظر خنثی» معمولاً به معنای «عدم عقیده» است.

سطح جنبه: نه در آنالیزهای سطح سند و نه در آنالیزهای سطح جمله، چیزی که مردم دقیقاً دوست دارند و دوست ندارند، کشف نمی‌شوند. به‌عبارت‌دیگر، آن‌ها نمی‌گویند که هر نظر در مورد چیست، یعنی هدف نظر. می‌توان گفت اگر بتوانیم جمله‌ای را مثبت طبقه‌بندی کنیم، همه‌چیز در جمله می‌تواند نظر مثبت را بدست آورد. بااین‌حال، این نیز مفید نخواهد بود، زیرا یک جمله می‌تواند چندین نظر داشته باشد، به‌عنوان‌مثال، «اپل در این اقتصاد ضعیف بسیار خوب عمل می‌کند.» منطقی نیست که این جمله را مثبت یا منفی طبقه‌بندی کنیم زیرا در مورد اپل مثبت است اما در مورد اقتصاد منفی است. برای به دست آوردن این سطح از نتایج ریز، باید به سطح جنبه برویم. این سطح از تجزیه‌وتحلیل قبلاً سطح ویژگی نامیده می‌شد، اکنون تجزیه‌وتحلیل احساسات مبتنی بر جنبه نامیده می‌شود [۳۱]. هدف این سطح از تجزیه‌وتحلیل کشف احساسات موجودیت‌ها و یا جنبه‌های آن‌ها است.

۳-۱-فرایند کلی تجزیه‌وتحلیل احساسات

تجزیه و تحلیل احساسات شامل پیش‌پردازش داده‌ها[۱۷]، انتخاب ویژگی-ها[۱۸]، طبقه‌بندی و سپس پیدا کردن قطبیت داده‌ها است (شکل ۱). پیش‌پردازش داده‌ها شامل توکن‌سازی، حذف کلمات توقف و ریشه-یابی[۱۹] است. توکن‌سازی عمل شکستن دنباله‌ای از کلمات به کلمات جداگانه به نام توکن است. کلمات توقف کلماتی هستند (is، am، are، in، to و غیره) که هیچ نظری ندارند، بنابراین حذف آن‌ها می‌تواند مفید واقع شود. ریشه‌یابی، عمل تبدیل اشکال مختلف کلمه، به شکل اصلی و پایه آن مانند helping به help است.



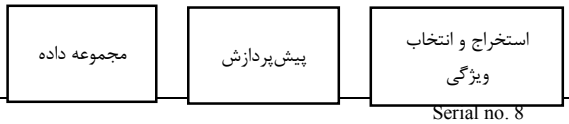
شکل ۱- فرایند تحلیل احساسات [۳۲]

۳-۲-تکنیک‌های تحلیل احساسات

روش‌های تجزیه‌وتحلیل احساسات مبتنی بر یادگیری ماشین، مبتنی بر واژه‌نامه و ترکیبی است. در روش یادگیری ماشین از مجموعه داده ویژگی‌ها را استخراج می‌کنیم و این ویژگی‌ها به طبقه‌بندی قطبیت جمله ورودی ناشناخته کمک می‌کند. روش‌های یادگیری ماشین به یادگیری تحت نظارت و یادگیری بدون نظارت تقسیم می‌شود [۳۲]. یادگیری با نظارت[۲۱]: این رویکرد زمانی مورداستفاده قرار می‌گیرد که داده‌های برچسب‌گذاری شده برای آموزش مدل وجود داشته باشد. دو مرحله در یادگیری با نظارت استفاده می‌شود: اول آموزش مدل و دیگری پیش‌بینی [۳۳]. در طول آموزش، مجموعه‌داده با برچسب‌های آن به الگوریتم طبقه‌بندی تغذیه می‌شود که یک مدل را به‌عنوان خروجی می‌دهد. پس‌ازآن داده‌های تست برای پیش‌بینی طبقه‌بندی وارد مدل می‌شوند.

یادگیری بدون نظارت[۲۲]: این روش زمانی مورداستفاده قرار می‌گیرد که قابلیت اطمینان داده‌های برچسب دار دشوار باشد. جمله بر اساس لیست کلمات کلیدی هر دسته طبقه‌بندی می‌شود. این الگوریتم‌ها الگوهای پنهان یا گروه‌بندی داده‌ها را بدون نیاز به دخالت انسان کشف می‌کنند. آنیسا[۲۳] و همکاران تجزیه‌وتحلیل احساسات را با استفاده از رویکرد بدون نظارت انجام دادند که در آن توییت‌ها با استفاده از روش خوشه‌بندی طیفی به خوشه مثبت و منفی تقسیم شدند. خوشه‌های طیفی از بیز ساده، ماشین بردار پشتیبان و حداکثر آنتروپی[۲۴] عملکرد بهتری داشتند [۳۴-۳۷].

یادگیری عمیق برای طبقه‌بندی احساسات: شبکه عصبی مصنوعی ساختار عصبی مغز انسان را تقلید می‌کند. واحد اصلی شبکه عصبی نورون است. شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی است. شبکه عصبی از تابع خطی $x(i) = A.(a(i))$ استفاده می‌کند. بردار "a (i)" به‌عنوان ورودی به نورون داده می‌شود که یک کلمه را در یک سند نشان می‌دهد. یک وزن "A"، مربوط به هر نورون وجود دارد که برای محاسبه عملکرد مورداستفاده قرار می‌گیرد. x (i) برای طبقه‌بندی کلاس استفاده می‌شود. در شبکه‌های عصبی مصنوعی، آموزش مدل شامل دو مرحله است: انتشار روبه‌جلو و انتشار عقبگرد. در انتشار روبه‌جلو، ورودی به لایه ورودی نورون‌ها داده می‌شود که در وزن‌هایی که اعداد تصادفی هستند ضرب می‌شود. توابع برای نرمال‌سازی مقدار خروجی بین ۰ و ۱ استفاده می‌شوند. سپس خروجی با مقدار موردنظر مقایسه می‌شود، اگر بین دو مقدار تفاوت (خطا) وجود داشته باشد، انتشار عقبگرد انجام می‌شود.

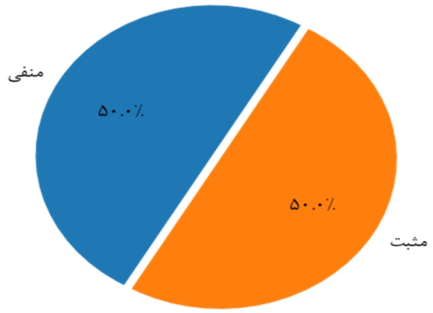


۴-پیاده‌سازی روش پیشنهادی

در این مقاله مدل‌های پایه با استفاده از توییت‌هایی که شکلک‌های آن‌ها حذف‌شده آموزش‌دیده و آزمایش‌شده‌اند. برای هر مدل پایه، یک مدل متناظر آموزش داده شد که شامل شکلک‌ها به‌عنوان ورودی بود. دقت[۲۵]، صحت[۲۶]، فراخوانی[۲۷] یا پوشش و امتیاز f1 مدل‌های پایه مربوطه که از شکلک استفاده نمی‌کنند با مدل‌های شامل شکلک‌ها مقایسه شدند.

۴-۱-مجموعه داده

استفاده از پایگاه‌های داده فارسی، دارای محدودیت‌هایی در مدت زمان پیاده‌سازی و تامین منابع دارد (که به عنوان یکی از کارهای آینده، در ادامه این تحقیق پیشنهاد می‌شود). در این مقاله، برای ارزیابی احساسات در پیام‌های توییتر، از مجموعه داده زبان انگلیسی Sentiment140 استفاده شده است. شکلک‌ها با استفاده از لیستی از شکلک‌های زبان انگلیسی ارایه شده در ویکی‌پدیا تهیه شده است، که در آن هر نماد از شکلک‌ها، با متن مناسب جایگزین شده است. برای مجموعه داده توییتری، در متن و شکلک‌ها، شش ویژگی در نظر گرفته شده است. این شش ویژگی شامل: قطبیت توییت، شناسه توییت، تاریخ توییت، پرس جوی کاربر، متن توییت و کاربری که توییت کرده است، می‌باشد.
شکل ۲، توزیع داده‌ها را برای برچسب‌های احساسات، نشان می‌دهد. هر دو برچسب احساسات مثبت و منفی، به‌طور مساوی در این مجموعه داده توزیع‌شده‌اند.



شکل ۲- توزیع برچسب احساسات مجموعه داده

در این تحقیق، فقط ردیف‌هایی که حداقل دارای یک شکلک باشند در نظر گرفته شده‌اند، علاوه بر این، دیتاست شامل: هشتگ، آدرس‌های اینترنتی، تگ‌های اچ‌تی‌ام‌ال[۲۸]، کاراکترهای خاص، کاراکترهای خط جدید (\n) و کاراکترهای انتهای خط (\t) نیز می‌باشد. در جدول ۲، چهل شکلک، به همراه تعداد وقوع آن‌ها در توییت‌ها، نشان داده شده است.

جدول ۲- شکلک‌ها با تعداد وقوع آن‌ها

شکلک	وقوع	شکلک	وقوع	شکلک	وقوع
;)	4790	:')	199	:-*	78
xp	5949	:$	1506	;D	1619
XP	203	:o	648	x3	1540
:3	1627	=]	469	X3	102
xD	1771	0:3	314	=3	48
XD	1830	8)	308	;3	13339
:p	1732	:*	281	:b	81
:O	1606	;]	170	=D	837
;-)	1101	=p	160	:L	667
:]	1019	:o)	83	=/	520
=\\	53	:-/	402	:[	211
:@	417	DX	215	;(	343
:'(	1411	:/	14646	:|	1083
:S	840				

ستونی از متن که شامل شکلک می باشد، با استفاده از لیستی از شکلک‌های زبان انگلیسی ذکرشده در ویکی‌پدیا تهیه‌شده، که در آن هر نماد شکلک با متن مناسب جایگزین شده است. به‌عنوان‌مثال، "[:" و "]=" با happy face و ");" با sad جایگزین شده‌اند. بیش از ۹۰٪ توییت ها یک یا دو شکلک داشتند و در تعداد کمی از توییت‌ها بیش از دو شکلک وجود داشت. تعداد ردیف‌های هر دو مجموعه‌داده، بر اساس قطبیت مثبت و منفی در جدول ۳ نشان داده‌شده است.

