سال سیزدهم، شماره‌های ۴۷ و ۴۸، بهار و تابستان ۱۴۰۰
صص: ۹۰_۱۰۴

دو فصلنامه علمی
فناوری اطلاعات و ارتباطات ایران

تحلیل احساسات برای پیش‌بینی بازار بورس با شبکه عصبی ژرف: مطالعه موردی برای پایگاه داده سهام شرکت‌های بین-المللی

حکیمه منصور* دکتر سعیده ممتازی دکتر کامران لایقی*****

*دانشجوی دکتری هوش مصنوعی-رباتیکز، گروه مهندسی کامپیوتر و برق ، شعبه شمال تهران ، دانشگاه آزاد اسلامی

** استادیار، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

*** استادیار، گروه مهندسی کامپیوتر و برق، شعبه شمال تهران، دانشگاه آزاد اسلامی

تاریخ دریافت: ۱۳۹۹/۰۶/۰۳ تاریخ پذیرش: ۱۳۹۹/۱۲/۱۳

نوع مقاله: پژوهشی

چکیده

امروزه تحلیل احساسات به عنوان یکی از ارکان اصلی در زمینه‌های مختلف از جمله مدیریت مالی، بازاریابی، پیش‌بینی تغییرات اقتصادی در کشورهای مختلف بکار گرفته می شود. به منظور ساخت یک تحلیل‌گر احساسات بر مبنای نظرات کاربران در رسانه‌های اجتماعی، بعد از استخراج ویژگی‌های مهم بین کلمات توسط شبکه پیچشی، از شبکه حافظه کوتاه-مدت بلند استفاده می‌کنیم تا رابطه نهفته در دنباله‌ای از کلمات را کشف و ویژگی‌های مهم متن را استخراج نماییم. با کشف ویژگی‌های استخراج شده جدید توسط شبکه برگشتی با حافظه کوتاه-مدت بلند، توانایی مدل پیشنهادی در طبقه‌بندی ارزش سهام شرکت‌ها افزایش می‌یابد و در نهایت به پیش‌بینی سهام بورس در روز بعد براساس تحلیل احساسات می‌پردازیم. این پژوهش بر اساس داده‌های مقاله انگویان و همکارانش انجام گرفته است و تنها از اطلاعات احساسی مردم در شبکه‌های اجتماعی برای پیش‌بینی سهام استفاده می‌کند. با توجه به اینکه هر یک از پیام‌های کاربران را در ۵ کلاس‌های احساسی طبقه‌بندی می‌کنیم، بنابراین این مدل ارزش سهام روز بعد را به دو حالت بالا یا پایین بودن آن می‌تواند پیش‌بینی کند. ساختار پیشنهادی شامل ۲۱ لایه شبکه عصبی ژرف و متشکل از شبکه‌های پیچشی و حافظه کوتاه-مدت بلند است که برای پیش‌بینی سهام بورس ۱۸ شرکت پیاده سازی شده است. اگرچه برخی مدل‌های ارائه شده قبل، از تحلیل احساسات به منظور پیش‌بینی بازار سرمایه بهره گرفته‌اند، اما از روش‌های ترکیبی و پیشرفته در شبکه‌های ژرف با میزان دقت پیش‌بینی بالا بهره نبرده‌اند. سنجش نتایج روش پیشنهادی با دیگر مطالعات نشان داده که عملکرد روش پیشنهادی در مقایسه با ۸ روش دیگر، بطور قابل ملاحظه‌ای خوب بوده و در معیار ارزیابی صحت در پیش‌بینی روزانه سهام با بهبود ۱۹/۸ درصدی نسبت به مدل شبکه پیچشی ژرف، ۲۴/۵ درصدی نسبت به مدل پیشنهادی انگویان و همکاران و ۲۳/۹۴ درصدی نسبت به مدل پیشنهادی درخشان و همکاران از روش‌های رقیب پیشی بگیرد.

واژگان کلیدی: تحلیل احساسات، شبکه عصبی ژرف، مدل‌های شبکه عصبی پیچشی، شبکه حافظه کوتاه-مدت بلند، بازار سهام

۱- نویسنده مسئول: دکتر سعیده ممتازی، momtazi@aut.ac.ir

۱ مقدمه

بازار مالی به عنوان یک سیستم پیچیده با نوسانات غیرخطی از جمله بخش‌هایی است که از مزایای الگوریتم‌های متن‌کاوی بهره‌مند شده

است. به عنوان مثال، متن‌کاوی در زمینه‌های مختلفی مانند درک و مدیریت ریسک مالی، مدیریت تسهیلات، رتبه‌بندی اعتباری مشتریان، تجزیه و تحلیل و رتبه‌بندی مشتریان بانکی، پیش‌بینی سود سهام، پیش‌بینی بازار بورس، پیش‌بینی نرخ ارز، پیش‌بینی ورشکستگی بانک‌ها، تحلیل احساسات سرمایه‌گذاران، تحلیل نظرات مشتریان و تشخیص کلاهبرداران مالی به ایفای نقش پرداخته است. عموماً سرمایه‌گذاران بازار مالی به منظور سود بیشتر، نظرات و حالت-های احساس خود را با دیگر سرمایه‌گذاران به اشتراک می‌گذارند و در نتیجه رفتار سرمایه‌گذاران مالی بر قیمت‌های بازار سهام تاثیرگذار خواهد بود. ولی از آنجاییکه غالب سرمایه‌گذاران از روند تغییرات بازارهای مالی اطلاع ندارند بنابراین بررسی حالت‌های احساسی سرمایه‌گذاران و شناسایی الگوهای مخفی بین نظرات مردم در رسانه-های اجتماعی، می‌تواند آنها را برای رسیدن به سودآوری بیشتر، شناخت تغییرات بازار و پیشگویی حوادث اقتصادی یاری کند.

امروزه محققان در صدد برآمدند تا با بکارگیری تکنیک‌های هوش مصنوعی مانند شبکه عصبی پیچشی[۱] و شبکه‌های ژرف به ارزیابی نظرات، تحلیل داده‌های احساسی مشتریان یا کاربران در فضاهای مجازی و نیز شناسایی و پیش‌بینی تغییرات در بازار بپردازند. سیستم‌های هوشمند تحلیلگر احساسات، وظیفه تحلیل و طبقه‌بندی نظرات، توصیفات و نگرش‌های مردم (نسبت به موضوعات مطرح در متن) در گروه‌های مثبت، منفی و خنثی را دارند. مطالعه حاضر با در نظر گرفتن این مهم، سعی نموده است تا با بکارگیری شبکه عصبی ژرف اقدام به پیش‌بینی تغییرات بازار بورس شرکت‌های بین‌المللی نموده و مدلی کارامد در این زمینه پیشنهاد نماید.

به منظور ارزیابی کمّی اطلاعات مالی و اقتصادی و نیز پیش‌بینی بازار سهام، محققان بازار مالی براساس داده‌های ورودی دو مقیاس کلی تحلیل بنیادی[۲] و تحلیل فنی[۳] را ترسیم می‌کنند[۱]. در تحلیل فنی متغیرهای سری‌های زمانی به‌کار می‌رود و بایستی تاریخچه ارزش سهام در روزهای قبلی برای پیش‌بینی تغییرات و یا ارزش سهام در روزهای آتی به‌کار رود. این متغیرها عبارتند از قیمت سهام، حجم معاملات، بالاترین و پایین‌ترین ارزش شاخص، ارزش شاخص در شروع و پایان معاملات در روز و غیره . بنابراین برای پیش‌بینی ارزش سهام کافیست الگوهای قبلی قیمت را بررسی کرده و موقعیت‌های مشابه را کشف نماییم [۱]. هدف از تحلیل فنی پیش‌بینی سیر صعودی و نزولی ارزش سهام در کوتاه مدت می‌باشد. به عبارتی پاسخ به سوال «چه زمانی باید یک سهام خاص خریده شود؟».

در تحلیل بنیادی داده‌های اقتصاد کلان، گزارش‌های مالی منظم، کیفیت مدیریتی، فعالیت‌های مالی دولت و شرایط سیاسی یا جغرافیایی مورد توجه قرار می‌گیرد. این مقیاس‌ها می‌تواند شامل قیمت طلا، نرخ ارز خارجی، نرخ بهره، نرخ مالیات، شاخص خرده‌فروشی، شاخص‌های تولید صنعتی، و نرخ ارز خارجی باشد. در این تحلیل بایستی با مطالعه این فاکتورهای بنیادی ارزش سهام پیش‌بینی شوند [۱]. هدف از تحلیل بنیادی این است که آیا ارزش فعلی سهام معقول و منطقی است و یک توسعه بلندمدت را نشان می‌دهد. به عبارتی پاسخ به سوال «چه سهامی امروز باید خرید؟».

نظرکاوی یا کاوش نظرات یکی از زمینه‌های نوین مطالعاتی در بازاریابی، مدیریت مالی و اقتصاد می‌باشد که وظیفه استخراج دانش پنهان در نظرات و دیدگاه‌های مصرف‌کنندگان برای کالاهای مختلف، خدمات شهری، خدمات شرکت‌ها، سرویس‌های عمومی یا مالی، خدمات بانک‌ها، تراکنش‌های مالی و... را بر عهده دارد. مزایای به‌کارگیری این تکنیک، استخراج اطلاعات به منظور شناسایی معایب محصولات و خدمات، بزرگنمایی توانمندی‌ها در تجارت، تعیین میزان علاقه‌مندی کاربران نسبت به خدمات، بهبود ارائه خدمات، طبقه-بندی رضایت‌مندی کاربران نسبت به محصولات یا خدمات ارائه شده، می‌باشد [۲].

