

تحلیل احساس در نظرات مشتریان صنعت داروسازی با استفاده از یادگیری عمیق

سید شایان طباطبائی*، مهدی براتی، مریم بیاتی

Shayantabatabaei@gmail.com

*نویسنده مسئول، دریافت: ۱۴۰۱/۸/۱۵، بازنگری: ۱۴۰۱/۹/۲۴، پذیرش: ۱۴۰۱/۱۰/۵

کارشناسی ارشد مهندسی فناوری اطلاعات، گرایش تجارت الکترونیک، موسسه عالی علوم و فناوری سپاهان، اصفهان، ایران
استادیار، دانشکده مهندسی کامپیوتر، موسسه آموزش عالی راغب، تهران، ایران
استادیار، موسسه موسسه عالی علوم و فناوری سپاهان، تهران، ایران

چکیده

شبکه های اجتماعی به عنوان مهمترین عضو دنیای مجازی برای جمع آوری منبع عظیمی از داده های غیر ساخت یافته و دانش پنهان وجود دارند. نتایج کاربردی حاصل از تحلیل و بررسی احساسات پنهان در متون غیر ساختار یافته کاربران، باعث توجه بیش از پیش سازمان های مالی و تجاری مختلف به این حوزه شده است. این مسئله باعث معرفی حوزه تحلیل احساس مبتنی بر تکنیک های یادگیری ماشین شد. مطالعات مبتنی بر شبکه های عصبی عمیق به صورت انفرادی و یا ترکیبی هستند. در بسیاری از روش ها دو شبکه عمیق به صورت مجزا ترکیب شده اند که نتایج موافقی ارائه نداده اند. در این تحقیق در راستای تحلیل احساس در حوزه داروسازی، ترکیب دو شبکه عصبی عمیق کانولوشنی و شبکه عصبی بازگشتی GRU در یک معماری واحد، انجام شده است. ترکیب مناسب لایه های دو شبکه فوق در یک معماری عمیق و همچنین استفاده از شبکه تعبیه کلمات پیش آموزش دیده FastText باعث شده معماری پیشنهادی در مقایسه با معماری های ترکیبی دیگر به نتایج موفق تری دست یابد. نتایج تجربی معماری عمیق ترکیبی پیشنهادی بر روی داده های Drug reviews نشان می دهد که مدل پیشنهادی توانسته به صحت ۹۰.۷۰٪ دست یابد که در مقایسه با روش های پیشین در حدود ۲.۲٪ صحت تحلیل احساس را در صنعت دارو بهبود بخشیده است. کلمات کلیدی: تحلیل احساس، صنعت داروسازی، احساسات پنهان، شبکه عصبی، تحلیل احساسات در صنعت داروسازی.

۱- مقدمه

ترجیحات سیاسی و رویدادهای اجتماعی دارد. برای مثال، تحلیل نظرات کاربران شبکه اجتماعی توییتر در پیش بینی محبوبیت احزاب سیاسی موثر است. امروزه تجزیه و تحلیل احساسات^۱ تبدیل به ابزاری قدرتمند برای مشاغل، دولت ها و محققان برای استخراج و تجزیه و تحلیل دیدگاه های عمومی، برای حصول بینش تجاری و تصمیم گیری بهتر شده است. اغلب روش هایی که در این حوزه تاکنون معرفی شده اند، مبتنی بر روش های یادگیری ماشین و یادگیری عمیق در شبکه های اجتماعی توییتر، فیسبوک هستند و مطالعات محدودی در این زمینه در حوزه دارو انجام شده است [۱]. از جمله آخرین پژوهش هایی که کاربرد تحلیل احساس را در صنعت دارو بررسی نموده است روش کلن رویز و همکاران در سال ۲۰۲۰ است [۲]. در این روش نظرات بیماران در

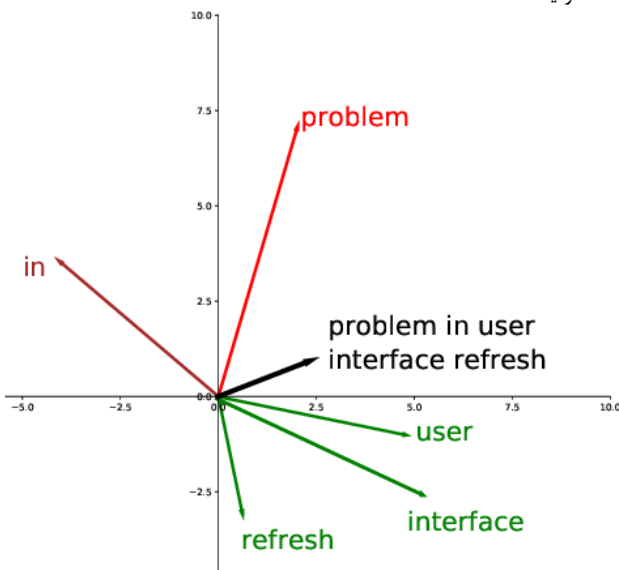
با افزایش روزانه کاربران شبکه های اجتماعی، ثبت نظرات آن ها در خصوص انواع خدمات و محصولات افزایش یافته است. این نظرات را می توان به عنوان منبعی از اطلاعات پنهان در نظر گرفت، که تحلیل و استخراج دانش از آن ها می تواند در راستای بهبود کیفیت محصولات و خدمات استفاده گردد. فواید متعدد حاصل از تحلیل نظرات کاربران و عدم امکان تحلیل دستی و سنتی آن ها، باعث شده تکنیک های تحلیل احساس بیش از پیش مورد توجه محققان و تحلیل گران واقع شود. نظرات کاربران تاثیر مستقیم در بهبود کیفیت محصولات، خدمات، بازاریابی،

^۱ sentiment analysis

آن‌ها یعنی شبکه‌های عصبی و شبکه‌های عصبی عمیق انجام شد. پس از بررسی و مطالعه روش‌های مختلف و بررسی محدودیت‌های آن‌ها راهکاری جهت تحلیل احساس در متون دارویی در قالب پروپزال مطرح شد. سپس مجموعه داده‌های مورد نیاز برای ارزیابی روش پیشنهادی با توجه به مجموعه داده استفاده شده در روش پایه [۲] جمع‌آوری شد. در ادامه پس از تایید پروپزال پیاده‌سازی ایده پیشنهادی در محیط متلب انجام پذیرفت و پس از پایان پیاده‌سازی ارزیابی و مقایسه روش‌ها بر روی پایگاه داده دریافت شده انجام شد و نتایج مدل‌ها با توجه متغیرهای وابسته این تحقیق مقایسه شدند.