جدول ۳- تعداد ردیف‌ها در دیتاست پایه و دیتاست شامل شکلک

	مثبت	منفی
مدل با شکلک	۳۲۱۰۹	۳۲۱۰۹
مدل پایه	۳۲۱۰۹	۳۲۱۰۹

۴-۲- پیش‌پردازش داده

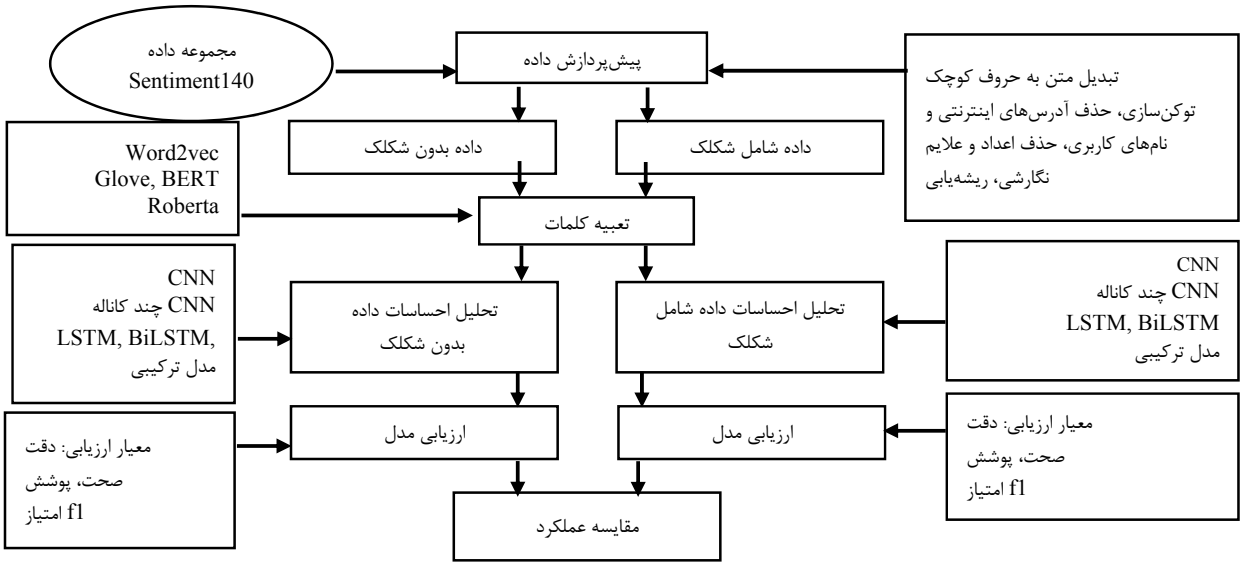
توییت‌های خام حاوی اطلاعات بی‌معنی و بدون ساختار و کلمات تکراری بودند. ازاین‌رو، هر آزمایش چند مرحله پیش‌پردازش را دنبال می‌کند. مجموعه داده‌ها مراحل پیش‌پردازش زیر را طی می‌کنند:

۱) تبدیل متن به حروف کوچک
۲) حذف تمام فضاهای خالی اضافی
۳) حذف همه اعداد و علائم نگارشی
۴) حذف آدرس‌های اینترنتی، تگ‌های html و نام‌های کاربری

برای مدیریت کلمات بن در ابتدا هر توییت را با استفاده از ابزار توکن‌ساز مجموعه ابزار زبان طبیعی[۲۹] به لیستی از کلمات تقسیم می‌شود. برای ریشه یابی از Stemming استفاده می‌گردد. اگر هر یک از توکن‌ها به کوله کلمات NLTK تعلق داشت، کلمه مناسب از بن NLTK برای آن توکن جایگزین می‌شود. عمل ریشه‌یابی این امکان را می‌دهد که فرم‌های مختلف یک کلمه به یک فرم واحد تبدیل گردد. با این کار تعداد ویژگی‌ها کمتر می‌شود و کامپیوتر می‌تواند شکل‌های مختلف یک کلمه را یکی در نظر بگیرد. همچنین کلماتی که آپاستروف می‌گیرند (I'm,It's و غیره)، آپاستروف آن‌ها حذف شدند، چون توکن‌ساز[۳۰] آن‌ها را به‌عنوان یک توکن جداگانه در نظر می‌گیرد.

۴-۳-معماری کلی مدل

در مدل ارایه شده، در ابتدا مجموعه داده sentiment140 پیش‌پردازش می‌شود، سپس بردارهای تعبیه از قبل آموزش داده‌شده Word2vec، glove، bert و Roberta روی مجموعه داده اعمال می‌شود. در ادامه داده‌ها به الگوریتم‌های شبکه عصبی کانولوشنی، شبکه عصبی کانولوشنی چندکاناله، حافظه طولانی کوتاه‌مدت، حافظه طولانی کوتاه-مدت دوجهته و مدل ترکیبی تزریق می‌گردد و درنهایت دقت هر مدل با استفاده از معیارهای ارزیابی اندازه‌گیری می‌شود. مراحل کلی مدل پیشنهادی در شکل ۳ نشان داده‌شده است.



شکل ۳- معماری کلی مدل

۴-۴-معیار ارزیابی

برای ارزیابی طبقه‌بندی احساسات، عملکرد مدل با استفاده از دقت، صحت، بازخوانی یا پوشش و امتیاز f1 اندازه‌گیری شد. دقت، معیار اندازه-گیری تمام کلاس‌های مشخص‌شده در برابر کلاس واقعی است. صحت کسری از نمونه مربوطه در برابر نمونه پیش‌بینی‌شده است. پوشش کسری از نمونه‌های مطلوب است که روی کل مقادیر نمونه‌های مطلوب پیش‌بینی‌شده است. امتیاز f1 به‌عنوان میانگین وزنی پوشش و صحت تعریف می‌شود. معادلات زیر دقت، فراخوانی و صحت و امتیاز F1 را توصیف می‌کنند:

$$Acc_i = \frac{TPi+TNi}{TPi+TNi+FPi+FNi} \quad (۱)$$

$$Rec_i = \frac{TPi}{TPi+FNi} \quad (۲)$$

$$Pre_i = \frac{TPi}{TPi+FPi} \quad (۳)$$

$$F1 = \frac{2\times Pre\times Rec}{Pre+Rec} \quad (۴)$$

که Acc_i، Rec_i و Pre_i به ترتیب دقت، فرخوانی و صحت برای کلاس i هستند. N تعداد کلاس‌ها در طبقه‌بندی احساسات است. TP مثبت واقعی، FP مثبت کاذب، FN منفی کاذب و TN منفی واقعی است. TP_i تعداد نمونه‌هایی در کلاس i هستند که پیش‌بینی شده‌اند در کلاس i هستند. TN_i تعداد نمونه‌هایی که در کلاس i نیستند و پیش‌بینی هم شده‌اند که در کلاس i نیستند. FP_i تعداد نمونه‌هایی که در کلاس i نیستند و پیش‌بینی شده‌اند که در کلاس i هستند. FN_i تعداد نمونه‌هایی که در کلاس i هستند و پیش‌بینی شده‌اند که در کلاس i نیستند.

لایه جاسازی یا تعبیه: قبل از ساختن مدل تجزیه‌وتحلیل احساسات، همه ویژگی‌های موجود در توییت‌ها باید استخراج‌شده و از طریق تعبیه کلمه در گروهی از کلمات مرتبط بامعنا شکل بگیرند. این مطالعه از چند روش مختلف تعبیه کلمات، ازجمله بردار تعبیه word2vec از قبل آموزش‌دیده، بردار تعبیه‌شده GloVe از قبل آموزش‌دیده، بردار تعبیه‌شده BERT از قبل آموزش داده‌شده و بردار تعبیه‌شده roberta از قبل آموزش داده‌شده استفاده می‌کند.

word2vec از قبل آموزش داده‌شده: برای طبقه‌بندی احساسات از بردار تعبیه word2vec از قبل آموزش‌دیده گوگل استفاده می‌گردد. دارای ۳۰۰ بعد بردار کلمات از قبل آموزش‌دیده است که با استفاده از ۱۰۰ میلیارد کلمه از یک مجموعه داده اخبار گوگل آموزش داده‌شده است. این بردار دارای اندازه واژگان حدود ۳ میلیون کلمه از اخبار گوگل است و به یک فرهنگ لغت با کلمات به‌عنوان کلید و ضرایب به‌عنوان مقادیر برای لایه تعبیه تبدیل‌شده است.

GloVe از قبل آموزش داده‌شده: برای طبقه‌بندی احساسات از بردار تعبیه GloVe از قبل آموزش‌دیده استنفورد استفاده می‌گردد. ۲۰۰ بعد از بردارهای کلمه از قبل آموزش‌دیده است که با استفاده از ۲ میلیارد توییت، ۲۷ میلیارد کلمه ساخته‌شده است و اندازه واژگان آن ۱,۲ میلیون است. این بردار به یک فرهنگ لغت با کلمات به‌عنوان کلید و ضرایب به‌عنوان مقادیر تبدیل‌شده است.