از لحاظ محاسباتی روش‌های نظرکاوی به سه دسته کلی سطح متن، سطح جمله و سطح ویژگی تقسیم شده است. مقصود از نظرکاوی در سطح متن تعیین گرایش معنایی کل متن و در سطح جمله، تعیین احساس یک جمله به صورت مستقل است. تحلیل سطح ویژگی یا جنبه‌ای احساس واقعی هر ویژگی را در مورد موضوعات هدف بیان می‌کند [۳]. به‌طورکلی بکارگیری روش‌های متن‌کاوی مبتنی بر تحلیل احساسات به عنوان رویکردی نوین در چرخه تولید تا مصرف، روندهای سرمایه‌گذاری و ارائه خدمات به‌شمار می‌آید. کشف دانش فنی و متن‌کاوی نظرات مشتریان و کاربران در تمامی فعالیت‌های چرخه مدیریت ارتباطات مالی اعم از شناخت، جذب، حفظ و توسعه خدمات و بهره‌برداری کاربرد دارد که مستقیماً با رضایت‌مندی مشتریان و کاربران ارتباط دارد [۴].

یادگیری ژرف که اخیراً در زمینه‌های مختلف علمی مورد توجه قرار گرفته است درواقع همان یادگیری شبکه عصبی است که دارای لایه‌های پنهانی زیادی بوده و برخی از مشکلات شبکه عصبی همانند بیش‌برازش و ناپدید شدن گرادیان[۴] توسط الگوریتم‌های خاصی برطرف شده‌اند. این نوع یادگیری تلاش می‌کند مفاهیم انتزاعی سطح بالا با ساختار پیچیده را با استفاده از یادگیری در لایه‌های مختلف مدل‌سازی نماید. هابیمانا و همکاران بعد از بررسی وظایف و کاربردهای تجزیه و تحلیل احساسات، به اشکالات و موانع موجود در استفاده از رویکردهای سنتی رایج در این زمینه اشاره کردند [۵]. آنها با معرفی پژوهش‌های انجام شده توسط روش‌های یادگیری ژرف در زمینه تجزیه و تحلیل احساسات، به مزایا، کاربردها، چالش‌ها و

---

۱ Convolutional neural network (CNN)

۲ Fundamental analyses

۳ Technical

۴ vanishing gradients

روند توسعه رویکردهای مذکور پرداختند. هابیمانا و همکاران اذعان داشتند که به منظور درک ژرف و شفاف احساسات موجود در داده-های کاربران، نیاز به رویکردهای یادگیری ژرف است که توانایی یادگیری متن بدون استخراج دستی ویژگی‌ها را داشته و نسبت به روش‌های سنتی عملکرد بهتری دارند. چانگ و همکاران، پرابا و همکاران، یاداو و همکاران از جمله محققانی هستند که نقش یادگیری ژرف را به عنوان یک تکنیک قدرتمند یادگیری در تجزیه و تحلیل احساسات بررسی کردند و به مزایا، معایب و نتایج پیشرفته آن در این حوزه اشاره کردند [۶] تا [۸].

روش‌های یادگیری ماشین در دو دسته اصلی با ناظر در صورت وجود برچسب و بدون ناظر در حالت بدون برچسب تقسیم می‌شوند که بیشتر کارهای انجام شده در سطح متن مبتنی بر روش‌های با ناظر می‌باشد همچنین روش‌های یادگیری ماشینی برای دسته‌بندی متون در سطح کلمه و در سطح جمله ارائه شده است. به طور کلی ، مدل پیشنهادی این مقاله از چندین دیدگاه با روش‌های موجود متفاوت است: ۱- تنظیم دقیق وزنهای ورودی برای لایه‌های بعدی و آرایش لایه‌های انتخاب شده موجب یادگیری بهتر این مدل شده است. طوریکه تحلیل احساسات نظرات کاربران دارای دقت قابل توجهی در پیش‌بینی سهام روز بعد ایجاد کرده است ۲- روش ترکیبی ارائه شده با استفاده از خصوصیت شبکه پیچشی به منظور استخراج ویژگی‌های محلی بین کلمات و شبکه حافظه کوتاه-مدت بلند[۱] در استخراج ویژگی‌های دنباله‌ای جملات متن و ترکیب این دو شبکه در ساخت تحلیلگر احساسات توانست روابط نهفته در دنباله‌ای از کلمات و جملات را کشف و ویژگی‌های مهم متن را استخراج نماید. این فرایند موجب بهبود در پیش‌بینی روند تغییرات سهام در فاز بعد گردید. اغلب کارهای انجام شده در زمینه تحلیل احساسات در بازار سهام مبتنی بر روش‌های گرافی، شبکه عصبی ساده و یا شبکه ژرف هستند که هرکدام دارای مزایا و معایب خاصی می‌باشند. به منظور حذف این معایب و افزایش پتانسیل‌های این پژوهش‌ها، این مقاله از شبکه عصبی ژرف[۲] ترکیبی با قدرت بالای پردازش و تحلیل احساسات بهره گرفته است. ساختار این مقاله به صورت زیر ادامه می‌یابد: در بخش دو به کارهای مرتبط اشاره می‌شود. فصل سوم به بررسی شبکه‌های ژرف عصبی پیچشی و حافظه کوتاه-مدت بلند می‌پردازد. ساختار مدل پیشنهادی در بخش چهار بیان خواهد شد و در بخش ۵ به دادگان و نتایج پژوهش اشاره می‌شود.

۲ کارهای مرتبط

تفاوت ساختاری داده‌های رسانه‌های اجتماعی با دیگر انواع داده‌های ورودی موجب ایجاد چالش‌های متفاوتی در مواجهه با این نوع داده‌ها می‌باشد. رسانه‌های اجتماعی دارای داده‌های نویزی، بدون ساختار، مبهم با بسامد بالا هستند. این داده‌ها اغلب اسناد کوتاه با محتوای کمتر و بدون ویرایش می باشند. تحقیقات اخیر ارتباط بین احساسات استخراج شده از رسانه‌های اجتماعی و بازار سهام را به خوبی بیان می‌کنند. چن ثابت کرد که ارزش سهام با گزارش خبری تغییر می‌کند. سهامی که دارای اخبار عمومی بدی هستند تاثیر منفی ۱۲ ماهه روی سهام داشته و سهام با اخبار خوب دارای افزایش قیمت و یا افت ارزش کمتری می‌باشند. همچنین سهام بدون خبر در یک ماه، عملکرد نامتعارف کمی در ماه بعدی داشته است [۹]. بولن و همکاران اثبات کردند که تغییرات احساسات عمومی در توییتر با میانگین صنعتی سهام داوجونز آمریکا مرتبط است [۱۰]. آنها با طراحی یک شبکه عصبی فازی خودسازمان‌ده[۳] که براساس داده‌های سهام داوجونز در روزهای قبلی و همچنین سری زمانی احساس عموم آموزش داده شده بود توانستند دقت پیشگویی ارزش سهام در روزهای آینده را بهبود دهند. لو و همکاران گزارش دادند که رسانه‌های اجتماعی شاخص مهمی در پیش‌بینی ارزش سهام شرکت‌ها نسبت به سایر رسانه‌های آنلاین به شمار می‌آیند و آن‌ها مقدار پیش‌بینی را در سریع‌ترین زمان ارائه می‌دهند [۱۱]. یو و همکاران با بررسی تنها دو شاخص ارزیابی عملکرد شرکت‌ها یعنی سود سهام و ریسک آن به این نتیجه رسیدند که نظرات رسانه‌های اجتماعی تأثیر بسیار زیادی نسبت به رسانه‌های عادی روی عملکرد سهام شرکت دارند [۱۲]. همچنین آن‌ها اذعان داشتند که نظرات هر کدام از رسانه‌های اجتماعی (وبلاگ‌ها، توییتر و غیره) تأثیر متفاوتی در عملکرد سهام شرکت داشته است. به عنوان مثال احساسات موجود در وبلاگ‌ها تاثیر مثبتی در بازدهی سهام دارد درحالیکه احساسات موجود در انجمن‌های وب تاثیر منفی در بازدهی سهام داشته است [۱۲]. لی و همکاران با طراحی یک سیستم معاملاتی مبتنی بر رسانه‌های اجتماعی تاثیر این رسانه‌ها را بر روی بازار سهام به خوبی نشان داد [۱۳]. این محققان اثبات کردند که با افزایش اطلاعات سرمایه‌گذاران براساس اخبار مختص شرکت‌ها، تصمیمات آنها تحت تاثیر قرار می‌گیرد. در ضمن احساسات عمومی موجب ایجاد نوسانات احساسی در سرمایه‌گذاران شده‌اند و در تصمیم‌گیری‌های آنها دخالت می‌کنند[۱۳]. مقالات فوق و سایر مطالعات اخیر به خوبی توانسته است ارتباط مستقیم بین ارزش سهام شرکت‌ها و رسانه‌های اجتماعی را نمایان سازد. بنابراین آنچه که در اینجا به عنوان چالش اساسی مطرح می‌شود اینست که «چگونه می‌توان پل ارتباطی بین رسانه‌های اجتماعی و ارزش سهام شرکت‌ها را برقرار کرد؟». محققان با بهره‌گیری از انواع روش‌های آماری و داده‌کاوی به بررسی رفتار بازار مالی براساس تحلیل نظرات مردم

[۱] Long Short-Term Memory (LSTM)