شبکه تعبیه کلمات

تکنیک تعبیه کلمات به فرآیندی اطلاق می‌شود که طی آن برای هر کلمه در متن یک بردار با مقادیر واقعی تشکیل می‌شود که این بردار و مقادیر آن نشان دهنده کلمه مورد نظر بر اساس محتوایی است که در آن حضور دارد. تشکیل بردار اعداد برای هر کلمه، امکان نگاشت آن را، به یک فضای برداری فراهم می‌کند. این مفهوم در شکل ۱ نشان داده شده است. در این تکنیک بردارهای تشکیل شده برای کلمات مشابه و نزدیک بهم که از نظر معنایی یا نحوی به یکدیگر مشابه‌تر هستند بهم نزدیک‌تر است و در مقابل بردار کلمات غیر مشابه نیز از هم دورتر است [۳]. یکی از اصلی‌ترین مزایای تکنیک تعبیه کلمات این است که، می‌توان آن را مستقیماً از یک مجموعه متن یاد گرفت و به هیچ برچسب گذاری دستی یا استخراج ویژگی نیاز ندارد در واقع این شبکه‌ها در دسته روش‌های بدون ناظر قرار می‌گیرند. در نتیجه امکان اعمال شبکه‌های تعبیه کلمات در هر مجموعه داده متنی براحتی امکان پذیر است [۳]. در شکل ۲ عملکرد تعبیه کلمات قابل مشاهده است. در این تکنیک بردار ورودی یک بردار one-hot است و خروجی، بردار یک کلمه است که توسط شبکه تعبیه کلمات تولید شده است.



شکل ۱: نمونه ایجاد تعبیه کلمات برای جمله "problem in user interface refresh"

۲- تعبیه کلمات و رمزگذاری

۲-۱- رمزگذاری داغ: روش‌های یادگیری تعبیه کلمات از یک متن، با تبدیل متن به رمزگذاری داغ شروع می‌شوند. منظور از رمزگذاری داغ فرآیندی است که در آن یک بردار خلوت با طولی برابر با کلمات منحصر بفرد در متن تشکیل می‌شود که هر کلمه منحصر به فرد را با یک شاخص منحصر به فرد نشان می‌دهد. به عبارت

خصوص داروهای مختلف به دو شبکه عمیق بازگشتی BiLSTM و شبکه کانولوشن ارائه شده است و ویژگی‌های استخراج شده از دو شبکه با یکدیگر پیوست شده و در نهایت با استفاده از یک لایه تمام متصل و یک لایه طبقه‌بندی نظرات بیماران در خصوص داروها دسته‌بندی شده است. این روش در بهترین حالت توانسته است به معیار $F 82.37\%$ دست یابد. یکی از مراحل که در موفقیت سیستم‌های تحلیل احساس بسیار موثر است مرحله پیش پردازش و تبدیل متن به ماتریسی از اعداد است.

۱-۲- اشاره به مراجع

هدف ارائه روشی است که در مقایسه با روش کلن رویز و همکاران [۲] منجر به بهبود دقت و صحت تشخیص تحلیل احساس در متن و نظرات بیماران در خصوص داروها گردد. در این راستا برای رسیدن به هدف فوق از ترکیب شبکه کانولوشن و بازگشتی در یک معماری واحد استفاده می‌شود به نحوی که لایه‌های کانولوشن در ابتدای معماری و لایه‌های بازگشتی در انتهای معماری شبکه تعریف شده درج شوند. همچنین برای مرحله آماده‌سازی متن از شبکه تعبیه کلمات پیش آموزش دیده FastText استفاده می‌شود.

متغیرهای تحقیق:

مستقل:

- تعداد واحدهای پنهان: تعداد واحدهای پنهان در لایه پنهان شبکه عصبی عمیق بازگشتی از جمله متغیرهای مستقل تحقیق است که می‌تواند باعث کاهش یا افزایش صحت مدل شود.

- تعداد فیلتر: تعداد فیلتر در لایه کانولوشن در شبکه عمیق ترکیبی پیشنهادی از دیگر پارامترهای ورودی مدل است که می‌تواند بر عملکرد مدل تاثیر بگذارد. تعداد فیلتر باید به نحوی تعریف شود که مرحله آموزش شبکه طولانی نشود و از طرف دیگر ویژگی‌های مناسبی از داده‌های ورودی استخراج نماید از این‌رو در این متغیر تعداد بین ۳ تا ۵۰ فیلتر در لایه کانولوشن بررسی می‌شود.

- سائز ورودی: در مدل ترکیبی پیشنهادی ورودی‌ها به صورت ماتریس‌های دو بعدی هستند (مثل تصاویر سطح خاکستری) سائز ورودی می‌تواند در نتایج مدل تاثیر داشته باشد. در نتیجه این پارامتر مدل نیز به عنوان متغیر مستقل در نظر گرفته شد تا تاثیر مقادیر مختلف آن بر نتایج صحت مدل بررسی شود.