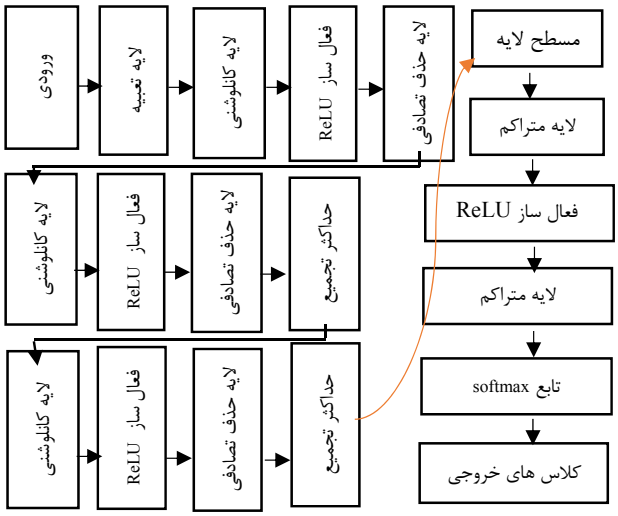
نمایش رمزگذار دوطرفه از ترانسفورمرها: نمایش رمزگذار دوطرفه از ترانسفورمرها توسط دولین[۳۱] و همکاران (۲۰۱۸) مطرح شد که از رمزگذارها در یک ترانسفورمر به‌عنوان زیرساختی برای مدل‌های از قبل آموزش‌دیده برای کارهای پردازش زبان طبیعی مانند تحلیل احساسات، خلاصه‌سازی متن و غیره استفاده می‌کند [۳۸]. توانایی BERT برای وارد کردن دو جمله به‌طور هم‌زمان و تعیین اینکه آیا جمله دوم بعد از جمله اول می‌آید یا خیر، برای پیش‌بینی جمله بعد[۳۲] (NSP) بسیار پرکاربرد می باشد. این توانایی به مدل کمک می‌کند تا روابط طولانی بین متون را حفظ کند. BERT دو مدل دارد، یعنی مدل‌های BERT-base و BERT-large. مدل BERT-base شامل بلوک‌های رمزگذار ترانسفورمر ۱۲ لایه است که هر بلوک حاوی ۱۲ لایه خود توجه[۳۳] و ۷۶۸ لایه پنهان است و درمجموع ۱۱۰ میلیون پارامتر تولید می‌کند. از سوی دیگر، BERT-large از بلوک‌های رمزگذار ترانسفورمر ۲۴ لایه تشکیل شده است که هر بلوک حاوی ۲۴ لایه خود توجه است و درمجموع ۳۴۰ میلیون پارامتر تولید می‌کند. عملکرد BERT به نوع مدل بستگی دارد، BERT-large می‌تواند دقت بالاتری نسبت به BERT-base داشته باشد. بااین‌حال، این بهبود در دقت با استفاده از BERT-large به منابع بیشتری نیاز دارد. مزیت BERT شامل توانایی آن در مدیریت استخراج اطلاعات متنی به دلیل توانایی دو جهته آن است. سریع‌تر آموزش می‌بیند و در طیف گسترده‌ای از برنامه‌های مدل زبانی استفاده شده است. در این تحقیق از BERT-base استفاده می‌گردد.

رویکرد BERT از قبل آموزش‌دیده بهینه‌شده قوی: مدل RoBERTa توسعه‌ای از BERT است. BERT و RoBERTa تحت خانواده ترانسفورمرها قرار می‌گیرند که برای رسیدگی به مشکل وابستگی‌های بلندمدت توسعه یافته است [۳۹]. RoBERTa همچنین روی مجموعه داده‌های بزرگ‌تری نسبت به BERT آموزش‌دیده است. این امر باعث می‌شود تا نمایش‌های RoBERTa در مقایسه با BERT بهتر تعمیم پیدا کنند. در این تحقیق از roberta-large-mnli استفاده می‌گردد.

در مرحله بعدی، لایه تعبیه از طریق چهار بردار تعبیه‌شده فوق آماده گردید. از تابع توکن‌ساز کراس استفاده شد که می‌تواند روی توییت‌های مجموعه‌داده آموزشی برازش شود که متن را به رشته تبدیل می‌کند. این بردار نمایش کلمات به‌عنوان داده ورودی برای مدل‌های CNN و LSTM استفاده می‌شود که به‌نوبه خود برای طبقه‌بندی احساسات بکار برده می‌شود.

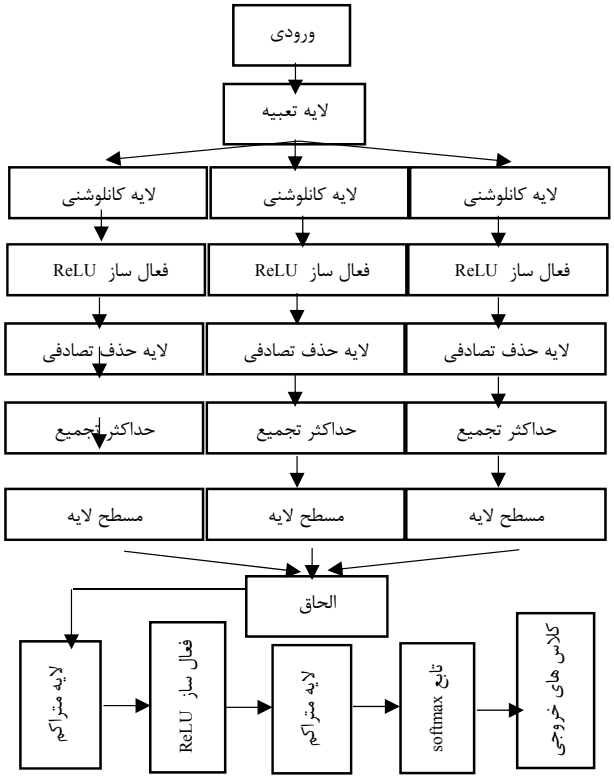
۴-۵-مدل شبکه عصبی کانولوشنی

با چند تغییر در مدل شبکه عصبی کانولوشنی کیم و با تنظیم پارامترها دو نوع CNN، ایجاد می‌گردد: یک مدل ساده CNN و یک مدل چندکاناله CNN. یک مدل ترتیبی CNN ، دارای یک پشته خطی از لایه‌ها است که در شکل ۴ نشان داده‌شده است. حداکثر طول توییت ها ۴۶ بود، حداکثر طول در این دو مجموعه داده توسط توکن‌ساز کراس کنترل می‌شود. توکن‌ساز توییت‌ها را بردار می‌کند و آن‌ها را به دنباله‌ای از اعداد صحیح تبدیل می‌کند، سپس توکن‌ساز برای استفاده از رایج‌ترین کلمات محدود می‌شود. عمل پد گذاری با اضافه کردن مقدار ۰ در انتها، طول رشته‌های متنی را برابر با حداکثر طول یعنی طول بزرگ‌ترین رشته متنی می‌رساند. اولین لایه در این CNN یک لایه تعبیه[۳۴] است که در قسمت لایه تعبیه قبل ساختیم. لایه تعبیه، خروجی را به یک لایه کانولوشنی[۳۵]منتقل می‌کند. این لایه ماتریس خروجی را از طریق چند پارامتر مانند تعداد فیلترهای خروجی در کانولوشن‌ها، طول پنجره‌های کانولوشنی، تابع فعال‌سازی و تابع زیان از طریق روش میزان‌سازی[۳۶] تولید می‌کند. از فیلترهایی استفاده می‌شود که در محدوده ۳۲-۸، اندازه هسته ۵-۳، تابع فعال‌سازی 'relu' یا 'tanh' و میزان‌سازی با وزن ۰/۰۰۰۲ هستند. خروجی لایه کانولوشنی به یک لایه حذف تصادفی با نرخ حذف ۰/۵ وارد می‌شود و به‌طور تصادفی به ۵۰ درصد نورون‌های شبکه وزن ۰ اختصاص می دهد. این عملیات حذف تصادفی شبکه باعث می‌شود حساسیت شبکه در هنگام واکنش به تغییرات کوچک‌تر در توییت کمتر شود که باعث افزایش دقت مدل می‌شود. سپس خروجی لایه حذف تصادفی به یک لایه کانولوشنی دوم وارد می‌شود.



شکل ۴ - مدل شبکه عصبی کانولوشنی ساده

سپس این لایه کانولوشنی به یک لایه حذف تصادفی با نرخ حذف ۰/۳ منتقل شد. به همین طریق، لایه حذف تصادفی به لایه حداکثر تجمیع و لایه حداکثر تجمیع به لایه کانولوشنی سوم منتقل شد. این لایه کانولوشنی سوم به یک لایه حذف تصادفی با نرخ حذف تصادفی ۰/۱ و این لایه حذف تصادفی به یک لایه حداکثر تجمیع منتقل شد. شکل ۵ مدل CNN چندکاناله را که در این مقاله استفاده شده است به تصویر می‌کشد.



شکل ۵- شبکه عصبی کانولوشنی چندکاناله

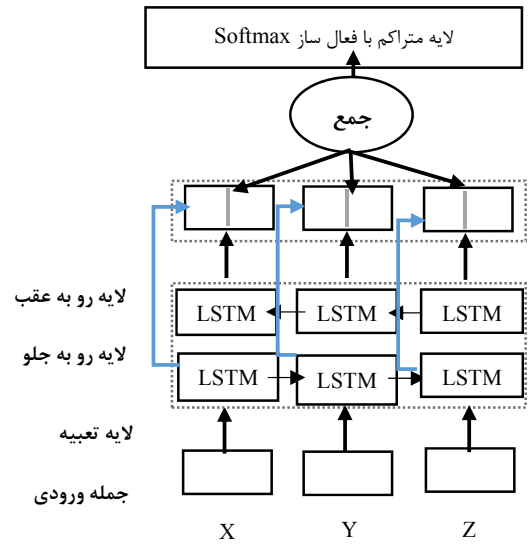
این مدل دارای سه کانال است. هر کانال دارای یک لایه ورودی است. لایه ورودی به یک لایه تعبیه، یک لایه یک‌بعدی کانولوشنی، یک لایه حذف تصادفی، یک لایه حداکثر تجمیع و یک لایه Flatten ارسال می‌شود. بردارهای خروجی لایه Flatten هر سه کانال به هم متصل شده و به یک لایه متراکم با تابع فعال‌سازی 'relu' یا 'tanh' تغذیه می‌شود. خروجی لایه فعال‌سازی به یک لایه متراکم با تابع فعال‌سازی softmax منتقل می‌شود. این دو مدل CNN با استفاده از تابع کراس categorical_crossentropy به‌عنوان تابع هزینه و "adam" به‌عنوان تابع بهینه‌ساز از طریق تابع کامپایل کراس کامپایل شدند. انواع مدل‌های ساده و چندکاناله CNN در جدول ۴ نشان داده‌شده است.