[۲] Deep learning

[۳] Self-Organizing Fuzzy Neural Network

جهت پیش‌بینی ارزش سهام پرداخته‌اند. الگوریتم‌های داده‌کاوی مختلفی مانند ماشین بردار پشتیبان[۱]، شبکه‌های عصبی مصنوعی[۲]، رگرسیون خطی[۳]، درخت تصمیم[۴]، بیز ساده[۵] و مبتنی بر فازی جهت پیش‌بینی بازار مالی به‌کار رفته است. اوه و همکاران اثبات کردند که حالات و احساسات رفتاری مردم که در رسانه‌های عمومی انعکاس می‌یابند نقش بسیار مهمی در فرایند تصمیم‌گیری سرمایه‌گذاران ایفا می‌کنند [۱۴]. روان و همکاران پیشنهاد داد که می‌توان از اعتماد بین کاربران رسانه‌های اجتماعی به عنوان مکانیسمی برای تقویت یا تضعیف رسانه‌های اجتماعی استفاده کرد تا ارتباط این رسانه‌ها با داده‌های مالی بازار سهام افزایش یابد [۱۵]. آنها با جمع‌آوری داده‌های توییتر ارتباط بین احساسات در توییتر و بازده غیر معمول سهام را بررسی کردند. روان اعتقاد داشت که نویسنده توییت دارای قدرت و اعتبار خاص خود در جامعه است که می‌تواند بر اعتماد بین دو کاربر نسبت به همدیگر و درنتیجه در بازار سهام تاثیرگذار باشد. انگویان و همکاران مدلی برای پیش‌بینی روند قیمت سهام ۱۸ شرکت در Yahoo Finance براساس احساسات رسانه‌های اجتماعی پیشنهاد کردند. آنها موضوعات و احساسات مرتبط با موضوع خاص هر شرکت را از متون در تابلوی پیام‌ها استخراج و ۶ نوع ویژگی مختص داده‌ها تعریف نمودند و در نهایت با استفاده از بردار پشتیبان مبتنی بر کرنل پیام‌ها در۵ دسته «حتماً بخر»[۶]، «بخر»[۷]، «دست نگه دار»[۸]، «بفروش» و «حتماً بفروش»[۹] طبقه‌بندی شدند [۱۶]. شانگ و همکاران از یک شبکه عمیق سه لایه‌ای برای پیش‌بینی بازار سهام ۳۸ سهام کشور کره استفاده کردند [۱۷]. ایشان از بازده سهام روزهای قبل و با استفاده از سه روش تجزیه و تحلیل مؤلفه‌های اصلی، کدگذاری خودکار[۱۰] و ماشین محدود شده بولتزمن[۱۱] هفت مجموعه ویژگی استخراج کردند. آنها معتقد هستند که یکی از مزایای اصلی یادگیری شبکه عمیق توانایی استخراج ویژگی‌ها از یک مجموعه بزرگ داده‌های خام بدون دانش قبلی[۱۲] است. این باعث می‌شود تا یادگیری عمیق برای پیش‌بینی بازار سهام مناسب‌تر باشد چون عوامل متعدد به صورت غیرخطی و پیچیده بر قیمت سهام تاثیرگذار است. چن و لیو مدل‌هایی مبتنی بر شبکه‌های عصبی برگشتی ارائه دادند تا نوسانات سهام در بازار چین را پیش‌بینی کنند [۱۸] و [۱۹]. چن و همکارانش بعد از خزش[۱۳]در وب سایت اجتماعی به تحلیل احساسات مردم از پست‌هایشان پرداختند. سپس براساس احساسات مردم و ارائه یک ویژگی نفوذ احساسی[۱۴] دقت مدل را افزایش دادند. ویژگی جدید با دیگر ویژگی‌های مؤثر در قیمت سهام ترکیب شده و به عنوان ورودی به شبکه عصبی برگشتی داده می‌شود. لیو با ایجاد یک فرهنگ لغت کلمات احساسی-مالی به محاسبه نمره احساسی هر پست آنلاین مربوط به سهام خاص اقدام کردند و سپس براساس اطلاعات احساسی برای پیش‌بینی سهام، مدل شبکه عصبی برگشتی را بکار بردند. نصراله‌پور با استفاده از ابزارهای یادگیری ماشینی از جمله پرسپترون چند لایه و ماشین بردار پشتیبان به پیش‌بینی نوسانات نرخ ارز در متون خبری اقتصادی فارسی پرداخت [۲۰]. صدر و همکاران با استفاده از ترکیب شبکه-های عصبی پیچشی و برگشتی مدل پیشنهادی خود را ارائه دادند [۲۱]. آنها لایه ادغام در شبکه پیچشی را با شبکه عصبی برگشتی جایگزین کردند تا بتوانند وابستگی های طولانی مدت را استخراج کرده و و زیان[۱۵] اطلاعات محلی را کاهش دهند. در مقاله دیگر صدر وهمکاران، آنها با پیشنهاد یک شبکه ژرف چند نمایی[۱۶] از ویژگی‌های میانی استخراج شده از شبکه‌های عصبی پیچشی و برگشتی برای انجام طبقه‌بندی استفاده می‌کنند [۲۲]. آنها اذعان داشتند که شبکه‌های عصبی ژرف مختلف به دلیل ساختارهای متمایز قادر به استخراج انواع مختلف ویژگی‌ها هستند، بنابراین در مدل پیشنهادی، ویژگی‌های استخراج شده از شبکه‌های عصبی ناهمگن را با استفاده از طبقه‌بندی کننده‌های چندنمایه ترکیب کردند تا با در نظر گرفتن ارتباط بین آنها، بتوانند عملکرد تحلیلگر احساسات در سطح متنی را بهبود دهند. درخشان و همکاران روش گرافیکی از اجزای کلام[۱۷] بنام "LDA-POS" جهت استخراج نظرات کاربران ارائه دادند و آن را بر داده‌های انگلیسی و فارسی پیاده‌سازی نمودند [۲۳]. در این مدل برچسب‌های اجزای کلام بر روی روش‌های مدل‌سازی موضوع[۱۸] قرار گرفت که میانگین دقت این روش در مجموعه داده‌های به زبان انگلیسی و فارسی به ترتیب ۵۶/۲۴٪ و ۵۵/۳۳٪ گزارش شد.

محققان در سال‌های اخیر با ایجاد سیستم‌های جدید تحلیلگر احساسات بر مبنای تکنیک‌های مختلفی از شبکه عصبی ژرف، دقت پیش‌بینی سهام بورس را بهبود دادند تا سرمایه‌گذاران بورس بتوانند با آگاهی بهتر از روند تغییرات بورس، بازده سهام خود را افزایش دهند [۲۴] تا [۲۸]. اگرچه عملکرد شبکه پیچشی در انجام وظایف طبقه‌ندی و شبکه کوتاه مدت-بلند در پیش‌بینی سری های زمانی دارای نتایج خوبی بوده‌اند اما در سال‌های اخیر محققان متعددی به خوبی اثبات کردند که نتایج عملکرد حاصل از ترکیب شبکه‌های

---

۱ Support Vector Machine (SVM)
۲ Artificial Neural Network (ANN)
۳ Linear Regression
۴ Decision tree
۵ Naive Bayesian Classification(NBC)
۶ Strong buy
۷ Buy
۸ Hold
۹ Strong sell
۱۰ Autoencoder
۱۱ restricted boltzmann machine
۱۲ Prior knowledge
۱۳ crawl
۱۴ sentiment influence feature
۱۵ loss
۱۶ multi-view deep network
۱۷ part-of-speech
۱۸ subject modeling methods

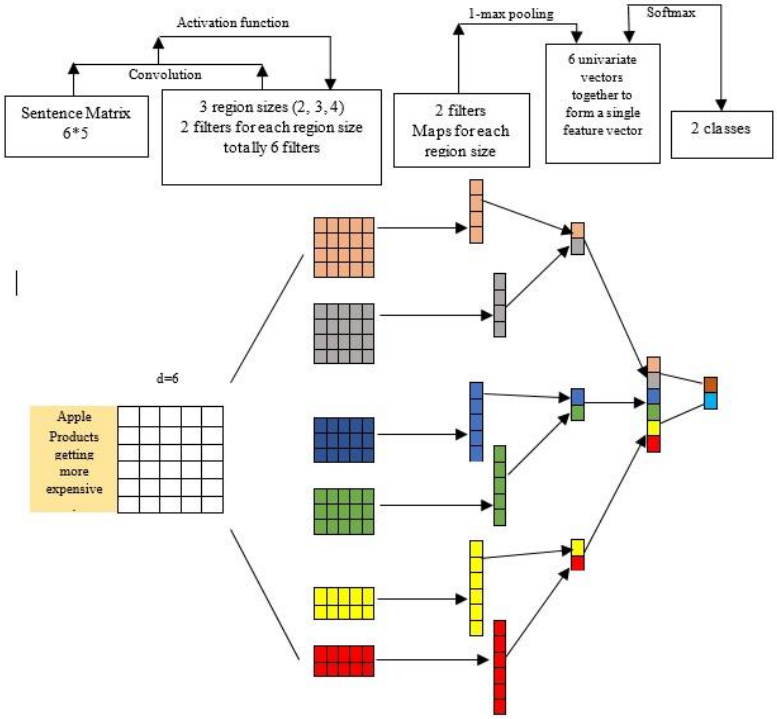
عصبی ژرف پیچشی و حافظه کوتاه-مدت بلند در تحلیل و طبقه-بندی احساسات بسیار بهتر از عملکرد منفرد آنها بوده است [۲۹] تا [۳۳]. بنابراین ما بر آن شدیم تا از شبکه ترکیبی پیچشی و حافظه کوتاه-مدت بلند به منظور پیش‌بینی بازار بورس توسط تحلیل احساسات کاربران بهره بگیریم. اما این مقاله دارای وجه تمایز عمده-ای با سایر پژوهش‌های انجام شده در این زمینه می‌باشد. اول اینکه پژوهش‌های قبلی در زمینه بازار بورس عموماً براساس مقدار داده کمی بوده است و کمتر مقاله‌ای دارای این حجم وسیع از داده‌ها به منظور تحلیل احساسات بوده است. دوم اینکه کارهای انجام شده بر روی مجموعه داده این مقاله کارهای گرافی بوده و این اولین باری است که شبکه ژرف برای این داده‌ها طراحی شده‌اند. سوم اینکه تحلیل احساسات در ۵ مجموعه با چنین دقت بالایی در هیچ مقاله‌ای مشاهده نشده است و عموماً طبقه‌بندی‌هایی با دقت بالا در دو نوع کلاس‌بندی بوده‌اند [۳۴] تا [۳۶]. مطالعه حاضر نیز با تکیه بر پژوهش‌های انجام شده قبل، به منظور بهبود عملکرد این سیستم‌ها

۳ شبکه‌های عصبی ژرف

در این تحقیق، از دو شبکه پیچشی و شبکه برگشتی با حافظه کوتاه – مدت بلند برای تحلیل احساسات نظرات کاربران در رسانه‌های اجتماعی استفاده شده است. که در بخش‌های زیر هرکدام رابه صورت مختصر توضیح می‌دهیم.