وابسته:

- دقت Precision: این معیار بیانگر دقت مدل در پیشگویی داده‌های صحیح می‌باشد.

- فراخوانی Recall: این معیار نسبت تعداد داده‌های صحیح پیشگویی شده، به تعداد کل داده‌های قابل پیشگویی است.

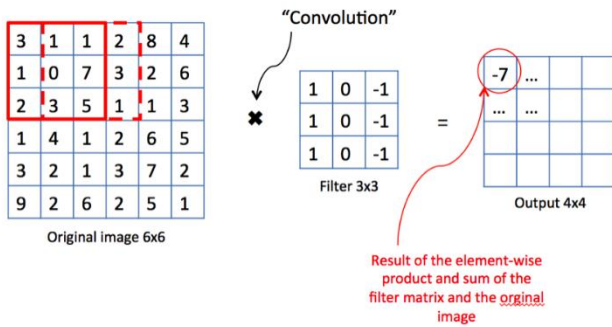
- صحت Accuracy: این معیار بیانگر صحت مدل در کلیه داده‌های پیشگویی شده است.

- میانگین هارمونیک F-Measure: این معیار از میانگین دو معیار دقت و فراخوانی حاصل می‌گردد. در جدول (۱-۱) مشخصات متغیرها و نحوه محاسبه آن‌ها بیان شده است.

۱-۲-۱- روش کار و روش شناسی تحقیق

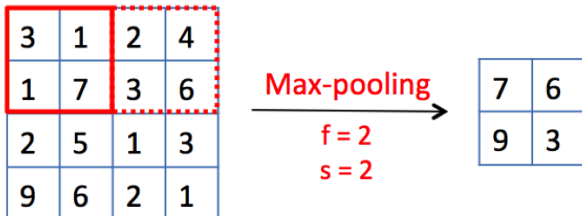
ابتدا مطالعه و بررسی مقالات خارجی و داخلی مرتبط با روش‌های تحلیل احساس به منظور شناخت روش‌های موجود و تاریخچه این موضوع انجام پذیرفت. اطلاعات از طریق اینترنت و مقالاتی که در مجلات معتبر چاپ شده‌اند، جمع‌آوری گردید. در ادامه با توجه به تکنیک‌های استفاده شده در برخی از آخرین مقالات، به صورت جزیی مطالعاتی در خصوص الگوریتم‌های یادگیری ماشین و زیر مجموعه

^۱ One-hot encoding



شکل ۴: عملیات در لایه کانولوشن

• لایه کاهش ابعاد: این لایه در معماری کانولوشنی پس از لایه کانولوشن درج می‌شود که دارای دو رویکرد اصلی تجمع حداکثر و تجمع میانگین است. این لایه، اندازه ماتریس ویژگی‌های خروجی لایه کانولوشن را کاهش می‌دهد. عملیات کاهش ابعاد ماتریس ویژگی‌ها، علاوه بر اینکه پیچیدگی‌های محاسباتی و تعداد پارمترهای شبکه را کاهش می‌دهد مشکل بیش برآزش را نیز حل می‌کند. در رویکرد تجمع حداکثر و میانگین، به ترتیب بیشترین مقدار و میانگین بلوک از تصویر، به عنوان خروجی این لایه انتخاب می‌شود [۹]. در شکل ۵ کاهش ابعاد ماتریس ویژگی با رویکرد تجمع حداکثر نشان داده شده است.



شکل ۵: نحوه کاهش ابعاد به روش تجمع حداکثر

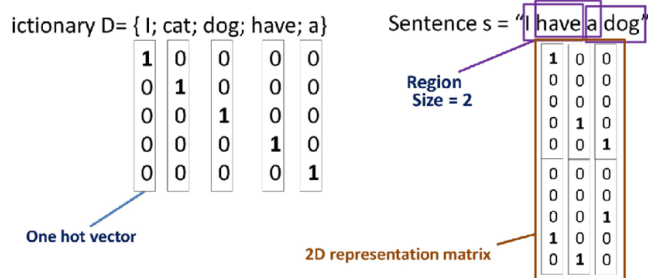
• لایه‌های تمام متصل: در معماری شبکه‌های کانولوشن آخرین لایه‌ها، لایه‌های تمام متصل هستند که پس از پشته‌ای از لایه‌های کانولوشن و کاهش بعد در شبکه درج می‌شوند. این لایه ویژگی‌های استخراج شده را تفسیر کرده و کلاس آن‌ها را پیش‌بینی می‌کند. این لایه کاملاً مشابه با شبکه عصبی پیشخور است و عملیات درون نرونی در این لایه انجام می‌شود. ورودی‌ها با یک بردار وزن ضرب شده و با یک بردار بایاس جمع می‌شوند. در این لایه امتیاز هر داده در کلیه کلاس‌ها محاسبه می‌شود و داده مورد نظر به کلاسی که بیشترین امتیاز را در آن دارد تخصیص می‌یابد [۱۱].

۳-۲ شبکه عصبی عمیق بازگشتی

شبکه‌های عصبی عمیق بازگشتی برای طبقه‌بندی داده‌های توالی معرفی شدند. شبکه‌های بازگشتی بر خلاف شبکه‌های پیشخور و کانولوشن دارای حلقه در معماری خود هستند که داده‌ها را برای آموزش مجدداً به شبکه ارائه می‌دهد. این شبکه‌ها از مجموعه‌ای از سلول‌ها تشکیل شده‌اند که عملیات یکسانی را بر روی یک توالی انجام می‌دهند. در شبکه بازگشتی هر سلول دارای ورودی به نام x ، خروجی به نام y و سه ماتریس وزن U برای ورودی، W برای واحد پنهان و V برای خروجی است. سه ماتریس وزن مذکور در ورودی‌های مختلف و در تمام مراحل زمانی دارای مقادیر یکسانی هستند [۱۲].