جدول ۴- مدل‌های مختلف شبکه عصبی کانولوشنی

مدل	Word2vec	GloVe	BERT
Cnn ساده	✓	✓	✓
Cnn چندکاناله	✓	✓	✓

۴-۶-مدل حافظه طولانی کوتاه‌مدت

این تحقیق از یک معماری برای هر دو مدل حافظه طولانی کوتاه‌مدت و حافظه طولانی کوتاه‌مدت دو جهته استفاده می‌کند که از معماری مدل BiLSTM کلیشی[۳۸] پیروی می‌کند که در شکل ۶ نشان داده‌شده است. این مدل دو LSTM برای آموزش در توییت های ورودی ایجاد می‌کند. طول توییت‌های مجموعه‌داده آموزشی و تست برای این مدل LSTM در همه توییت‌ها نرمالایز شد. نرمال‌سازی متن ورودی مشابه مدل CNN بود که در قسمت قبل توضیح داده شد. شبکه LSTM یک پشته خطی از لایه‌ها است که با استفاده از کراس ساخته شده است. لایه اول یک لایه تعبیه است که رشته ورودی را به رشته‌ای از بردارهای عددی کد می‌کند. خروجی لایه تعبیه به یک لایه LSTM متصل است. LSTM رشته بردارها را بر اساس پارامتر اندازه بعد خروجی که حاوی اطلاعات مربوط به کل رشته است، به یک بردار واحد تبدیل می‌کند. لایه LSTM دوطرفه از طریق کراس پیاده‌سازی شد. خروجی لایه LSTM به یک لایه متراکم منتقل‌شده است و تابع فعال‌سازی این لایه "relu" یا "tanh" است. خروجی این لایه به یک لایه متراکم می‌رود که از softmax به‌عنوان تابع فعال‌سازی استفاده می‌کند، که باعث ایجاد برچسب احساسات شد.



شکل ۶- مدل BiLSTM

هر دو آزمایش با مدل LSTM و مدل BiLSTM، همراه با چهار مدل مختلف تعبیه که در جدول ۵ نشان داده‌شده است، مورد بررسی قرار گرفت.

جدول ۵- مدل‌های حافظه طولانی کوتاه‌مدت

مدل	Word2vec	GloVe	BERT	roberta
LSTM	✓	✓	✓	✓
BiLSTM	✓	✓	✓	✓

۴-۷-مدل ترکیبی مبتنی بر توجه شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت دوجهته (۴۸ACBR)

معماری پیشنهادی از هفت بخش اصلی تشکیل شده است، که شامل پیش‌پردازش داده، جاسازی کلمه، لایه کانولوشنی، لایه تجمیع، لایه‌های شبکه حافظه طولانی کوتاه‌مدت دوطرفه، و دو لایه دیگر می‌باشد (شکل ۷). در لایه جاسازی برای بدست آوردن نمایش برداری نظرات متنی و ایجاد بردار ویژگی‌ها RoBERTa مورداستفاده قرار می‌گیرد. هدف RoBERTa ایجاد یک جاسازی کلمه معنی‌دار به‌عنوان نمایش ویژگی است، به‌طوری‌که لایه‌های بعدی به‌راحتی می‌توانند اطلاعات مفید را از جاسازی کلمه دریافت کنند. خروجی لایه جاسازی به‌عنوان ورودی به لایه کانولوشنی وارد می‌شود که یک نگاشت ویژگی ایجاد می‌کند که حضور ویژگی‌های شناخته شده در ورودی را خلاصه می‌کند و ویژگی-های ذاتی و محلی را بدست میاورد. سپس لایه تجمیع جهت کاهش ابعاد نگاشت ویژگی مورداستفاده قرار می‌گیرد. همچنین یک مکانیسم توجه بر روی خروجی لایه‌های BILSTM برای تأکید بیشتر بر کلمات مهم‌تر و تأکید کمتر بر کلمات کم‌اهمیت‌تر اعمال می‌شود [۴۰]. خروجی لایه توجه به یک لایه متراکم منتقل می‌شود و تابع فعال‌سازی این لایه "relu" است. درنهایت یک لایه متراکم مورداستفاده قرار می‌گیرد که از softmax به‌عنوان تابع فعال‌سازی استفاده می‌کند؛ که باعث ایجاد برچسب احساسات می‌شود. مدل پیشنهادی در شکل ۷ نشان داده‌شده است. در توییت بعضی از کلمات ممکن است عبارات خاص را نادرست طبقه‌بندی کند، برای این منظور یک لایه حذف تصادفی جهت جلوگیری از پیش برازش به شبکه اضافه شد، که خروجی لایه کانولوشن را دریافت می‌کند و کلمات را تصادفی خارج می‌کند. همچنین یک مدل CNN چند کاناله شامل استفاده از چندین ورژن از مدل CNN با اندازه هسته‌های مختلف ایجاد می‌گردد. لایه حذف تصادفی محبوب‌ترین رویکرد برای میزان‌سازی CNN و کاهش مسائل پیش برازش است. این لایه از انطباق همزمان سلول‌های عصبی جلوگیری کرده و آن‌ها را مجبور به یادگیری ویژگی‌های ارزشمند می‌کند. این مدل CNN از "adagrad" یا "adam" به‌عنوان بهینه‌ساز استفاده می‌کند.

۴-۸- نتایج مدل‌های CNN

مدل‌های ساده و چندکاناله CNN برای طبقه‌بندی احساسات مورد آزمایش قرار گرفتند. این مدل‌ها با استفاده از سه بردار تعبیه مختلف برای مجموعه داده‌های مدل پایه و مدل شامل شکلک آموزش و ارزیابی شدند. در مدل‌های CNN بر اساس زمان اجرا و عملکرد مدل در هر مجموعه داده چند پارامتر تنظیم گردید.

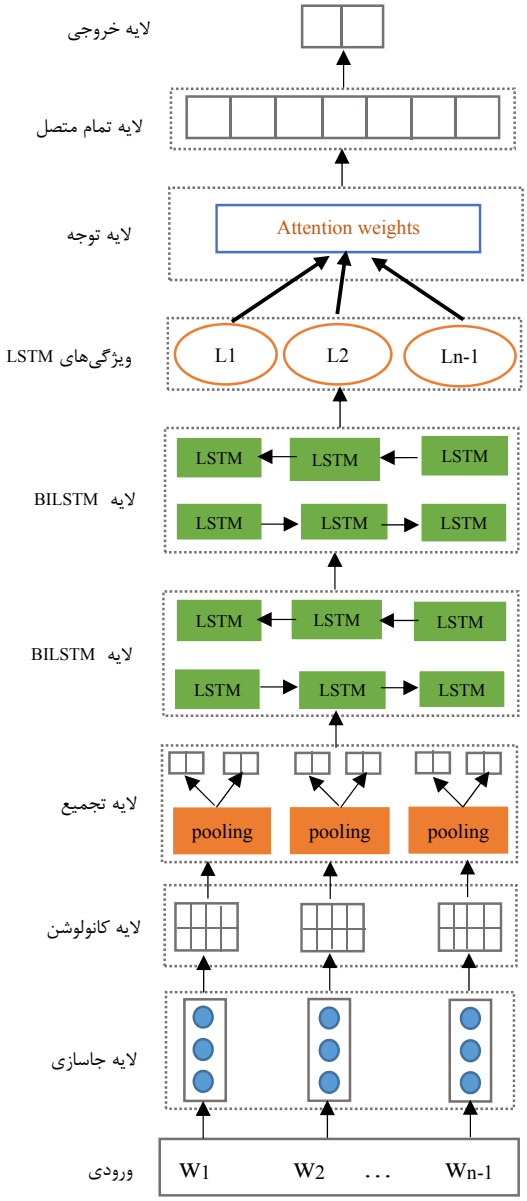
۴-۸-۱- CNN با word2vec، Glove و BERT از قبل آموزش‌دیده

سه مدل CNN پیاده‌سازی شد که در هر کدام یک لایه تعبیه با استفاده از یک مورد از بردارهای word2vec، Glove و BERT از قبل آموزش داده‌شده ایجاد گردید. این مدل‌ها با استفاده از هر دو مجموعه‌داده آموزشی آموزش داده شدند. سپس دقت پیش‌بینی مجموعه داده تست برای هر دو مجموعه داده اندازه‌گیری شد. جدول ۶ دقت هر کدام از مدل‌ها را نشان می‌دهد.

جدول ۶- دقت مدل‌های CNN روی مجموعه‌داده پایه و شامل شکلک

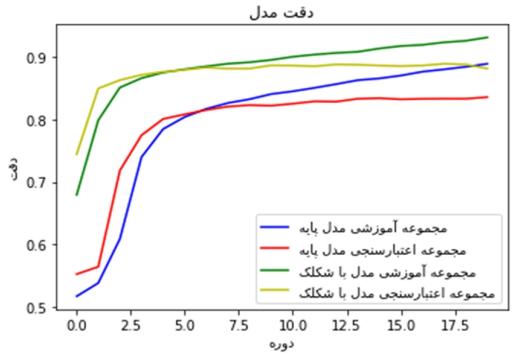
مدل	دقت (Accuracy) (٪)	
	مدل پایه	مدل با شکلک
word2vec + CNN	۷۸/۹۶	۸۴/۸۷
GloVe + CNN	۸۰/۱۶	۸۵/۶۶
BERT + CNN	۸۲/۹۲	۸۷/۶۶

نتایج نشان می‌دهد مدل CNN با بردار تعبیه BERT از قبل آموزش داده‌شده نسبت به دو مدل دیگر عملکرد بهتری دارد.