۳. ۱ شبکه عصبی پیچشی

شبکه‌های عصبی پیچشی توسط لیکان معرفی گردیده که به عنوان یکی از روش‌های یادگیری ژرف برای اولین بار برای طبقه‌بندی تصاویر و بینایی ماشین[۱] از آن استفاده شد تا بتواند مشکل مقیاس‌پذیری شبکه‌های سنتی نسبت به کل تصویر را از بین ببرد [۳۷]. این شبکه‌ها بعد از قطعه‌بندی تصویر به قسمت‌های کوچک آن را با قطعات تصویر واقعی مقایسه کرده و با یافتن ویژگی‌های برجسته تقریباً در مکان‌های یکسانی از دو تصویر نسبت به طبقه‌بندی کل تصویر بسیار بهتر عمل می‌کنند. در این شبکه‌ها طی عملیاتی



شکل ۱. شمایی از عملکرد شبکه پیچشی در پردازش زبان طبیعی (الهام گرفته از مقاله [۴۳])

در راستای افزایش دقت پیش‌بینی ارزش سهام توانسته است معیار صحت را به ۸۰/۱۸ درصد بهبود دهد که این نتیجه در مقایسه با پژوهش‌های قبلی بسیار حائز اهمیت است و نشان می‌دهد که تکینیک یادگیری ژرف بکار رفته می‌تواند روند آینده قیمت سهام را به طور موثرتری پیش‌بینی کند و بهتر از سایر روش‌های سنتی یا رویکردهای یادگیری ژرف مشابه، به سرمایه گذاران در تصمیم گیری صحیح کمک نماید.

که تصویر وارد شده به لایه ورودی را به لایه خروجی انتقال می‌دهند، از عرض و ارتفاع تصویر کاسته و به عمق آن اضافه می‌شود. بعدها نشان داده شد که عملکرد این شبکه‌ها در پردازش زبان طبیعی و طبقه‌بندی جملات نیز بسیار موفق بوده است [۳۸] تا [۴۱]. همانطور که شکل ۱ روند تحلیل توسط شبکه‌های عصبی پیچشی را نشان می‌دهد، یک شبکه عصبی پیچشی از یک معماری سه لایه شامل لایه پیچشی، لایه ادغام و لایه تماماً متصل تشکیل شده است که

[۱] Computer vision

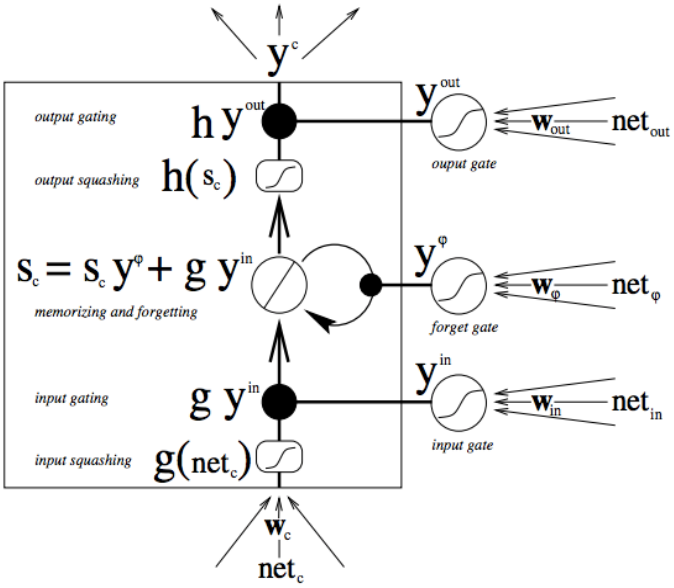
هرکدام وظایف خود را در راستای استخراج بهتر ویژگی‌ها در متن، انجام می‌دهند.

۳. ۲ شبکه برگشتی با حافظه کوتاه-مدت بلند

شبکه‌های برگشتی از واحدهای تکرارشونده تشکیل شده‌اند که قابلیت یادگیری روابط دنباله‌ای از واژه‌ها را دارد. حافظه کوتاه-مدت بلند نوعی شبکه بازگشتی عصبی است که در آن مشکل ناپدید شدن گرادیان[۱] و عدم یادگیری توالی‌های طولانی از بین رفته است و توانایی به یادسپاری اطلاعات برای بازه‌های زمانی بلندمدت را دارد [۴۲]. همانطور که در شکل ۲ مشخص شده است، این شبکه‌ها همانند هر شبکه برگشتی از واحدهای تکرارشونده تشکیل شده‌اند که هر واحد از۴ لایه تشکیل شده است.

۴ مدل پیشنهادی

در این تحقیق، از رویکردی ترکیبی مبتنی بر شبکه عصبی پیچشی و حافظه کوتاه-مدت بلند برای تحلیل احساسات نوشته‌های کاربران در شبکه‌های اجتماعی استفاده شده است. مجموعه داده به‌کار رفته در این پژوهش بر اساس داده‌های مقاله انگویان و همکاران می‌باشد که مشتمل بر دو نوع مجموعه داده برای پیش‌بینی سهام است. اولین مجموعه شامل داده‌های قیمت‌های سهام در روزهای قبل و آنها هم از سایت یاهو فایننس استخراج کرده‌اند و مجموعه دوم نظرات مردم در شبکه‌های اجتماعی را بیان می‌کند که در بخش‌های بعدی به بررسی این دو نوع مجموعه داده می‌پردازیم. باید توجه داشت مدل-های مبتنی بر شبکه‌های عصبی بخاطر ماهیتشان چندان نیاز به پیش‌پردازش ندارند، با این حال برای نرمال‌سازی داده‌ها و متناسب-سازی داده ورودی، ابتدا حذف داده‌های زائد، ایست واژه‌ها، علامت تعجب، نقطه، حروف ربط و غیره صورت می‌گیرد. در روش پیشنهادی، داده‌های ورودی به دو قسمت داده آموزشی و آزمایشی تقسیم شده است. مجموعه آموزشی به دو قسمت مجموعه آموزش و مجموعه اعتبارسنجی[۲] تقسیم می‌شوند، که در کل ۶۷٪ داده برای آموزش و اعتبارسنجی و ۳۳٪ برای آزمون استفاده شده است. رویکرد اصلی تحقیق، استفاده از شبکه‌های عصبی ژرف، ارائه یک تحلیل‌گر احساسات و کلاس‌بندی نظرات کاربران برای پیش‌بینی ارزش سهام شرکت‌های بورس در روزهای آتی از مجموعه داده ورودی است. ابتدا کلمات کل متن به بردار تبدیل شده و سپس با توجه به ارزیابی تمام بردارها، روش پیشنهادی با استفاده از چندین لایه پشت سرهم و یک روش تلفیقی از یادگیری ژرف سعی در بهبود ویژگی‌های یادگرفته شده برای پیشنهاد یک تحلیل‌گر احساسات دارد.



شکل۲. یک سلول حافظه کوتاه-مدت بلند که دارای یک واحد خطی خودبازگشتی با وزن یک (CEC) است. دروازه ورودی و خروجی وظیفه کنترل دسترسی خواندن و نوشتن روی سلول را دارد که حالت آن توسط S_c بیان می‌شود. تابع g ورودی سلول را کنترل می-کند و تابع h، خروجی سلول را کنترل می کند [۴۴]

مدل پیشنهادی یک مدل مبتنی بر شبکه عصبی ژرف شامل شبکه عصبی پیچشی و شبکه حافظه کوتاه-مدت بلند می‌باشد که شامل ۲۱ لایه است. ابتدا پیش‌پردازش‌هایی برای حذف اطلاعات زاید از ویژگی‌های ورودی انجام می‌گیرد. اطلاعات زائد نه تنها اطلاعات مفیدی در خصوص طبقه‌بندی داده‌ها تولید نمی‌کنند، بلکه می‌تواند منجر به طبقه‌بندی نادرست آن‌ها گردد و در نتیجه کارایی یادگیری

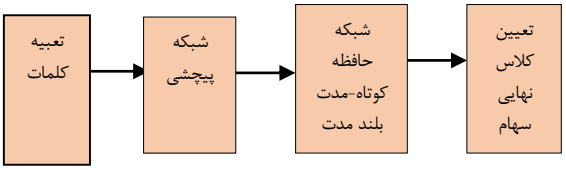
[۱] vanishing gradient

[۲] Validation

ژرف و خود شبکه حافظه کوتاه-مدت بلند کاهش می‌یابد. چرا که خروجی لایه قبل به عنوان ورودی لایه بعدی بوده و شناسایی نادرست ویژگی‌ها در هر لایه موجب آموزش نادرست داده‌ها خواهد بود. ساختار کلی مدل پیشنهادی در شکل ۳ نمایش داده شده است. همان‌طور که مشاهده می‌شود این ساختار از بخش‌های زیر تشکیل شده است و سعی دارد با یک مدل ترکیبی مشتمل بر دو شبکه عصبی ژرف موفق در حوزه پردازش متن به نام‌های شبکه پیچشی و حافظه کوتاه-مدت بلند به پیش‌بینی بازار بورس بپردازد.