شبکه بازگشتی حافظه بلند مدت کوتاه: این دسته از شبکه‌های بازگشتی برای حل مشکلات گرادیان معرفی شدند. شبکه حافظه بلند مدت کوتاه، داده‌هایی با وابستگی طولانی را به خوبی تحلیل می‌کند. سلول حافظه بلند مدت

ساده همانطور که در شکل ۲ نشان داده شده است، یک کلمه به صورت برداری از صفرها نشان داده می‌شود و در موقعیت کلمات، مقدار یک قرار می‌گیرد

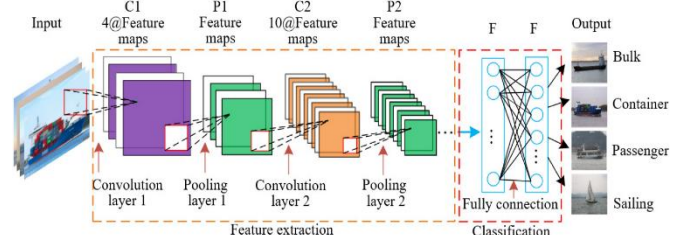


شکل ۲- نمونه ای از نمایش متن با بردار رمزگذاری داغ

۳-۱ معرفی شبکه‌های عصبی مورد تحقیق

۳-۱-۳ شبکه عصبی کانولوشنی:

شبکه‌های عصبی کانولوشنی در دسته شبکه‌های پیشخور عمیق قرار می‌گیرند که از واحدهای پردازش غیر خطی تشکیل شده‌اند و به صورت ویژه برای داده‌های دو بعدی و سه بعدی مثل تصاویر و ویدئو استفاده می‌شوند در طول فرآیند یادگیری شبکه‌های کانولوشنی، در لایه‌های ابتدایی شبکه استخراج ویژگی‌هایی مثل لبه‌ها و خطوط انجام می‌شود و در لایه‌های انتهایی شبکه طبقه‌بندی ویژگی‌های استخراج شده از تصاویر انجام می‌شود [۴]. شبکه‌های کانولوشنی برای اولین بار در سال ۱۹۹۴ ارائه و مطرح شدند [۵]. تفاوت اصلی شبکه‌های کانولوشنی با دیگر شبکه‌های عمیق در ورودی آن‌ها است [۶]. شبکه‌های کانولوشنی بدلیل ساختار و معماری که دارند در مقایسه با مدل‌های یادگیری ماشین سنتی، پیش‌پردازش کمتری بر روی ورودی خود انجام می‌دهند. در واقع در این شبکه‌ها برخلاف روش‌های سنتی مرحله استخراج ویژگی یک مرحله جداگانه نیست و در لایه‌های شبکه انجام می‌شود [۷]. در سال‌های اخیر مزایای متعدد این شبکه‌ها باعث شده از آن‌ها بطور گسترده در وظایف مختلف مثل تشخیص چهره، عنبیه، صدا، دستخط و طبقه‌بندی تصاویر استفاده شود [۴، ۸] در ادامه این سه لایه بطور دقیق‌تر معرفی شده‌اند.



شکل ۳: معماری یک شبکه کانولوشن

• لایه کانولوشن: این لایه از فیلترهایی با ابعاد قابل تنظیم تشکیل شده است که از آن‌ها برای استخراج ویژگی از تصاویر ورودی استفاده می‌شود. در این لایه با توجه به ابعاد فیلترها و گام حرکت فیلترها در تصویر، ضرب داخلی بین وزن‌ها (فیلترها) و تصویر انجام می‌شود و مقدار حاصل شده با ماتریس بایاس جمع می‌شود. فیلترها با توجه به گام تعیین شده برای آن‌ها که بطور پیش فرض برابر با 2×2 است در طول و عرض تصویر حرکت می‌کنند. شکل ۴ عملکرد لایه کانولوشن را با یک فیلتر 3×3 نشان می‌دهد در این شکل ماتریس تصویر ورودی دارای ابعاد 6×6 است. خروجی عملیات لایه کانولوشن از تابع فعال‌ساز مثل ReLU عبور می‌کند و ماتریس ویژگی‌های تصویر ورودی، تولید می‌گردد [۹، ۱۰].

کوتاه دارای سه دروازه ، ورودی ، فراموشی و خروجی است که مقدار حالت سلول را کنترل می کنند. سه دروازه فوق و وضعیت سلول حافظه به ترتیب از روابط (۲-۴) تا (۲-۷) حاصل می شوند.

$$i_t = \sigma(W_i h_{t-1} + W_i x_t + b_i) \quad 4-2$$

$$f_t = \sigma(W_f h_{t-1} + W_f x_t + b_f) \quad 5-2$$

$$o_t = \sigma(W_o h_{t-1} + W_o x_t + b_o) \quad 6-2$$

$$h_t = o_t \tanh(c_t) \quad \text{where} \quad c_t = f_t c_{t-1} + i_t \tanh(W_c x_t + W_c h_{t-1}) \quad 7-2$$

به ترتیب دروازه ورودی، فراموشی، خروجی و c_t و f_t, o_t, i_t در روابط فوق x_t نماد بردار بایاس و b نماد تابع فعال سازی سیگموئید، σ حالت سلول هستند را نشان می دهد و t حالت پنهان سلول در زمان h_t است t نماد ورودی در زمان [59] نیز نماد وزن است W .