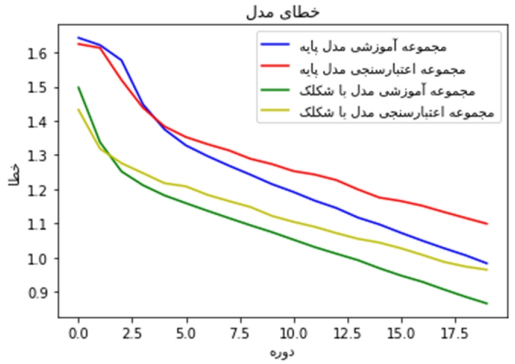
شکل ۷- مدل ترکیبی مبتنی بر توجه شبکه‌های عصبی کانولوشنی و شبکه‌های حافظه طولانی کوتاه‌مدت دوجهته

نتایج نشان می‌دهد مدل CNN با بردار تعبیه BERT از قبل آموزش داده‌شده نسبت به دو مدل دیگر عملکرد بهتری دارد. شکل ۸ و شکل ۹ دقت و خطا را برای این مدل نشان می‌دهد. دقت برای مدل پایه و مدل با شکلک به ترتیب ۸۲/۹۲ و ۸۷/۶۶ درصد است. جدول ۷ و جدول ۸ صحت، پوشش و امتیاز f1 را برای مجموعه داده تست مدل پایه و مدل شامل شکلک شرح می‌دهد.



شکل ۸ – CNN با BERT از قبل آموزش‌دیده – دقت مدل

شکل ۸، همگرایی مدل را نشان می‌دهد که بر اساس آزمایش‌های انجام‌شده بر روی مجموعه داده‌های آموزشی با تکرار ۲۰ دور بدست آمده است. دقت بدست آمده ۸۷/۸۰ درصد می‌باشد که برای تکرارهای بعدی نیز تقریباً ثابت می‌ماند.



شکل ۹ – CNN با BERT از قبل آموزش‌دیده – خطای مدل

جدول ۷ – CNN با BERT از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل پایه)

	مدل پایه			
	صحت	پوشش	معیار f1	پشتیبانی
۰	۸۵	۸۲	۸۳	۶۶۷۴
۱	۸۱	۸۴	۸۲	۶۱۷۰
دقت			۸۳	۱۲۸۴۴
میانگین کلان	۸۳	۸۳	۸۳	۱۲۸۴۴
میانگین وزنی	۸۳	۸۳	۸۳	۱۲۸۴۴

جدول ۸ – CNN با BERT از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل شامل شکلک)

مدل با شکلک				
پشتیبانی	معیار f1	پوشش	صحت	
۶۶۲۲	۸۸	۸۷	۸۹	۰
۶۲۲۱	۸۷	۸۸	۸۶	۱
۱۲۸۴۴	۸۸			دقت
۱۲۸۴۴	۸۸	۸۸	۸۸	میانگین کلان
۱۲۸۴۴	۸۸	۸۸	۸۸	میانگین وزنی

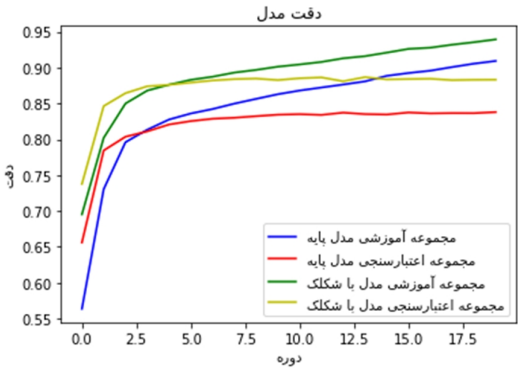
۴-۸-۲-چندکاناله با word2vec، Glove و BERT از قبل آموزش‌دیده

سه مدل CNN چندکاناله پیاده‌سازی شد که در هر کدام یک لایه تعبیه با استفاده از یک مورد از بردارهای word2vec، Glove و BERT از قبل آموزش داده‌شده ایجاد گردید. این مدل‌ها با استفاده از هر دو مجموعه داده آموزشی آموزش داده شدند سپس دقت پیش‌بینی مجموعه داده تست برای هر دو مجموعه داده اندازه‌گیری شد. جدول ۹ دقت هر کدام از مدل‌ها را نشان می‌دهد.

جدول ۹- دقت مدل‌های CNN چندکاناله روی مجموعه‌داده پایه و مجموعه‌داده شامل شکلک

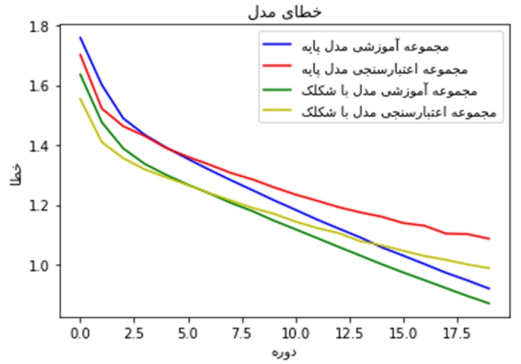
مدل	دقت (Accuracy) (٪)	
	مدل پایه	مدل با شکلک
CNN چندکاناله + word2vec	۷۹/۷۲	۸۵/۰۳
CNN چندکاناله + GloVe	۸۱/۰۸	۸۷/۴۲
CNN چندکاناله + BERT	۸۳/۰۶	۸۸/۰۳

نتایج نشان می‌دهد مدل CNN چندکاناله با بردار تعبیه BERT از قبل آموزش داده‌شده نسبت به دو مدل دیگر عملکرد بهتری دارد. شکل ۱۰ و شکل ۱۱ دقت و خطا را برای این مدل نشان می‌دهد. دقت برای مدل پایه و مدل شامل شکلک به ترتیب ۸۳/۰۶ و ۸۸/۰۳ درصد است. جدول ۱۰ و جدول ۱۱ صحت، پوشش و امتیاز f1 را برای مجموعه داده تست مدل پایه و مدل شامل شکلک نشان می‌دهند.



شکل ۱۰ – CNN چندکاناله با BERT از قبل آموزش‌دیده – دقت مدل

شکل ۱۰، همگرایی مدل را نشان می‌دهد که بر اساس آزمایش‌های انجام‌شده بر روی مجموعه داده‌های آموزشی با تکرار ۲۰ دور بدست آمده است. دقت بدست آمده ۸۸/۲۰ درصد می‌باشد که برای تکرارهای بعدی نیز تقریباً ثابت می‌ماند.



شکل ۱۱ – CNN چندکاناله با BERT از قبل آموزش‌دیده – خطای مدل

جدول ۱۰ – CNN چندکاناله با BERT از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل پایه)

مدل پایه				
پشتیبانی	معیار f1	پوشش	صحت	
۶۶۱۶	۸۳	۸۳	۸۴	۰
۶۲۲۸	۸۳	۸۴	۸۲	۱
۱۲۸۴۴	۸۳			دقت
۱۲۸۴۴	۸۳	۸۳	۸۳	میانگین کلان
۱۲۸۴۴	۸۳	۸۳	۸۳	میانگین وزنی

جدول ۱۱ – CNN چندکاناله با BERT از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل شامل شکلک)

مدل با شکلک

پشتیبانی	معیار f1	پوشش	صحت	
۶۷۹۰	۸۸	۸۶	۹۰	۰
۶۰۵۴	۸۸	۹۰	۸۶	۱
۱۲۸۴۴	۸۸			دقت
۱۲۸۴۴	۸۸	۸۸	۸۸	میانگین کلان
۱۲۸۴۴	۸۸	۸۸	۸۸	میانگین وزنی

۵-نتایج مدل‌های LSTM

یک مدل LSTM و یک مدل LSTM دوطرفه (BiLSTM) برای طبقه-بندی احساسات ایجاد گردید. هر دو مدل آموزش داده‌شده و با استفاده از چهار بردار تعبیه‌شده برای مجموعه داده پایه و مجموعه داده شامل شکلک پیش‌بینی شده‌اند.

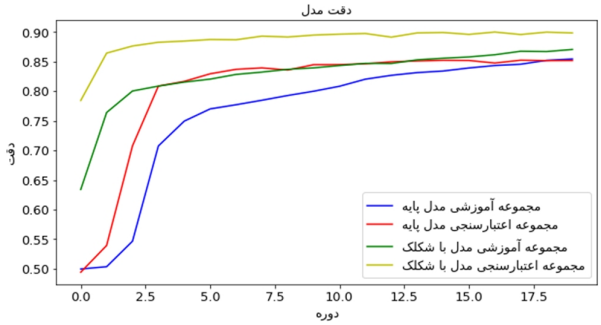
۵-۱-LSTM با word2vec، Glove، BERT و RoBERTa از قبل آموزش‌دیده

چهار مدل LSTM پیاده‌سازی شد که در هر کدام یک لایه تعبیه با استفاده از یک مورد از بردارهای word2vec، Glove، BERT و RoBERTa از قبل آموزش داده‌شده ایجاد گردید. این مدل‌ها با استفاده از هر دو مجموعه داده آموزشی آموزش داده شدند سپس دقت پیش‌بینی مجموعه داده تست برای هر دو مجموعه داده اندازه‌گیری شد. جدول ۱۲ دقت هر کدام از مدل‌ها را نشان می‌دهد.