اجزای اصلی ساختار پیشنهادی عبارتند از:

- تعبیه کلمات
- شبکه پیچشی
- شبکه حافظه کوتاه- مدت بلند

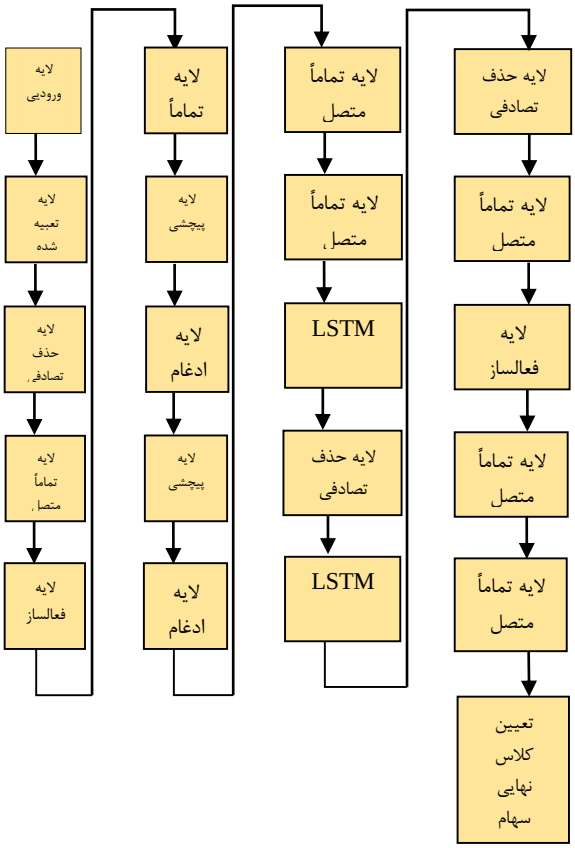


شکل ۳. ساختار کلی روش پیشنهادی

در ساختار پیشنهادی (شکل ۴) بعد از استفاده از لایه تعبیه کلمات و دو لایه پیچشی، دو لایه حافظه کوتاه-مدت بلند و در نهایت یک لایه تماماً متصل استفاده شده است. در کنار لایه‌های فوق لایه‌های دیگر از جمله لایه حذف تصادفی[1]، لایه فعال‌ساز و لایه ادغام نیز به‌کار رفته است که در نهایت، حاصل آن یک ساختار ۲۱ لایه می‌باشد که در زیر به شرح آن پرداخته می‌شود. در گام نخست مدل، به منظور انتخاب بردار کلمات از لایه ورودی مورد استفاده قرار می‌گیرد که در آن به هر واژه به یک بردار حقیقی نسبت داده می‌شود. این کار به دلیل حجم بسیار زیاد داده‌های موجود بسیار اهمیت دارد. در ادامه از یک لایه حذف تصادفی برای کنترل بیش‌برازش، لایه تماماً متصل همراه با فعال‌ساز Relu، استفاده شده‌اند. لایه پنجم یک لایه فعال‌ساز است، لایه ششم دومین لایه تماماً متصل برای بهبود عملکرد مدل است. لایه هفتم اولین لایه پیچشی برای استخراج ویژگی‌های ورودی از جملات می‌باشد و لایه هشتم لایه ادغام است که بلافاصله بعد از لایه پیچشی برای کاهش ابعاد خروجی لایه پیچشی استفاده شده است. این لایه تاثیری بر محاسبات ندارد و فقط برای کاهش ابعاد و جایگزینی حداکثر مقادیر استفاده شده است. دولایه بعد نیز به ترتیب پیچشی دوم و ادغام دوم هستند. لایه یازدهم و دوازدهم دو لایه تماماً متصل با فعال‌ساز Relu هستند. در سیزدهمین لایه مدل پیشنهادی، اولین لایه حافظه کوتاه-مدت بلند برای پردازش بازگشتی داده‌ها و بردارهای ورودی برای تحلیل احساسات بورس به‌کار رفته است. بعد از این لایه یک لایه حذف تصادفی برای بهبود یادگیری سیستم و جلوگیری از حفظ کردن

جدول۲. ورودی‌ها، خروجی‌ها و تعداد پارامترهای هر لایه در شرکت ابا ،

شبکه قرار گرفته است. ترتیب قرارگیری لایه‌های ۱۵ تا ۲۰ بدین صورت است که لایه حافظه کوتاه-مدت بلند، لایه حذف تصادفی، لایه تماماً متصل، لایه فعال‌ساز و لایه تماماً متصل قرار گرفته است. در نهایت، بیست و یکمین لایه تعیین کننده کلاس نهایی و انتخاب یکی از شش کلاس احساسی برای تحلیل احساسات در داده‌های



شکل ۴. ترتیب لایه‌های مدل شبکه عصبی بکار گرفته شده در این مطالعه

بورس است. این کلاس‌ها عبارتند از «حتما بخر»، «بخر»، «دست نگه‌دار»، «بفروش»، «حتما بفروش». همانطور که قبلاً اشاره شد پیش‌بینی این پژوهش براساس بالا یا پایین بودن سهام در روز آتی است و سرمایه‌گذاران با استفاده از نتایج حاصل از این تقسیم‌بندی، توانایی بهتری در پیش‌بینی تغییرات بازار بورس در روز آتی را خواهند داشت. به این معنا که پیش‌بینی «حتماً بخر» به مفهوم بالا بودن قیمت سهام در روز آتی بوده و خرید سهام با اطمینان بیشتری موجب سودآوری خواهد بود و پیش‌بینی «حتماً بفروش» به سرمایه‌گذار اطمینان می‌دهد که قیمت سهام در روزهای آتی افت زیادی

[1] Dropout

خواهد داشت و فروش امروز آن می‌تواند میزان ضرر سرمایه‌گذار را به حداقل برساند. گزینه‌های «بخر» و «بفروش» به ترتیب بالا و پایین بودن سهام در روز بعد را پیش‌بینی می‌کنند با این تفاوت که میزان سود و زیان سرمایه‌گذار شاید کاملاً محسوس نباشد و در گزینه «دست نگه دار»، سرمایه‌گذار بیشتر به دنبال مانیتورینگ بازار خواهد بود و از انجام هر نوع معامله در این روز خودداری خواهد کرد.

شکل ۴ ترتیب لایه‌های مدل ارائه شده مبتنی بر شبکه عصبی را نشان داده است. جدول۱ مشخصات هر لایه و نحوه تنظیم پارمترهای شبکه پیشنهادی در تمام لایه‌های آن را ارائه داده است و جدول۲ شمایی از ورودی، خروجی و تعداد پارامترهای شبکه برای شرکت

جدول ۱. نحوه تنظیم پارامتر / هایپرپارامترهای شبکه

اپل را نشان می‌دهد

	مشخصه شبکه	نحوه تنظیم پارامتر/هایپرپارامتر شبکه
C-LSTM	ورودی شبکه	طول هر رشته ورودی به شبکه ۱۵۰ عدد است
	خروجی: پیش‌بینی بازار سهام براساس تحلیل احساسات و طبقه‌بندی در ۵ کلاس	ابعاد بردار one-hot: ۵*۱
	وزن گره‌ها	یادگیری در مرحله آموزش
	ابعاد لایه تعبیه	۳۲*۱۵۰
	تعداد کل لایه‌های شبکه CLGN	۳۵ لایه
	احتمال در لایه رها کردن	۰/۲ درصد
	تابع فعالساز	ReLU
	احتمال در لایه نویزی گاوسین	۰/۱ -۰/۲
	ابعاد لایه تماماً متصل	۱۲۸- ۶۴- ۳۲- ۵
CNN	ورودی لایه کانولوشن	۳۲-۶۴
	ابعاد لایه کانولوشن اول	۳۲*۱۵۰
	تعداد فیلترها در لایه کانولوشن ۱ و ۲	filters=۳۲ or ۶۴
	اندازه هر فیلتر در لایه کانولوشن ۱ و ۲	kernel_size=۲
	تعداد کلمات ادغامی در لایه ماکزیمم ادغام ۱و۲	pool_size:۲
	ابعاد لایه ماکزیمم ادغام در لایه کانولوشن اول	۳۲*۷۵
	تابع فعالساز کانولوشن ۱ و ۲	ReLU
	ابعاد لایه کانولوشن دوم	۶۴*۷۵
	ابعاد لایه ماکزیمم ادغام	۶۴*۳۷
LSTM	ورودی لایه LSTM	۱۲۸
	ابعاد لایه LSTM اول	۱۲۸*۳۷
	ابعاد لایه LSTM دوم	۱۲۸
	تعداد تکرارهای شبکه[۱]	۱۰
	تابع فعالساز	sigmoid-tanh
	تابع بهینه ساز	آدام[۲]
	تابع زیان	binary_crossentropy

Layer (type)	Output Shape	Param #
embedding_۲ (Embedding)	(۱۵۰, ۳۲)	۱۰۸۲۹۱۲
dropout_۷ (Dropout)	(۱۵۰, ۳۲)	۰
dense_۷ (Dense)	(۱۵۰, ۱۲۸)	٤۲۲٤
activation_٥ (Activation)	(۱۵۰, ۱۲۸)	۰
dense_۸ (Dense)	(۱۵۰, ٦٤)	۸۲٥٦
conv۱d_۳ (Conv۱D)	(۱۵۰, ۳۲)	٤۱۲۸
)max_pooling۱d_۳ (MaxPooling	(۷٥, ۳۲)	۰
conv۱d_٤ (Conv۱D)	(۷٥, ٦٤)	٤۱٦۰
)max_pooling۱d_٤ (MaxPooling	(۳۷, ٦٤)	۰
dense_۹ (Dense)	(۳۷, ۳۲)	۲۰۸۰
activation_٦ (Activation)	(۳۷, ۳۲)	۰
dense_۱۰ (Dense)	(۳۷, ۳۲)	۱۰٥٦
lstm_۳ (LSTM)	(۳۷, ۱۲۸)	۸۲٤۳۲
dropout_۸ (Dropout)	(۳۷, ۱۲۸)	۰
lstm_٤ (LSTM)	(۱۲۸)	۱۳۱٥۸٤
dropout_۹ (Dropout)	(۱۲۸)	۰
dense_۱۱ (Dense)	(۳۲)	٤۱۲۸
activation_۷ (Activation)	(۳۲)	۰
dense_۱۲ (Dense)	(۳۲)	۱۰٥٦
dense_۱۳ (Dense)	(٥)	۱٦٥
	Total params: ۱،۳۲٦،۱۸۱ Trainable params: ۱،۳۲٦،۱۸۱ Non-trainable params: ۰	