این شبکه الهام گرفته : (GRU) شبکه بازگشتی واحد بازگشتی گیتی است اما در مقایسه با آن دارای پارامترهای کمتر، ساختار LSTM از ساختار شبکه ساده تر و اجزای سلول کمتری است که اولین بار در سال ۲۰۱۴ معرفی شد. این شبکه در مقابل مشکل محو شدگی گرادیان مقاوم است و دارای سرعت آموزش و (Z) بیشتری است. شبکه واحد بازگشتی گیتی دارای دروازه های بروزرسانی است که از روابط (۲-۸) تا (۲-۱۰) محاسبه می شوند (r) تنظیم مجدد

$$z_t = \sigma(W_z h_{t-1} + W_z x_t + z) \quad 8-2$$

$$r_t = \sigma(W_r h_{t-1} + W_r x_t + b_r) \quad 9-2$$

$$c_t = \tanh(W_c h_{t-1} + W_c x_t + b_c) \quad \text{where} \quad h_t = (z \otimes c) \oplus ((1-z) \otimes h_{t-1}) \quad 10-2$$

نماد دروازه بروزرسانی، تنظیم مجدد، h_t و c_t, r_t, z_t در روابط فوق به ترتیب نماد بایاس و b نماد تابع سیگموئید، σ حالت سلول و حالت پنهان سلول هستند [۱۴، ۱۳] است t نماد ورودی در زمان x_t

۲-۲- خلاصه تحقیقات پیشین

در جدول زیر خلاصه تحقیقات مشابه آورده شده است:

کریلو و همکاران [۱۵] ۲۰۱۸	ماشین بردار پشتیبان، بیز ساده، جنگل تصادفی	استفاده از تکنیک	رای گیری	عدم استفاده از شبکه های عصبی	صحت
۶۸,۷٪	نظرات در خصوص سلامت حصول نتایج بالاتر	اکثریت برای حصول نتایج بالاتر	رای گیری	رای گیری	صحت

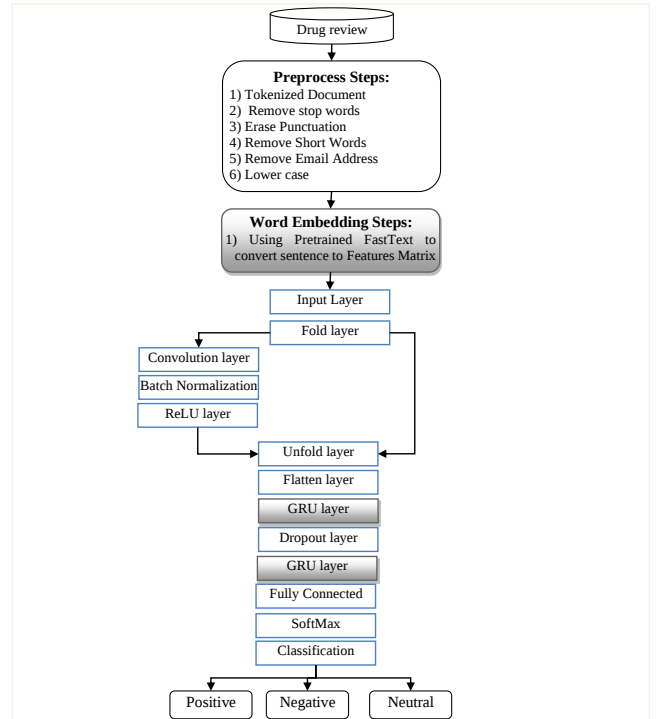
۲	یداو و همکاران [۱۶] ۲۰۱۸	شبکه کانولوشن	نظرات در خصوص دارو و سلامت	عمیق و مقایسه با یادگیری ماشین سنتی	عدم حصول نتایج بالا به دلیل پیش پردازش ناکافی	میانگین F ۸۲٪
۳	مین و همکاران [۱۷] ۲۰۱۹	شبکه کانولوشن و بازگشتی	نظرات دارو	یادگیری عمیق به صورت ترکیبی	عدم بررسی روش بر روی پایگاه داده های عمومی	صحت ۸۶,۷۲٪
۴	پنگ و همکاران [۱۸] ۲۰۱۹	استفاده از تکنیک های پالایش یادگیری مبتنی بر یادگیری عمیق	متون چینی	ویژگی های آوایی و ترکیب آن با ویژگی های متن	محدود در آنالیز استفاده شده و تنها بسنده به متون چینی شده است	صحت ۹۰ درصد بالاتر از ۹۰ درصد
۱/۸	سانگ و همکاران [۱۹] ۲۰۲۰	ماشین بردار پشتیبان	داده های متن	مناسب و عملکرد برای مجموعه های بندی منجر به کم شدن متوسط از داده	عدم تنظیم پارامترهای مدل طبقه - آنها بالاتر از ۸۵ درصد	صحت کار در تحقیق
۵	لیو و همکاران [۲۰] ۲۰۲۰	استفاده از الگوی CBOW و یادگیری عمیق	COAE2014 IMDB و	ویژگی ها از حجم فرایند آموزش کاسته است	امکان آنالیز تعداد کلاس های مختلف احساسات برآمده از متن به یک شبکه سریع تر و کارآمدتر نیاز دارد	در دو مجموعه داده به ترتیب به ۹۰٪ و ۸۷٪ درصد بدست آمده است
۶	کلن رویز و همکاران [۷] ۲۰۲۰	استفاده از شبکه کانولوشن و BiLSTM	Drug Review	استفاده از یادگیری عمیق	ترکیب نامناسب شبکه ها و مرحله پیش پردازش ضعیف	میانگین F ۸۲,۳۷٪ رسیده است

همانطور که در نتایج بالا مشخص است بالاترین درصد موفقیت مربوط به روش کلن رویز و همکارانش در سال ۲۰۲۰ بوده است که به میانگین F ۸۲,۳۷٪ رسیده و فقط در ترکیب نتایج موفق نبوده است.