جدول ۱۲- دقت مدل‌های LSTM روی مجموعه‌داده پایه و مجموعه-داده شامل شکلک

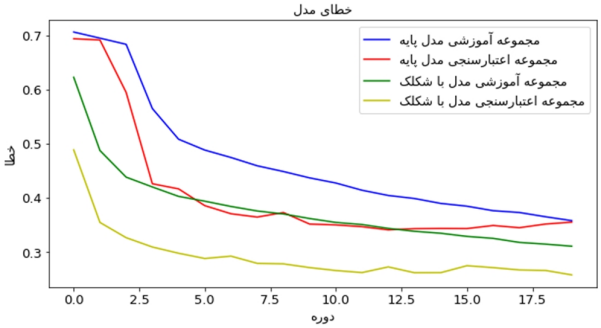
مدل	دقت (Accuracy) (٪)	
	مدل پایه	مدل با شکلک
word2vec + LSTM	۸۰/۴۹	۸۶/۹۰
GloVe + LSTM	۸۰/۹۱	۸۷/۱۶
BERT + LSTM	۸۲/۷۰	۸۸/۵۰
RoBERTa + LSTM	۸۵/۳۰	۸۹/۳۷

نتایج نشان می‌دهد مدل LSTM با بردار تعبیه RoBERTa از قبل آموزش داده‌شده نسبت به سه مدل دیگر عملکرد بهتری دارد. شکل ۱۲ و شکل ۱۳دقت و خطا را برای این مدل نشان می‌دهد. دقت برای مدل پایه و مدل شامل شکلک به ترتیب ۸۵/۳۰ و ۸۹/۳۷ درصد است. جدول ۱۳و جدول ۱۴ صحت، پوشش و امتیاز f1 را برای مجموعه داده تست مدل پایه و مدل شامل شکلک شرح می‌دهد.



شکل ۱۲ – LSTM با RoBERTa از قبل آموزش‌دیده – دقت مدل

شکل ۱۲، همگرایی مدل را نشان می‌دهد که بر اساس آزمایش‌های انجام‌شده بر روی مجموعه داده‌های آموزشی با تکرار ۲۰ دور بدست آمده است. دقت بدست آمده ۸۹/۴۸ درصد می‌باشد که برای تکرارهای بعدی نیز تقریباً ثابت می‌ماند. شکل ۱۳ خطای مدل را نشان می‌دهد که که در تکرار۱۴ به ۲۶/۲۳ درصد کاهش می‌یابد.



شکل۱۳ – LSTM با RoBERTa از قبل آموزش‌دیده – خطای مدل

جدول ۱۳ – LSTM با RoBERTa از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل پایه)

مدل پایه				
پشتیبانی	معیار f1	پوشش	صحت	
۶۶۸۲	۸۶	۸۵	۸۷	۰
۶۱۸۲	۸۵	۸۶	۸۴	۱
۱۲۸۴۴	۸۵			دقت
۱۲۸۴۴	۸۵	۸۵	۸۵	میانگین کلان
۱۲۸۴۴	۸۵	۸۵	۸۵	میانگین وزنی

جدول ۱۴ - LSTM با RoBERTa از قبل آموزش‌دیده -صحت، پوشش، امتیاز f1 (مدل شامل شکلک)

مدل با شکلک				
پشتیبانی	معیار f1	پوشش	صحت	
۶۷۴۶	۹۰	۸۸	۹۲	۰
۶۰۸۰	۸۹	۹۱	۸۷	۱
۱۲۸۴۴	۸۹			دقت
۱۲۸۴۴	۸۹	۸۹	۸۹	میانگین کلان
۱۲۸۴۴	۸۹	۸۹	۸۹	میانگین وزنی

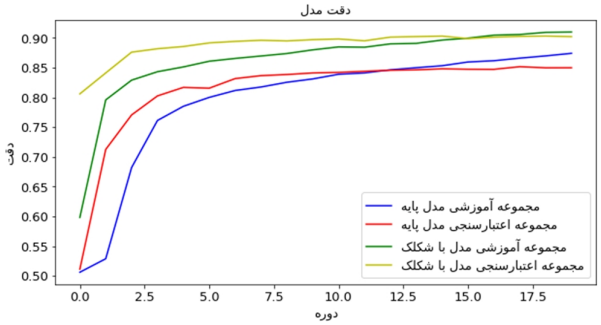
۵-۲- BiLSTM با word2vec، Glove، BERT و RoBERTa از قبل آموزش‌دیده

چهار مدل BiLSTM پیاده‌سازی شد که در هر کدام یک لایه تعبیه با استفاده از یک مورد از بردارهای word2vec، Glove، BERT و RoBERTa از قبل آموزش داده‌شده ایجاد گردید. این مدل‌ها با استفاده از هر دو مجموعه داده آموزشی آموزش داده شدند، سپس دقت پیش-بینی مجموعه داده تست برای هر دو مجموعه داده اندازه‌گیری شد. جدول ۱۵ دقت هر کدام از مدل‌ها را نشان می‌دهد.

جدول ۱۵- دقت مدل‌های BiLSTM روی مجموعه‌داده پایه و مجموعه‌داده شامل شکلک

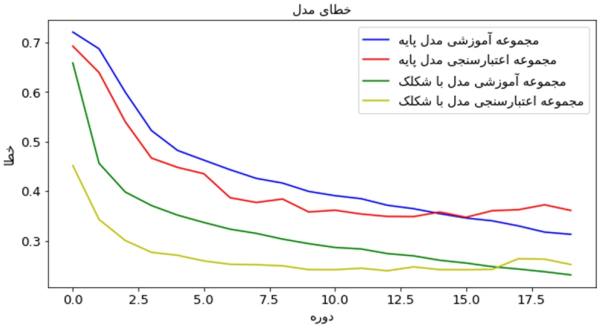
مدل	دقت (Accuracy) (٪)	
	مدل پایه	مدل با شکلک
word2vec + BiLSTM	۸۱/۲۱	۸۷/۲۴
GloVe + BiLSTM	۸۱/۲۸	۸۷/۰۱
BERT + BiLSTM	۸۳/۳۲	۸۸/۵۳
RoBERTa + BiLSTM	۸۵/۳۷	۸۹/۵۹

نتایج نشان می‌دهد مدل BiLSTM با بردار تعبیه RoBERTa از قبل آموزش داده‌شده نسبت به سه مدل دیگر عملکرد بهتری دارد. شکل ۱۴ و شکل ۱۵ دقت و خطا را برای این مدل نشان می‌دهد. دقت برای مدل پایه و مدل با شکلک به ترتیب ۸۵/۳۰ و ۸۹/۳۷ درصد است. جدول ۱۵ و جدول ۱۶ صحت، پوشش و امتیاز f1 را برای مجموعه داده تست مدل پایه و مدل شامل شکلک شرح می‌دهد.



شکل۱۴- BiLSTM با RoBERTa از قبل آموزش‌دیده - دقت مدل

شکل ۱۴، همگرایی مدل را نشان می‌دهد که بر اساس آزمایش‌های انجام‌شده بر روی مجموعه داده‌های آموزشی با تکرار ۲۰ دور بدست آمده است. دقت بدست آمده ۸۹/۵۲ درصد می‌باشد که برای تکرارهای بعدی نیز تقریباً ثابت می‌ماند. شکل ۱۵ خطای مدل را نشان می‌دهد که که در تکرار ۱۵ به ۲۳/۹۶ درصد کاهش می‌یابد.



شکل۱۵- BiLSTM با RoBERTa از قبل آموزش‌دیده - خطای مدل

جدول ۱۵ - BiLSTM با RoBERTa از قبل آموزش‌دیده -صحت، پوشش، امتیاز f1 (مدل پایه)

مدل پایه				
پشتیبانی	معیار f1	پوشش	صحت	
۶۴۸۹	۸۶	۸۵	۸۶	۰
۶۳۵۵	۸۵	۸۵	۸۵	۱
۱۲۸۴۴	۸۵			دقت
۱۲۸۴۴	۸۵	۸۵	۸۵	میانگین کلان
۱۲۸۴۴	۸۵	۸۵	۸۵	میانگین وزنی

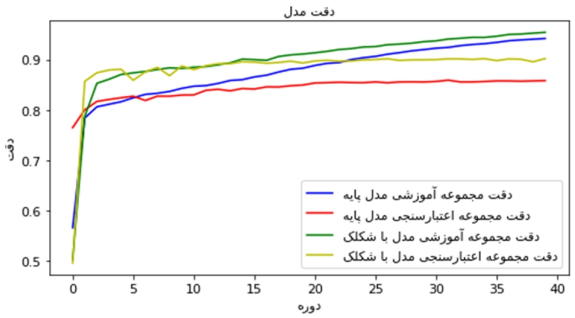
جدول ۱۶ – BiLSTM با RoBERTa از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل شامل شکلک)

مدل با شکلک				
پشتیبانی	معیار f1	پوشش	صحت	
۶۷۳۹	۹۰	۸۸	۹۲	۰
۶۱۰۵	۸۹	۹۱	۸۷	۱
۱۲۸۴۴	۹۰			دقت
۱۲۸۴۴	۹۰	۹۰	۹۰	میانگین کلان
۱۲۸۴۴	۹۰	۹۰	۹۰	میانگین وزنی

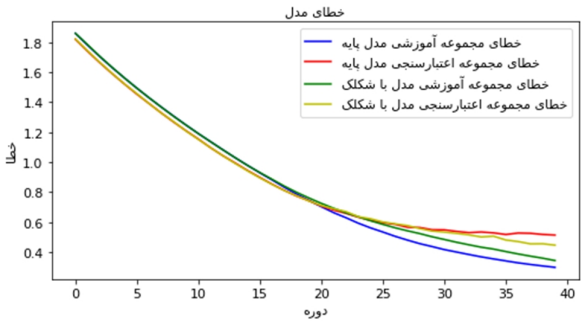
۵-۳-نتایج مدل ACBR با roberta از قبل آموزش‌دیده

در این مدل، یک لایه تعبیه با استفاده از بردار RoBERTa از قبل آموزش داده‌شده برای مدل ACBR ایجاد گردید. این مدل با استفاده از هر دو مجموعه‌داده آموزشی آموزش داده شد و با استفاده از مجموعه داده اعتبار سنجی، دقت، صحت، پوشش، امتیاز f1 و خطای مدل برای هر دو آزمایش اندازه‌گیری شد.

شکل ۱۶ و شکل ۱۷ این معیارها را برای این مدل نشان می‌دهد. سپس دقت پیش‌بینی مجموعه داده تست برای هر دو مجموعه داده اندازه‌گیری شد. دقت برای مدل پایه و مدل شامل شکلک به ترتیب ۸۵/۸۰ و ۹۰/۱۸ درصد است. جدول ۱۷ و جدول ۱۸ صحت، پوشش و امتیاز f1 را برای مجموعه داده تست مدل پایه و مدل شامل شکلک شرح می‌دهد.