[۱] number of epochs
[۲] Adaptive Moment Estimation (Adam)

۵ دادگان پژوهش و نتایج

در مطالعه حاضر، مدلی مبتنی بر رویکرد هوش مصنوعی و با به‌کارگیری دو نوع شبکه عصبی پیچشی و حافظه کوتاه-مدت بلند برای ارزیابی و پیش‌بینی تغییرات قیمت‌های بورس برای ۱۸ شرکت بین‌المللی تهیه شده است. در این راستا از مجموعه نوسانات قیمت سهام در روزهای قبل و نظرات کاربران ارائه شده در شبکه‌های اجتماعی برای ۱۸ سهام مندرج بهره گرفته شده است. برای هر تاریخ معامله ۵ قیمت شامل قیمت آغازین، بالاترین قیمت، پایین‌ترین قیمت، قیمت پایانی و قیمت تنظیم‌شده پایانی در نظر گرفته می‌شود. قیمت تنظیم شده پایانی همان قیمت پایانی است که براساس سود و زیان سهام تنظیم می‌شود و در برخی مقالات استفاده شده است و در این پژوهش به عنوان ارزش قیمت سهام در تاریخ معامله انتخاب می‌شود جدول ۳ تعداد متون، کلمات، داده‌های آموزشی، داده‌های تست و داده‌های ارزیابی را نشان می‌دهد.

در تابلوی پیام، کاربران به بحث در مورد اخبار شرکت‌ها، پیش‌بینی در مورد بالا و پایین بودن سهام، حقایق و پیام‌های مربوط به مدیران شرکت‌ها می‌پردازند. در ۱۵/۶٪ از پیام‌های این مجموعه داده، زمانیکه کاربـران پیام‌ها را در این تابلـوها ارسال کردند هر پیام را با یکی از گزینه‌های احساسی «حتما بخر»، «بخر»، «دست نگه‌دار»، «بفروش»، «حتماً بفروش» برچسب‌گذاری نمودند که ما در این پژوهش از آنها برای آموزش شبکه استفاده می‌کنیم ولی برای تعبیه کلمات از کل مجموعه داده (برچسب‌دار و غیر برچسب‌دار) استفاده شده است. با توجه به اینکه بورس در روزهای آخر هفته و تعطیلات رسمی بسته است بنابراین تعداد کل روزهای کاری بورس در دوره یکساله مذکور، ۲۴۹ روز می‌باشد. همچنین شرکت اپل دارای بیشترین پیامها در سایت‌های اینترنتی بوده و حجم وسیعی از داده‌ها را در برگرفته است، بنابراین برای این شرکت تنها دوره هفت ماهه به منظور جمع آوری داده‌ها در نظر گرفته شده بود. برای ارزیابی روش پیشنهادی، صحت مدل پیشنهادی با روش‌های ارائه شده توسط انگویان و همکارانش و همچنین درخشان و همکاران به‌عنوان روش‌های مطرح پایه مورد توجه قرار می‌گیرند [۱۶] و [۲۳].

معیار صحت به صورت زیر تعریف شده است:

$$Accuracy = \frac{Number\ of\ truly\ predicted\ classes}{Total\ number\ of\ instances} \qquad (۱)$$

برای ارزیابی روش پیشنهادی، همچنین در کنار این روش‌ها، از نتایج حاصل از روش‌های تحلیل انسانی و روش تخصیص دیریکله نیز بهره گرفته شده است. همچنین در یک مدل پایه ساده با بهره‌گیری از تحلیل قیمت به تنهایی نتایج به دست آمده مورد مقایسـه قرار می‌گیـرد. برای سنجـش میزان کارآمدی مدل، صحت روش پیشنهادی با روش‌های انگویان و همکاران و درخشان و همکاران که

جدول ۳. تعداد متون، کلمات، داده‌های آموزشی، داده‌های تست و داده‌های ارزیابی

سهام	تعداد متون	تعداد کلمات	تعداد داده‌های آموزشی	تعداد داده‌های تست	تعداد داده‌های ارزیابی
AAPL	۴۸۸۶۲	۱۹۰۶۸۴۵	۱۲۷۷۵۸۶	۶۲۹۲۵۸	۱۲۷۷۵
AMZN	۶۰۲۴	۲۰۷۳۷۵	۱۳۸۹۴۱	۶۸۴۳۳	۱۳۸۹۴
BA	۳۳۷۸	۱۷۱۵۱۸	۱۱۴۹۱۷	۵۶۶۰۰	۱۱۴۹۱
BAC	۹۹۳۵	۲۴۴۱۹۲	۱۶۳۶۰۸	۸۰۵۸۳	۱۶۳۶۰
CSCO	۲۰۲۵	۷۳۱۸۲	۴۹۰۳۱	۲۴۱۵۰	۴۹۰۳
DELL	۱۸۰۷	۵۱۰۶۱	۳۴۲۱۴	۱۶۸۴۶	۳۴۲۱
EBAY	۵۶۷	۱۷۳۶۰	۱۱۶۳۱	۵۷۲۸	۱۱۶۳
ETFC	۲۴۷۲	۱۰۹۹۹۹	۷۳۶۹۹	۳۶۲۹۹	۷۳۶۹
GOOG	۳۵۶۳	۹۳۱۱۱۴	۶۲۳۸۴۶	۳۰۷۲۶۷	۶۲۳۸۴
IBM	۷۰۸	۱۸۰۱۳	۱۲۰۶۸	۵۹۴۴	۱۲۰۶
INTC	۵۸۷۸	۳۶۸۲۱۲	۲۴۶۷۰۲	۱۲۱۵۰۹	۲۴۶۷۰
KO	۴۶۳	۹۶۰۰	۶۴۳۲	۳۱۶۸	۶۴۳
MSFT	۱۰۵۷۰	۴۰۲۶۱۹	۲۶۹۷۵۴	۱۳۲۸۶۴	۲۶۹۷۵
NVDA	۲۳۵۹	۱۵۲۸۰۶	۱۰۲۳۸۰	۵۰۴۲۶	۱۰۲۳۸
ORCL	۱۱۵۸	۴۲۶۱۱	۲۸۵۴۹	۱۴۰۶۱	۲۸۵۴
T	۱۶۸۱	۸۳۳۰۳	۵۵۸۱۳	۳۲۴۸۹	۵۵۸۱
XOM	۷۰۲	۳۶۵۰۲	۲۴۴۵۶	۱۲۰۴۵	۲۴۴۵
YHOO	۵۵۰۰	۱۸۴۵۱۶	۱۲۳۶۲۵	۶۰۸۹۰	۱۲۳۶۲
مجموع	۱۰۷۶۵۲	۵۰۱۸۲۸	۳۳۵۷۲۵۲	۱۶۵۸۵۶۰	۲۲۰۷۳۴

برای این مجموعه داده ارائه شده است، مورد بررسی قرار گرفته است [۱۶] و [۲۳].

در جدول ۴ نتایج حاصل از ارزیابی صحت مدل و مقایسه آن با نتایج انگویان و شبکه ژرف پیچشی نشان داده شده است. همانطور که مشاهده می‌گردد، مدل ارائه شده یک مدل کارامد می‌باشد که توانسته است کلاس داده‌های آزمایش را با دقت بیشتری در مقایسه با سایر مدل‌ها پیش‌بینی کند و توانسته است به‌طور میانگین با موفقیت بیشتری عمل کند.

با بررسی جدول ۴ که مقایسه معیار صحت مدل پیشنهادی با ۶ روش دیگر است، به این نتیجه می‌رسیم که مدل پیشنهادی در ۱۸ شرکت مورد بررسی، دارای بالاترین مقدار صحت بین ۶ روش مقایسه‌ای می‌باشد. همچنین میانگین صحت مدل پیشنهادی بهتر از سایر مدل‌ها بوده و با اختلاف ۱۹/۷۹۲ و ۲۴/۵ درصدی قوی‌تر از مدل شبکه ژرف پیچشی و مدل پیشنهادی انگویان و همکارانش عمل می‌کند. اگرچه ترکیب شبکه‌های مختلف در یادگیری ژرف در کارهای مشابه و بر روی تحلیل احساسات انجام شده است اما کاربرد چنین شبکه‌هایی در پیش‌بینی سهام بورس شرکت‌ها نبوده است. عملکرد خوب این رویکرد ترکیبی را می‌توان ناشی از تنظیم درست وزن‌های ورودی لایه‌های بعد، انتخاب درست چیدمان لایه‌ها، افزایش نرخ یادگیری و یادگیری بهتر دانست. در گام دیگر برای مقایسه مدل پیشنهادی با نتایج درخشان و همکاران، با توجه به اینکه در مقاله مربوطه ۱۵ سهام از مجموعه سهام‌های ارائه شده مورد بررسی قرار گرفته‌است، ما نیز نتایج خود را به این ۱۵ سهام محدود نمودیم تا بتوان در شرایط مساوی به مقایسه روش‌ها پرداخت (جدول ۵). در این بخش از مقایسه نیز به نتایج تقریباً مشابهی دست یافتیم و مدل پیشنهادی ما بهتر از ۲ روش ارائه شده در مقاله درخشان در مقایسه با دیگر روش‌ها قابل ملاحظه بوده و در معیار ارزیابی صحت با بهبود ۱۹/۸۰ درصدی نسبت به مدل شبکه پیچشی ژرف، ۲۴/۵ درصدی نسبت به مدل پیشنهادی انگویان و همکاران و ۲۳/۹۴ درصدی نسبت به مدل پیشنهادی درخشان و همکاران پیشی گرفته است.