۴-روش انجام تحقیق

پس از بررسی و معرفی مفاهیم بنیادی تحقیق و بررسی مدل های تحلیل احساس در حوزه مسئله این مقاله، در این روش تحلیل احساس پیشنهادی در حوزه داروسازی بطور کامل شرح داده می شود.

مدل ارائه شده در این تحقیق در دسته روش های بانظر و طبقه بندی است که یک معماری ترکیبی متشکل از دو شبکه عصبی عمیق کانولوشن و بازگشتی برای تحلیل احساس در صنعت دارو ارائه نموده است. در شکل ۶ دیاگرام مدل پیشنهادی ارائه شده است. این مدل شامل سه گام پیش پردازش و آماده سازی متن و گام آموزش و یادگیری مدل و در نهایت ارزیابی مدل است.



شکل ۶: دیاگرام مدل عمیق پیشنهادی برای تحلیل احساس در صنعت دارو

۱-۴ گام پیش پردازش و آماده سازی متن

اولین مرحله در مدل عمیق پیشنهادی آماده سازی متن است که این گام خود شامل ۶ بخش می شود. در این بخش برای واضحتر شدن ۶ عملیات انجام شده بر روی متن خروجی کد متلب روش به صورت تصویر ارائه می شود. در اولین گام نظرات بیماران نشانه گذاری می شوند. در گام نشانه گذاری هر کامنت تبدیل به لیستی از کلمات تشکیل دهنده خود می شود. به عنوان مثال اولین نظر در دو بخش آموزش و آزمایش در پایگاه داده به صورت زیر است که بعد از نشانه گذاری به مجموعه کلمات و علائم تشکیل دهنده آن تفکیک می شود. اولین کامنت در بخش آزمایش همانطور که در شکل ۸ نشان داده شده است به ۹۳ توکن و اولین کامنت در داده های آموزش همانطور که در شکل ۷ نشان داده شده است به ۲۰ توکن تبدیل می شود.

```
documents: 161297x1 tokenizedDocument =
161297x1 tokenizedDocument:
20 tokens: " I & # 039 ; ve tried a few antidepressants over the years ( citalopram , fluoxetine
164 tokens: " My son is halfway through his fourth week of Intuniv . We became concerned when h
172 tokens: " I used to take another oral contraceptive , which had 21 pill cycle , and was ver
104 tokens: " This is my first time using any form of birth control . I & # 039 ; m glad I went
159 tokens: " Suboxone has completely turned my life around . I feel healthier , I & # 039 ; m
95 tokens: " 2nd day on 5mg started to work with rock hard erections however experienced heada
41 tokens: " He pulled out , but he cummed a bit in me . I took the Plan B 26 hours later , an
171 tokens: " Abilify changed my life . There is hope . I was on Zoloft and Clonidine when I fi
47 tokens: " I Ve had nothing but problems with the Keppera : constant shaking in my arms & am
176 tokens: " I had been on the pill for many years . When my doctor changed my RX to chateal ,
172 tokens: " I have been on this medication almost two weeks , started out on 25mg and working
94 tokens: " I have taken anti-depressants for years , with some improvement but mostly modera
```

شکل ۷: مرحله نشانه گذاری نظرات بیماران در بخش آموزش

```
documents: 53766x1 tokenizedDocument =
53766x1 tokenizedDocument:
93 tokens: " I & # 039 ; ve tried a few antidepressants over the years ( citalopram , fluoxetine
65 tokens: " My son has Crohn & # 039 ; s disease and has done very well on the Asacol . He has n
6 tokens: " Quick reduction of symptoms "
197 tokens: " Contrave combines drugs that were used for alcohol , smoking , and opioid cessation
170 tokens: " I have been on this birch control for one cycle . After reading some of the reviews
74 tokens: " 4 days in on first 2 weeks . Using on arms and face . Put vaseline on lips , under e
159 tokens: " I & # 039 ; ve had the copper coil for about 3 months now . I was really excited at
87 tokens: " This has been great for me . I & # 039 ; ve been on it for 2 weeks and in the last w
101 tokens: " Ive been on Methadone for over ten years and currently , I am trying to get off of th
108 tokens: " I was on this pill for almost two years . It does work as far as not getting pregnan
```

شکل ۸: مرحله نشانه گذاری نظرات بیماران در بخش آزمایش

پس از نشانه گذاری متون در دومین مرحله برای هر نظر کلیه علائم نگارشی شامل نقطه، کاما، ویرگول آدرس ایمیل، وب سایت و اعداد از نظرات نشانه گذاری شده حذف می شود. حذف علامت های بی ارزش فوق باعث می شود تعداد توکن های هر نظر کاهش یابد. به عنوان مثال نظر سوم در شکل ۷ دارای ۱۷۲ توکن است اما پس از حذف علائم در شکل ۸ تعداد توکن های آن به ۱۶۸ کاهش یافته است. در داده های آزمایشی نظر اول از ۹۳ به ۹۰ توکن کاهش یافته است. پس از حذف اعداد، نظرات کاربران در بخش آموزش و صورت شکل ۸ و ۹ می شوند.

```
documents: 161297x1 tokenizedDocument =
161297x1 tokenizedDocument:
19 tokens: " It has no side effect , I take it in combination of Bystolic
163 tokens: " My son is halfway through his fourth week of Intuniv . We b
168 tokens: " I used to take another oral contraceptive , which had pill c
100 tokens: " This is my first time using any form of birth control . I &
155 tokens: " Suboxone has completely turned my life around . I feel heal
86 tokens: " day on started to work with rock hard erections however exp
39 tokens: " He pulled out , but he cummed a bit in me . I took the Plan
165 tokens: " Abilify changed my life . There is hope . I was on Zoloft a
47 tokens: " I Ve had nothing but problems with the Keppera : constant
171 tokens: " I had been on the pill for many years . When my doctor chang
166 tokens: " I have been on this medication almost two weeks , started c
```