شکل ۱۶– ACBR با RoBERTa از قبل آموزش‌دیده – دقت مدل



شکل ۱۷– ACBR با RoBERTa از قبل آموزش‌دیده – خطای مدل

در ارزیابی نتایج شکل‌های ۱۶ و ۱۷، دقت و خطای مدل ACBR پیشنهادی پس از تعداد دوره‌های ذکرشده همگرا می‌شود، بنابراین، این تحلیل‌ها بیانگر این است که مدل ACBR مشکل پیش برازش را کاهش داده و دقت را افزایش می‌دهد.

جدول ۱۷ – ACBR با RoBERTa از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل پایه)

مدل پایه				
پشتیبانی	معیار f1	پوشش	صحت	
۶۵۱۳	۸۶	۸۶	۸۶	۰
۶۳۳۱	۸۶	۸۶	۸۵	۱
۱۲۸۴۴	۸۶			دقت
۱۲۸۴۴	۸۶	۸۶	۸۶	میانگین کلان
۱۲۸۴۴	۸۶	۸۶	۸۶	میانگین وزنی

جدول ۱۸ – ACBR با RoBERTa از قبل آموزش‌دیده –صحت، پوشش، امتیاز f1 (مدل شامل شکلک)

مدل با شکلک				
پشتیبانی	معیار f1	پوشش	صحت	
۶۷۱۹	۹۰	۸۹	۹۲	۰
۶۱۲۵	۹۰	۹۲	۸۸	۱
۱۲۸۴۴	۹۰			دقت
۱۲۸۴۴	۹۰	۹۰	۹۰	میانگین کلان
۱۲۸۴۴	۹۰	۹۰	۹۰	میانگین وزنی

۶-ارزیابی مدل‌ها

همه مدل‌های CNN و LSTM با استفاده از مجموعه داده‌های آموزشی آزمایش‌های مدل پایه و مدل شامل شکلک با پارامترهایی نشان داده‌شده در جداول ۱۹ و ۲۰ آموزش داده شدند. در حین برازش هر مدل، عملکرد

مدل با استفاده از مجموعه‌داده اعتبارسنجی موردسنجش قرار گرفت. برای جلوگیری از پیش برازش یک تنظیم‌کننده L2 با نرخ ۰/۰۰۰۲ مورداستفاده قرار گرفت، همچنین نرخ یادگیری کاهش یافت و لایه‌های حذف تصادفی به مدل‌ها اضافه گردید. عملکرد مدل با استفاده از دقت، صحت، پوشش و امتیاز $F1$ اندازه‌گیری شد. برای مدل‌های CNN و CNN چندکاناله از پارامترهای نشان داده‌شده در جدول ۱۹ استفاده گردید.

جدول ۱۹ - پارامترهای پیشنهادشده مدل‌های شبکه عصبی کانولوشنی

پارامتر	مقدار
حداکثر تعداد کلمات	**۴۶-۴۵**
تعداد کلاس‌های لایه خروجی	**۲**
اندازه واژه‌نامه	**۴۹۲۲۷-۳۸۱۹۶**
اندازه دسته	**۱۲۸**
توابع فعال‌سازی	**relu**
تعداد تکرار (چرخش)	**۲۰**
نرخ حذف تصادفی	**۰/۱ -۰/۳ -۰/۵**
بهینه‌ساز	**Adam**
تابع زیان	**categorical_crossentropy**
نوع Pooling	**2,max**

برای مدل‌های LSTM، BiLSTM و ACBR پارامترهای نشان داده‌شده در جدول ۲۰ مورداستفاده قرار گرفت.

جدول ۲۰ - پارامترهای پیشنهادشده مدل‌های حافظه طولانی کوتاه‌مدت

پارامتر	مقدار (LSTM)	مقدار (ACBR)
حداکثر تعداد کلمات	۴۶-۴۵	**۴۶-۴۵**
تعداد کلاس‌های لایه خروجی	۲	**۲**
اندازه واژه‌نامه	۴۹۲۲۷-۳۸۱۹۶	**۴۹۲۲۷-۳۸۱۹۶**
اندازه دسته	۱۲۸	**۱۲۸**
توابع فعال‌سازی	Relu	**Relu**
تعداد تکرار (چرخش)	۲۰	**۴۰**
نرخ حذف تصادفی	۰/۵	**۰/۵**
بهینه‌ساز	Adam	**Adam**
تابع زیان	categorical crossentropy	**Categorical crossentropy**
تعداد نودهای لایه متراکم	۱۰	**۲۰**

۶-۱- مقایسه

همه مدل‌ها با معیارهای عملکرد مختلف مانند دقت، صحت، پوشش و امتیاز f1 مقایسه شدند. جدول ۲۱ عملکرد مدل‌ها را با استفاده از معیارهای کارایی ذکرشده بیان می‌کند. با مقایسه نتایج، معلوم شد که مدل ACBR با بردار تعبیه roberta دارای بالاترین دقت طبقه‌بندی برای هر دو مجموعه‌داده پایه و مجموعه‌داده شامل شکلک است که دقت آن روی مجموعه‌داده پایه ٪۸۵/۸۰ و روی مجموعه‌داده شامل شکلک ٪۹۰/۱۸ است.

جدول ۲۱ - دقت مدل‌ها روی مجموعه‌داده پایه و مجموعه‌داده شامل شکلک

مدل	دقت (Accuracy)		اختلاف (%)
	مدل شامل شکلک	مدل پایه	
word2vec + CNN	۸۴/۸۷	۷۸/۹۶	۵/۹۱
GloVe + CNN	۸۵/۶۶	۸۰/۱۶	۵/۵
BERT + CNN	۸۷/۶۶	۸۲/۹۲	۴/۷۴
word2vec + CNN چندکاناله	۸۵/۰۳	۷۹/۷۲	۵/۳۱
GloVe + CNN چندکاناله	۸۷/۴۲	۸۱/۰۸	۶/۳۴
BERT + CNN چندکاناله	۸۸/۰۳	۸۳/۰۶	۴/۹۷
word2vec + LSTM	۸۶/۹۰	۸۰/۴۹	۶/۴۱
GloVe + LSTM	۸۷/۱۶	۸۰/۹۱	۶/۲۵
BERT + LSTM	۸۸/۵۰	۸۲/۷۰	۵/۸
roberta + LSTM	۸۹/۳۷	۸۵/۳۰	۴/۰۷
word2vec + BiLSTM	۸۷/۲۴	۸۱/۲۱	۶/۰۳
GloVe + BiLSTM	۸۷/۰۱	۸۱/۲۸	۵/۷۳
BERT + BiLSTM	۸۸/۵۳	۸۳/۳۲	۵/۲۱
roberta + BiLSTM	۸۹/۵۹	۸۵/۳۷	۴/۲۲
roberta + ACBR	۹۰/۱۸	۸۵/۸۰	۴/۳۲

نتایج نشان می‌دهد عملکرد همه مدل‌ها برای مجموعه داده شامل شکلک بهتر از مجموعه‌داده پایه است. مدل‌های LSTM و BiLSTM نسبت به مدل‌های CNN و CNN چندکاناله برای این طبقه‌بندی احساسات بهتر عمل کردند. BiLSTM نسبت به LSTM کمی بهتر عمل کرده است و مدل ACBR دارای بالاترین دقت در هر دو مجموعه داده است. همچنین بردارهای از قبل آموزش‌دیده roberta و BERT از بردارهای از قبل آموزش‌دیده Glove و word2vec و بردار roberta از بردار BERT دقت بهتری را ارائه کردند.

برای دریافت بینش بهتر از بهترین رویکردهای کمی، صحت، پوشش، امتیاز f1، پشتیبانی، دقت و میانگین کلان در همه مدل‌ها در نظر گرفته شد. مدل ACBR با بردار تعبیه roberta از پیش آموزش‌دیده بهترین نتایج را در بین مدل‌های آزمایش‌شده به دست آورد. صحت، پوشش و امتیاز f1 برای مجموعه داده پایه ۰/۸۶، ۰/۸۶، ۰/۸۶ و برای مجموعه داده شامل شکلک به ترتیب ۰/۹۰، ۰/۹۰ و ۰/۹۰ بود. ازاین‌رو، مدل ACBR را با بردار تعبیه roberta از پیش آموزش‌دیده به‌عنوان بهترین مدل برای مجموعه‌داده پایه و مجموعه‌داده شامل شکلک توصیه می‌شود.

مقایسه مدل ACBR با نتایج بررسی‌شده قبلی در جدول ۲۲ نشان داده‌شده است.

جدول ۲۲ - مقایسه مدل‌های تحلیل احساسات

نویسنده / سال	مدل	مجموعه داده	دقت (٪)
بصیری و همکاران (۲۰۲۰)	ABCDM	Sentiment140	۸۱/۸۲
مورنو و همکاران (۲۰۲۱)	L-C	Sentiment140	۸۴/۲
کامیاب و همکاران (۲۰۲۱)	ACL-SA	Sentiment140	۸۷/۱۲
آیتوگ اونان (۲۰۲۲)	RCNNGWE	Sentiment140	۸۴/۰۶
مدل پیشنهادی	ACBR	Sentiment140	۹۰/۱۸

۷- نتیجه‌گیری و پیشنهاد برای تحقیقات آتی

این مقاله نظرات نوشته‌شده به زبان انگلیسی که حاوی شکلک بودند را در نظر گرفت. برای پرداختن به این موضوع، از مجموعه داده sentiment140 استفاده شد و شکلک‌ها برای تعیین قطبیت احساسات در دو کلاس مثبت و منفی در نظر گرفته شدند. مجموعه داده توییتر با استفاده از دو آزمایش موردسنجش قرار گرفت، که هر دو مجموعه داده بدست آمده با مدل‌های شبکه عصبی CNN و LSTM آزمایش شدند. در این مدل‌ها از چهار بردار تعبیه از قبیل آموزش‌دیده word2vec، GloVe، BERT و roberta استفاده شد. درنهایت، عملکرد مدل‌ها با استفاده از چند معیار عملکرد مختلف، مانند دقت، صحت، پوشش، امتیاز $F1$ و میانگین کلان مقایسه شدند. پس از بررسی معیارهای عملکرد برای همه مدل‌ها، نتایج نشان داد که مدل ACBR با بردار تعبیه از پیش آموزش‌دیده roberta، دقت بهتری را هنگام طبقه‌بندی احساسات برای این دو مجموعه‌داده ارائه می‌کند. همچنین نتایج نشان داد که شکلک‌ها ارزش بیشتری را هنگام تعیین طبقه‌بندی احساسات اضافه می‌کنند. ازآنجایی‌که کاربران بیشتری هستند که نظرات و بازخورد خود را با استفاده از شکلک‌ها ارائه می‌دهند، باید این سیگنال‌ها را هنگام شناسایی احساسات در نظر گرفت.