۶ نتیجه گیری

در این تحقیق با پیاده‌سازی مبتنی بر شبکه عصبی ژرف پیچشی و شبکه برگشتی حافظه کوتاه-مدت بلند به منظور تحلیل احساسات در بررسی سهام بورس ۱۸ شرکت ورودی از مجموعه داده پرداختیم. روش پیشنهادی، از یک مدل ۲۱ لایه‌ای تشکیل شده است که در ابتدا کلمات نوشته‌های کاربران در شبکه‌های اجتماعی به بردار تبدیل می‌شود و سپس با استفاده از یک روش تلفیقی از یادگیری ژرف سعی در بهبود ویژگی‌های یادگرفته شده برای پیشنهاد احساسات مناسب نوشته‌های کابران می‌نماید. خروجی این معماری سپس به‌عنوان پیش‌بینی بازار سهام مورد ارزیابی قرار گرفته است. مدل

جدول ۴. مقایسه معیار صحت مدل پیشنهادی با پیچشی ژرف و ۵ روش انگویان و همکاران

نام شرکت	CNN	CNN-LSTMروش پیشنهادی	فقط قیمت	تحلیل انسانی	JST	LDA	انگویان
AMZN	۲۵/۷۲	۸۱/۸۱	۴۶/۰۵	۴۸/۶۸	۵۹/۲۱	۵۱/۳۲	۷۱/۰۵
APPL	۴۲/۳۴	۵۸/۳۴	۳۹/۵۱	۵۶/۷۹	۵۸/۰۲	۵۸/۰۲	۵۴/۳۲
BA	۶۲/۰۳	۸۰	۶۳/۱۶	۶۰/۵۳	۶۳/۱۶	۵۵/۲۶	۵۹/۲۱
BAC	۴۸/۹۶	۷۱/۵۱	۵۶/۵۸	۵۹/۲۱	۵۶/۵۸	۵۵/۲۶	۴۴/۷۴
CSCO	۶۰/۱۰	۸۰/۶۲	۵۵/۲۶	۴۴/۷۴	۵۱/۳۲	۴۷/۳۷	۴۶/۰۵
DELL	۵۶/۰۶	۷۹/۴۹	۵۳/۹۵	۵۹/۲۱	۴۳/۴۲	۵۱/۳۲	۶۴/۴۷
EBAY	۷۷/۷۹	۸۰/۸۴	۵۹/۲۱	۴۶/۰۵	۴۰/۷۹	۵۶/۵۸	۵۷/۸۹
ETFC	۷۴/۳۹	۸۱/۵۱	۵۷/۸۹	۵۹/۲۱	۴۳/۴۲	۴۸/۶۸	۵۵/۲۶
GOOG	۴۵/۳۱	۸۰	۵۰	۵۶/۵۸	۵۳/۹۵	۵۶/۵۸	۵۲/۶۳
IBM	۵۶/۵۱	۷۹/۱۱	۴۸/۶۸	۴۷/۳۷	۴۴/۷۴	۵۳/۹۵	۵۵/۲۶
INTC	۶۱/۸۵	۸۱/۸۴	۴۴/۷۴	۴۶/۰۵	۴۸/۶۸	۵۰	۵۲/۶۳
KO	۵۷/۷۸	۷۹/۸۱	۴۰/۷۹	۴۸/۶۸	۵۱/۳۲	۵۶/۵۸	۴۴/۷۴
MSFT	۵۷/۷۹	۸۳/۲۷	۵۷/۸۹	۶۵/۷۹	۵۵/۲۶	۵۵/۲۶	۵۲/۶۳
NVDA	۶۹/۹۷	۷۹/۸۱	۶۰/۵۳	۵۷/۸۹	۵۰	۳۹/۴۷	۵۳/۹۵
ORCL	۶۵/۴۲	۸۰/۷۷	۴۸/۶۸	۵۲/۶۳	۵۰	۵۲/۶۳	۵۳/۹۵
YHOO	۶۳/۸۹	۸۷/۱۰	۵۵/۲۶	۵۹/۲۶	۴۴/۷۴	۵۲/۶۳	۵۵/۲۶
T	۶۱/۱۴	۷۵/۷۲	۵۵/۲۶	۴۷/۳۷	۵۶/۵۸	۵۰	۵۱/۳۲
XOM	۷۶/۹۰	۷۸/۸۱	۴۸/۶۸	۶۴/۴۷	۵۶/۵۸	۴۳/۴۲	۵۳/۹۵
میانگین	**۵۹/۱۰۸**	**۷۸/۹۰**	**۵۲/۳۴**	**۵۴/۴۷**	**۵۱/۵۴**	**۵۱/۹۰**	**۵۴/۴۰**

جدول ۵. مقایسه معیار صحت مدل پیشنهادی با دو روش درخشان و همکاران

نام شرکت	CNN-LSTM روش پیشنهادی	LDA-POS	مدل زبانی روی LDA
AMZN	۸۱/۸۱	۵۷/۶۹	۳۸/۴۶
BA	۸۰	۶۱/۵۴	۵۵/۱۲
BAC	۷۱/۵۱	۵۸/۸۹۷	۶۱/۵۳
CSCO	۸۰/۶۲	۵۵/۱۳	۵۳/۸۴
EBAY	۸۰/۸۴	۵۱/۲۸	۵۳/۸۴
ETFC	۸۱/۵۱	۵۵/۱۳	۵۵/۱۲
GOOG	۸۰	۵۳/۸۵	۵۵/۱۲
IBM	۷۹/۱۱	۵۵/۱۳	۴۷/۴۳
INTC	۸۱/۸۴	۵۷/۶۹	۴۱/۰۲
MSFT	۸۳/۲۷	۵۳/۸۵	۴۷/۴۳
NVDA	۷۹/۸۱	۵۳/۸۵	۴۸/۷۱
ORCL	۸۰/۷۷	۵۱/۲۸	۵۰
T	۷۵/۷۲	۵۵/۱۳	۵۵/۱۲
XOM	۷۸/۸۱	۶۰/۲۶	۵۷/۶۹
YAHOO	۸۷/۱۰	۶۲/۸۲	۵۳/۸۴
میانگین	**۸۰/۱۸۱۳**	**۵۶/۲۴**	**۵۱/۶۲**

پیشنهادی یک مدل مبتنی بر شبکه عصبی پیچشی و شبکه برگشتی حافظه کوتاه-مدت بلند می‌باشد. ابتدا پیش پردازش‌هایی برای حذف اطلاعات زاید از ویژگی‌های ورودی که اطلاعات بیشتری در خصوص دسته‌بندی تولید نمی‌کنند، انجام شده است. با توجه به کارکرد یادگیری ژرف و شبکه برگشتی حافظه کوتاه-مدت بلند، خروجی هر لایه ورودی لایه بعدی بوده و اگر در هر لایه شناسایی ویژگی نادرست آموزش داده شود در نهایت شناسایی دقیق نخواهد بود. مطالعه حاضر با در نظر گرفتن این مهم، سعی نموده است تا با بکارگیری شبکه عصبی ژرف و مدل‌های پیچشی اقدام به پیش‌بینی تغییرات بازار بورس شرکت های بین‌المللی نموده و مدلی کارامد در این زمینه پیشنهاد نماید. برپایه نتایج حاصل از ارزیابی، مدلی ژرف مشتمل بر ۲۱ لایه شبکه عصبی پیچشی و شبکه برگشتی حافظه کوتاه-مدت بلند که برای ۱۸ شرکت وارد شده در بورس پیاده‌سازی شده، مشخص گردیده است که نتایج حاصل با ۷ روش تحلیل دیگر مورد سنجش قرار گرفته است. مدل‌سازی نشان داده که عملکرد روش پیشنهادی در مقایسه با دیگر روش‌ها قابل ملاحظه بوده و در معیار ارزیابی صحت با بهبود ۱۹/۸۰ درصدی نسبت به مدل شبکه پیچشی ژرف، ۲۴/۵۰ درصدی نسبت به مدل پیشنهادی انگویان و همکاران و ۲۳/۹۴ درصدی نسبت به مدل پیشنهادی درخشان و همکاران پیشی گرفته است همانطور که ذکر شد، تمامی نتایج ارائه شده در این پژوهش بر مبنای تحلیل نظرات کاربران بوده است و بهره‌گیری از تاریخچه قیمت سهام‌ها به عنوان یک فاکتور تاثیرگذار دیگر بخشی از کارهای آتی این پژوهش می‌باشد.

مراجع

[۱] Roberts, H.V., ۱۹۵۹. Stock-market" patterns" and financial analysis: methodological suggestions. *The Journal of Finance*, ۱٤(۱), pp.۱-۱۰.

[۲] Salloum, S.A., Khan, R. and Shaalan, K., ۲۰۲۰, April. A survey of semantic analysis approaches. In *Joint European-US Workshop on Applications of Invariance in Computer Vision* (pp. ٦۱-۷۰). Springer, Cham.

[۳] Liu, B., ۲۰۲۰. Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge university press.

[۴] Rambocas, M. and Pacheco, B.G. ۲۰۱۸, "Online sentiment analysis in marketing research: a review", *Journal of Research in Interactive Marketing*, ۱۲(۲), pp ۱٤٦–۱٦۳.