شکل ۹: حذف علائم و اعداد از نظرات کاربران بخش داده های آموزش

```
documents: 53766x1 tokenizedDocument =
53766x1 tokenizedDocument:
90 tokens: " I & # ; ve tried a few antidepressants over the
63 tokens: " My son has Crohn & # ; s disease and has done v
6 tokens: " Quick reduction of symptoms "
191 tokens: " Contrave combines drugs that were used for alco
165 tokens: " I have been on this birth control for one cycle
71 tokens: " days in on first weeks . Using on arms and face
146 tokens: " I & # ; ve had the copper coil for about months
80 tokens: " This has been great for me . I & # ; ve been on
95 tokens: " Ive been on Methadone for over ten years and cu
105 tokens: " I was on this pill for almost two years . It do
167 tokens: " Holy Hell is exactly how I feel . I had been tai
36 tokens: " Honestly its day one on the day treatment . Yes
22 tokens: " This is a waste of money . Did not curb my appe
10 tokens: " No problems , watch what you eat . "
```

شکل ۱۰: حذف علائم و اعداد از نظرات کاربران بخش داده های آزمایش

در ادامه گام های آماده سازی متن، یکی از مراحل تبدیل متن به حروف کوچک است در این گام کلیه کلمات به حالت lowercase تبدیل می شود. خروجی این بخش در شکل های ۱۰ و ۱۱ به ترتیب برای نظرات بخش آموزش و آزمایش شده است.

```
documents: 161297x1 tokenizedDocument =
161297x1 tokenizedDocument:
19 tokens: " it has no side effect , i take it in combination of bystolic
163 tokens: " my son is halfway through his fourth week of intuniv . we be
168 tokens: " i used to take another oral contraceptive , which had pill c
100 tokens: " this is my first time using any form of birth control . i &
155 tokens: " suboxone has completely turned my life around . i feel health
86 tokens: " day on started to work with rock hard erections however expe
39 tokens: " he pulled out , but he cummed a bit in me . i took the plan
165 tokens: " abilify changed my life . there is hope . i was on zoloft ar
47 tokens: " i ve had nothing but problems with the keppera : constant sh
171 tokens: " i had been on the pill for many years . when my doctor chang
166 tokens: " i have been on this medication almost two weeks , started ou
93 tokens: " i have taken anti-depressants for years , with some improvem
83 tokens: " i had crohn & # ; s with a resection years ago and have beer
142 tokens: " have a little bit of a lingering cough from a cold . not giv
185 tokens: " started nexplanon months ago because i have a minimal amount
158 tokens: " i have been taking saxenda since july . i had severe nausea
```

شکل ۱۱: تبدیل کامنت کاربران به حروف کوچک در داده های آموزش

جملات کاربران کاهش یابد. در شکل‌های ۱۵ و ۱۶ خروجی این گام از مدل نشان داده شده است.

```
documents: 161297x1 tokenizedDocument =

161297x1 tokenizedDocument:

7 tokens: side effect take combination bystolic fish oil
62 tokens: son halfway fourth week intuniv became concerne
68 tokens: take another oral contraceptive pill cycle happ
35 tokens: first time form birth control glad went patch m
59 tokens: suboxone completely turned life around feel hea
39 tokens: day started work rock hard erections experiance
14 tokens: pulled cummed bit took plan hours later took pr
61 tokens: abilify changed life hope zoloft clonidine firs
22 tokens: nothing problems keppera constant shaking arms
64 tokens: pill many years doctor changed chateal effectiv
55 tokens: medication almost two weeks started working way
42 tokens: taken anti-depressants years improvement mostly
```

شکل ۱۵: حذف کلمات بی ارزش از نظرات کاربران در بخش آموزش

```
documents: 53766x1 tokenizedDocument =

53766x1 tokenizedDocument:

37 tokens: tried few antidepressants years citalopram f.
22 tokens: son crohn disease well asacol complaints sho
3 tokens: quick reduction symptoms
65 tokens: contrave combines drugs alcohol smoking opio.
60 tokens: birth control cycle reading reviews type sim.
29 tokens: days first weeks arms face put vaseline lips
47 tokens: copper coil months really excited thought ta.
24 tokens: great weeks last week headaches went away ty.
32 tokens: methadone ten years currently trying get off.
40 tokens: pill almost two years work far getting pregn.
65 tokens: holy hell exactly feel taking brisdelle year.
13 tokens: honestly day day treatment yes burns bit lea.
8 tokens: waste money curb appetite nor make feel full
```

شکل ۱۶: حذف کلمات بی ارزش از نظرات کاربران در بخش آزمایش

۲-۴ تبدیل متن به عدد با استفاده از شبکه تعبیه کلمات

پس از تکمیل فرآیند آماده سازی متن و انجام گام‌های مختلف بر روی متن‌ها برای حذف کلمات بی ارزش و استخراج کلمات ارزشمند؛ در دومین مرحله از مدل عمیق پیشنهادی کلمات استخراج شده از گام آماده‌سازی و پیش پردازش با استفاده از تکنیک تعبیه کلمات به برداری از اعداد تبدیل می‌شوند. در این تحقیق از شبکه پیش آموزش دیده FastText برای تبدیل متن به ماتریس اعداد استفاده شده است. این شبکه دارای بالغ بر ۱۶ میلیارد توکن است. این شبکه برای هر کلمه در هر جمله یک بردار با طول ۳۰۰ بُعد بطور پیش فرض تولید می‌کند.