کارهای آینده شامل موارد زیر است:

۱) در نظر گرفتن بازخورد یا نظرات چندزبانه کاربران

۲) توسعه مدل‌هایی که می‌توانند طعنه را تشخیص دهند

۳) در نظر گرفتن تصاویر، صدا و کلیپ‌های ویدئویی، علاوه بر متن برای طبقه‌بندی احساسات

مراجع

[1]. S. Rosenthal, N. Farra, P. Nakov, "Sentiment analysis in twitter", In Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015), pp. 451- 463, Dec 2019.

[2]. S Poria, E. Cambria, N. Howard, GB. Huang and A. Hussain , "Fusing audio, visual and textual clues for sentiment analysis from multimodal content", Neurocomputing, Volume 174, Part A, Pages 50-59, 22 January 2016.

[3]. P. Chikersal, S. Poria, E. Cambria, A. Gelbukh and C.E. Siong, "Modelling public sentiment in twitter: using linguistic patterns to enhance supervised learning", in: International Conference on Intelligent Text Processing and Computational Linguistics, LNTCS, Springer, vol. 9042, pp. 49–65, 2015.

[4]. S. Kiritchenko, X. Zhu and SM. Mohammad, "Sentiment analysis of short informal texts". Journal of Artificial Intelligence Research, vol. 50, pp. 723-762, Aug 20, 2014.

[5]. Y. Kim, "Convolutional neural networks for sentence classification", In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1746-1751, 25 Aug 2014.

[6]. E. Cambria, S. Poria, A. Gelbukh and M. Thelwall, "Sentiment analysis is a big suitcase". IEEE Intelligent Systems, vol. 32, pp. 74-80, December 2017.

[7]. Y. Zhang and B. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification" IJCNLP, 13 Oct 2015.

[8]. C. dos Santos, M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts". In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics, pp. 69-78, August 2014.

[9]. M. Cliché, "Twitter sentiment analysis with cnns and lstms", Published in Proceedings of SemEval-2017, arXiv:1704.06125, 8 pages, 20 Apr 2017.

[10]. T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient Estimation of word representations in vector space" arXiv:1301.3781, Jan 2013.

[11]. A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong and U. Kaymak, "Exploiting emoticons in sentiment analysis", in: Proceedings of the 28th Annual ACM Symposium on Applied Computing, pp. 703–710, March 2013.

[12]. N. Hu, P.A. Pavlou, and J. Zhang, "Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online Word-of-Mouth communication", EC '06: Proceedings of the 7th ACM conference on Electronic commerce, Pages 324–330, June 2006.

[13]. M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar and M. Mahmoud Al-Ayyoub, " SemEval-2016 Task 5: Aspect Based Sentiment Analysis", in: 10th International Workshop

analysis", Future Generation Computer Systems, Vol. 115, pp. 279-294 September 2020.
[27]. M. Kamyab, G. Liu and M. Adjeisah, "Attention-Based CNN and Bi-LSTM Model Based on TF-IDF and GloVe Word Embedding for Sentiment Analysis", Applied Sciences, Vol. 11, November 2021.
[28]. A. Onan, "Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification", Journal of King Saud University - Computer and Information Sciences, Vol 34, pp. 2098-2117, May 2022.
[29]. M.A. Ullah, S.M. Marium, Sh. Begum, N.S. Dipa, "An algorithm and method for sentiment analysis using the text and emoticon", ICT Express, Vol. 6, pp. 606–615, December 2020.
[30]. B.Pang, , L.Lillian, and S. Vaithyanathan, " Thumbs Up? Sentiment classification using machine learning techniques", In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP), Computation and Language, pp. 79–86, May 2002.
[31]. M. Hu, and B. Liu, "Mining and summarizing customer reviews" In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 168-177, August 2004.
[32]. D. Sharma, M. Sabharwal, V. Goyal and M. Vij, "Sentiment analysis techniques for social media data: A review", First International Conference on Sustainable Technologies for Computational Intelligence, Springer, pp.75-90, January 2020.
[33]. k. jagmeet and M. sabharwal, " Spam detection in online social networks using feed forward neural network", In: RSRI Conference on Recent Trends in Science and Engineering, pp. 69–78, 2018.
[34]. M. Unnisa, A. Ameen and S. Raziuddin. " Opinion mining on twitter data using unsupervised learning technique", International Journal of Computer Applications, vol 148(12), 2016.
[35]. X. Zhang, J. Zhao and Y. LeCun. "Character-level convolutional networks for text classification", In Advances in Neural Information Processing Systems 28, Neural Information Processing Systems Foundation. pp. 649–657, 2015.
[36]. A. Karpathy, J. Johnson and L. Fei-Fei. "Visualizing and understanding recurrent networks", ICLR, arXiv:1506.02078, Jun 2015.
[37]. C. Baziotis, N. Pelekis and C. Doulkeridis, "Datastories at semeval 2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis", In Proceedings of the 11th International Workshop on Semantic Evaluation, Association for Computational Linguistics, pp. 747-754, August 2017.
[38]. J. Devlin, M.W. Chang, K. Lee and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding", Proceedings of NAACL-HLT, pp: 4171–4186, Oct 2018.
[39]. A.Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez,Ł. Kaiser and I. Polosukhin, ''Attention is all you need,'' in Proc. 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, pp. 6000–6010, 2017.
on Semantic Evaluation (SemEval), pp.19 – 30, Jan 2019
[14]. Y. Wan and Q. Gao, "An ensemble sentiment classification system of twitter data for airline services analysis", IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, pp. 1318–1325, November 2015.
[15]. A. Pak and P. Paroubek. "Twitter as a corpus for sentiment analysis and opinion mining", Conference: Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010,vol. 13, pp. 1320–1326, 17-23 May 2010.
[16]. W. Wolny, "Emotion analysis of twitter data that use emoticons and emoji ideograms", The 25th International Conference on Information Systems Development (ISD2016), European Language Resources Association (ELRA), Valletta, Malta, Corpus ID: 31794814, 2016.
[17]. E. Kouloumpis, T. Wilson and J. Moore, "Twitter sentiment analysis: The good the bad and the omg!", in: Fifth International AAAI Conference on Weblogs and Social Media, Vol. 5, pp. 538-541, 2011.
[18]. D. Jurafsky and J. H. Martin, "Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition", In Prentice Hall series in artificial intelligence, Prentice Hall, 2000.
[19]. J. Lee, W. Yoon, S. Kim, D. Kim, C. H. So and J. Kang, "Biobert: A pre-trained biomedical language representation model for biomedical text mining", Bioinformatics, Oxford, England, vol. 36(4), pp. 1234–1240, 15 February 2020.
[20]. Y. Goldberg and O. Levy, "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method", arXiv:1402.3722, 15 Feb 2014.
[21]. X. Fu, W. Liu, Y. Xu and L. Cui, "Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis". Neurocomputing, Vol 241, pp. 18–27, 7 June 2017.
[22]. E. Bender and A. Lascarides, "Linguistic fundamentals for natural language processing II: 100 essentials from semantics and pragmatics". Synthesis Lectures on Human Language Technologies, vol. 12(3), 1–268 pages, November 2019.
[23]. Y. Sharma, G. Agrawal, P. Jain and T. Kumar, "Vector representation of words for sentiment analysis using GloVe", International Conference on Intelligent Communication and Computational Techniques (ICCT), IEEE, ISBN:978-1-5386-3030-3, December 2017.
[24]. A. Varshney, Y. Kapoor, A. Thukral, R. Sharma and P. Bedi, "Performing Sentiment Analysis on Twitter Data Using Deep Learning Models: A Comparative Study" Advances in Data and Information Sciences, vol 318, pp 371–381, 08 February 2022.
[25]. C.N. Dang, M.N. Moreno-García and F. Prieta, "Hybrid Deep Learning Models for Sentiment Analysis", Vol. 2021, 16 pages, Aug 2021.
[26]. M.E. Basiri, Sh. Nemati , M. Abdar, E. Cambria and U.R. Acharrya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment

[40]. Attention layer Retrieved from: https://github.com/deepak-kaji/mimic-lstm

زیرنویس‌ها

1. Sentiment Analysis
2. Emoticons
3. Convolutional neural network (CNN)
4. Long short-term memory (LSTM)
5. Word embedding
6. Natural language processing
7. Context
8. Document level
9. Support vector machine(SVM)
10. Naive bayes
11. One-hot
12. Term frequency–inverse document frequency(TF-IDF)
13. Bidirectional long short-term memory
14. Bidirectional encoder representations from transformers (BERT)
15. Sentence level
16. Aspect level
17. Data preprocessing
18. features selection
19. Stemming
20. Stop words
21. Supervised learning
22. Unsupervised learning
23. Anisa
24. maximum entropy
25. Accuracy
26. Precision
27. Recall
28. HTML
29. Natural Language Toolkit
30. Tokenizer
31. Devlin
32. Next Sentence Prediction(NSP)
33. Self attention
34. Embedding layer
35. Convolutional layer
36. regularization
37. Dense layer
38. Cliche
39. Attention Baseb Cnn And Lstm Model With RoBERT