[۵] Habimana, O., Li, Y., Li, R., Gu, X. and Yu, G., ۲۰۲۰. Sentiment analysis using deep learning approaches: an overview. *Science China Information Sciences*, ٦۳(۱), pp.۱-۳٦.

[۶] Zhang, L., Wang, S. and Liu, B., ۲۰۱۸. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, ۸(٤), p.e۱۲۵۳

[۷] Prabha, M.I. and Srikanth, G.U., *Survey of sentiment analysis using deep learning techniques*. In ۲۰۱۹ ۱st International Conference on Innovations in Information and Communication Technology (ICIICT) (pp. ۱-۹). IEEE, ۲۰۱۹, April..

[۸] Yadav, A. and Vishwakarma, D.K., ۲۰۲۰. Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, ۵۳(٦), pp.٤۳۳۵-٤۳۸۵.

[۹] Chan, W.S., ۲۰۰۳. Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, ۷۰(۲):۲۲۳–۲٦۰.

[۱۰] Bollen, J., Mao, H. and Zeng, X., ۲۰۱۱. Twitter mood predicts the stock market. *Journal of computational science*, ۲(۱):۱–۸.

[۱۱] Luo, X., Zhang, J. and Duan, W., ۲۰۱۳. Social media and firm equity value. *Information Systems Research*, ۲٤(۱), pp.۱٤٦-۱٦۳

[۱۲] Yu, Y., Duan, W. and Cao, Q., ۲۰۱۳. The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision support systems*, ٥٥(٤), pp.۹۱۹-۹۲٦.

[۱۳] Li, Q., Wang, T., Li, P., Liu, L., Gong, Q. and Chen, Y., ۲۰۱٤. The effect of news and public mood on stock movements. *Information Sciences*, ۲۷۸, pp.۸۲٦-۸٤۰

[۱۴] Oh, C. & Sheng, O. (۲۰۱۱), Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement., in Dennis F. Galletta & Ting-Peng Liang, ed., 'ICIS' , Association for Information Systems.

[۱۵] Ruan, Y., Durresi, A. and Alfantoukh, L., ۲۰۱۸. Using Twitter trust network for stock market analysis. *Knowledge-Based Systems*, ۱٤٥, pp.۲۰۷-۲۱۸.

[۱۶] Nguyen, T. H., Shirai, K., & Velcin, J., ۲۰۱٥. Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, ٤۲(۲٤), ۹٦۰۳-۹٦۱۱.

[۱۷] Chong, E., Han, C. and Park, F.C., ۲۰۱۷. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, ۸۳, pp.۱۸۷-۲۰٥.

[۱۸] Chen, W., Zhang, Y., Yeo, C.K., Lau, C.T. and Lee, B.S., *Stock market prediction using neural network through news on online social networks*. In ۲۰۱۷ international smart cities conference (ISC۲) (pp. ۱-٦). IEEE, ۲۰۱۷, September.

[۱۹] Liu, Y., Qin, Z., Li, P. and Wan, T., *Stock volatility prediction using recurrent neural networks with sentiment analysis.* In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. ۱۹۲-۲۰۱). Springer, Cham, ۲۰۱۷, June.

[۲۰] نصرااله پور، ف.، بحرانی، م.، بیجن خان، م. (۱۳۹۸). استفاده از روش‌های یادگیری ماشین جهت پیش‌بینی نوسانات نرخ ارز در متون خبری اقتصادی فارسی. اولین همایش ملی هوش مصنوعی و محاسبات نرم در علوم انسانی، مقاله کامل چاپ شده،

[۲۱] Sadr, H., Pedram, M.M. and Teshnehlab, M., ۲۰۱۹. A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks. *Neural Processing Letters*, ٥۰(۳), pp.۲۷٤٥-۲۷٦۱

[۲۲] Sadr, H., Pedram, M.M. and Teshnehlab, M., ۲۰۲۰. Multi-View Deep Network: A Deep Model Based on Learning Features From Heterogeneous Neural Networks for Sentiment Analysis. *IEEE Access*, ۸, pp.۸٦۹۸٤-۸٦۹۹۷.

[۲۳] Derakhshan, A. and Beigy, H. ۲۰۱۹. Sentiment analysis on stock social media for stock price movement prediction. Engineering Applications of Artificial Intelligence ۸٥ (۲۰۱۹): ٥٦۹-٥۷۸

[۲۴] Xu, Y. and Keselj, V., *Stock Prediction using Deep Learning and Sentiment Analysis. In ۲۰۱۹ IEEE International* Conference on Big Data (Big Data) (pp. ٥٥۷۳-٥٥۸۰). IEEE, ۲۰۱۹, December

[۲۵] Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P. and Anastasiu, D.C., *Stock Price Prediction Using News Sentiment Analysis*. In ۲۰۱۹ IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService) (pp. ۲۰٥-۲۰۸). IEEE, ۲۰۱۹, April..

[۲۶] Shi, Y., Zheng, Y., Guo, K. and Ren, X., ۲۰۲۰. Stock movement prediction with sentiment analysis based on deep learning

networks. *Concurrency and Computation: Practice and Experience*, p.e٦٠٧٦.

[٢٧] Chun, J., Ahn, J., Kim, Y. and Lee, S., ٢٠٢٠. Using Deep Learning to Develop a Stock Price Prediction Model Based on Individual Investor Emotions. Journal of Behavioral Finance, pp.١-١٠.

[٢٨] Li, Y. and Pan, Y., ٢٠٢٠. A novel ensemble deep learning model for stock prediction based on stock prices and news. *arXiv preprint arXiv:* ٢٠٠٧,١٢٦٢٠.

[٢٩] Mohamed Ali, N., Abd El Hamid, M. M. and Youssif, A. ٢٠١٩. Sentiment analysis for movies reviews dataset using deep learning modes, *International Journal of Data Mining & Knowledge Management Process (IJDKP)* Vol.٩, No.٢/٣.

[٣٠] Ramadhani, A.M. and Goo, H.S. ٢٠١٧. [IEEE ٢٠١٧ ٧th International Annual Engineering Seminar (InAES) - Yogyakarta, Indonesia (٢٠١٧,٨,١-٢٠١٧,٨,٢)] *٢٠١٧ ١th International Annual Engineering Seminar (InAES) - Twitter sentiment analysis using deep learning methods*. , (), ١–٤.

[٣١] Goularas, D. and Kamis, S., *Evaluation of deep learning techniques in sentiment analysis from Twitter data*. In ٢٠١٩ International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML) (pp. ١٢-١٧). IEEE, ٢٠١٩, August.

[٣٢] Jin, Z., Yang, Y., & Liu, Y. (٢٠١٩). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications*. doi:١٠,١٠٠٧/s٠٠٥٢١-٠١٩-٠٤٥٠٤-٢

[٣٣] Akhtar, M.S., Kumar, A., Ekbal, A. and Bhattacharyya, P., *A hybrid deep learning architecture for sentiment analysis.* In Proceedings of COLING ٢٠١٦, the ٢٦th International Conference on Computational Linguistics: Technical Papers (pp. ٤٨٢-٤٩٣), ٢٠١٦, December.

[٣٤] Zhou Y., ٢٠٢٠. *A Review of Text Classification Based on Deep Learning*. In Proceedings of the ٢٠٢٠ ٣rd International Conference on Geoinformatics and Data Analysis (ICGDA ٢٠٢٠). Association for Computing Machinery, New York, NY, USA, ١٣٢–١٣٦. DOI:https://doi.org/١٠,١١٤٥/٣٣٩٧٠٥٦,٣٣٩٧٠٨٢

[٣٥] Zhou, C.; Sun, C.; Liu, Z. & Lau, F. C. M. (٢٠١٥), 'A C-LSTM Neural Network for Text Classification.' CoRR abs/١٥١١,٠٨٦٣٠

[٣٦] Sukheja, S., Chopra, S. and Vijayalakshmi, M., ٢٠٢٠, March. *Sentiment Analysis using Deep Learning–A survey*. In ٢٠٢٠ International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. ١-٤). IEEE.

[٣٧] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., ١٩٩٨. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, ٨٦(١١), pp.٢٢٧٨-٢٣٢٤.

[٣٨] Kim, Y., *Convolutional neural networks for sentence classification*, In Proceedings of the ٢٠١٤ Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar. Association for Computational Linguistics. PP: ١٧٤٦-١٧٥١, ٢٠١٤.

[٣٩] Wang, P., Xu, J., Xu, B., Liu, C., Zhang, H., Wang, F. and Hao, H. *Semantic clustering and convolutional neural network for short text categorization*. In Proceedings of the ٥٣rd Annual Meeting of the Association for Computational Linguistics and the ٧th International Joint Conference on Natural Language Processing (Volume ٢: Short Papers) (pp. ٣٥٢-٣٥٧), ٢٠١٥, July.

[٤٠] Zhang, Y., Roller, R. and Wallace, B. C., *Mgnc-cnn: A simple approach to exploiting multiple word embeddings for sentence classification*, Proceedings of the ٢٠١٦ Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, pp.١٥٢٢-١٥٢٧, ٢٠١٦.

[۴۱] Samat, N.A., Salleh, M.N.M. and Ali, H., ۲۰۲۰, January. *The Comparison of Pooling Functions in Convolutional Neural Network for Sentiment Analysis Task*. In International Conference on Soft Computing and Data Mining (pp. ۲۰۲-۲۱۰). Springer, Cham

[۴۲] Hochreiter, S. and Schmidhuber, J., ۱۹۹۷. Long short-term memory. Neural computation, ۹(۸), pp.۱۷۳۵-۱۷۸۰.

[۴۳] Zhang, Ye, and Byron Wallace. "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification." arXiv preprint arXiv: ۱۵۱۰,۰۳۸۲۰ (۲۰۱۵).

[۴۴] Gers, F.A., Schmidhuber, J. and Cummins, F., ۱۹۹۹. Learning to forget: Continual prediction with LSTM.