در این گام از مدل هر جمله به شبکه تعبیه کلمات پیش آموزش دیده FastText داده می‌شود. سپس شبکه تعبیه کلمات، ماتریسی از اعداد برای هر جمله تولید می‌کند. ستون‌های هر ماتریس برابر با ۳۰۰ است و سطرهاى آن با توجه به کلمات هر جمله (توکن‌ها) متفاوت است. در واقع شبکه‌های تعبیه کلمات برای هر کلمه یک بردار از اعداد تولید می‌کنند. با توجه به اینکه شبکه‌های عمیق ورودی‌هایی با اندازه مشخص دریافت می‌کنند؛ خروجی شبکه FastText را باید برای هر جمله به نحوی تغییر داد که برای کلیه جملات دارای اندازه یکسان باشد. با توجه به این در مدل این تحقیق معماری ترکیبی پیشنهادی از لایه‌های کانولوشن در ابتدای معماری شبکه و از لایه‌های شبکه بازگشتی در انتهای معماری تشکیل شده است؛ و از آنجایی که شبکه کانولوشن داده‌های دو بعدی به عنوان ورودی دریافت می‌کند، برای این بخش در مدل پیشنهادی، با هر جمله مشابه با تصاویر برخورد شده است. به این معنی که ماتریس دو بعدی خروجی شبکه تعبیه کلمات هر جمله، به ابعاد ۵۰×۵۰ تغییر سایز داده شد. این عمل با استفاده از تابع تغییر سایز تصاویر در متلب انجام پذیرفت. در شکل (۳-۱۵) بخشی از ماتریس خروجی شبکه تعبیه کلمات برای نظر اول در داده‌های آموزشی ارائه شده است. در این شکل، هر سطر مقادیر معنایی

```
documents: 53766x1 tokenizedDocument =

53766x1 tokenizedDocument:

90 tokens: " i & # ; ve tried a few antidepressants over t
63 tokens: " my son has crohn & # ; s disease and has done
6 tokens: " quick reduction of symptoms "
191 tokens: " contrave combines drugs that were used for al
165 tokens: " i have been on this birth control for one cyc
71 tokens: " days in on first weeks . using on arms and fa
146 tokens: " i & # ; ve had the copper coil for about mont
80 tokens: " this has been great for me . i & # ; ve been
95 tokens: " ive been on methadone for over ten years and
105 tokens: " i was on this pill for almost two years . it
167 tokens: " holy hell is exactly how i feel . i had been
36 tokens: " honestly its day one on the day treatment . y
22 tokens: " this is a waste of money . did not curb my ap
10 tokens: " no problems , watch what you eat . "
```

شکل ۱۲: تبدیل کلمات کاربران به حروف کوچک در داده‌های آزمایش

یکی دیگر از گام‌های آماده‌سازی داده‌ها حذف کلمات با طول کمتر از ۲ است. این کلمات نمی‌توانند در فرآیند تحلیل احساس موثر باشند و فقط پیچیدگی‌های محاسباتی مدل را افزایش می‌دهند. از این‌رو حذف آن‌ها بسیار موثر است. خروجی این بخش در شکل‌های ۱۳ و ۱۴ برای دو بخش آموزش و آزمایش نشان داده شده است.

```
documents: 161297x1 tokenizedDocument =

161297x1 tokenizedDocument:

9 tokens: has side effect take combination bystolic and fish oil
106 tokens: son halfway through his fourth week intuniv became concer
108 tokens: used take another oral contraceptive which had pill cycle
63 tokens: this first time using any form birth control glad went wit
85 tokens: suboxone has completely turned life around feel healthier
52 tokens: day started work with rock hard erections however experia
18 tokens: pulled out but cummed bit took the plan hours later and to
97 tokens: abilify changed life there hope was zoloft and clonidine
26 tokens: had nothing but problems with the keppera constant shaking
92 tokens: had been the pill for many years when doctor changed chat
95 tokens: have been this medication almost two weeks started out an
64 tokens: have taken anti-depressants for years with some improveme
52 tokens: had crohn with resection years ago and have been mostly r
82 tokens: have little bit lingering cough from cold not giving much
102 tokens: started nexplanon months ago because have minimal amount
شکل ۱۳: حذف کلمات با طول کمتر از ۲، از نظرات کاربران در بخش در داده‌های آموزشی
```

```
documents: 53766x1 tokenizedDocument =

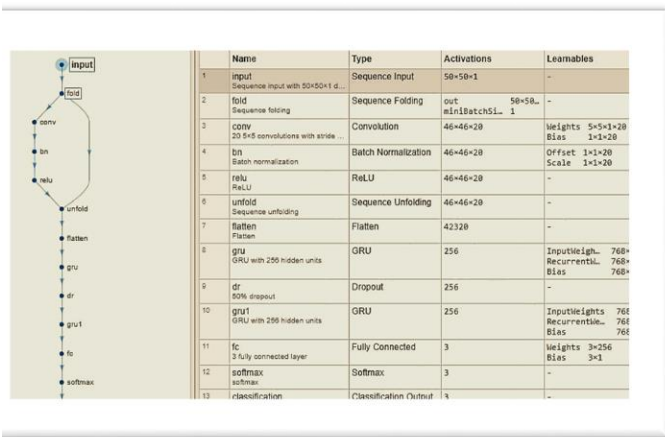
53766x1 tokenizedDocument:

53 tokens: tried few antidepressants over the years cita
37 tokens: son has crohn disease and has done very well
3 tokens: quick reduction symptoms
100 tokens: contrave combines drugs that were used for al
110 tokens: have been this birth control for one cycle af
41 tokens: days first weeks using arms and face put vase
82 tokens: had the copper coil for about months now was
45 tokens: this has been great for been for weeks and th
55 tokens: ive been methadone for over ten years and cur
66 tokens: was this pill for almost two years does work
104 tokens: holy hell exactly how feel had been taking br
24 tokens: honestly its day one the day treatment yes bu
12 tokens: this waste money did not curb appetite nor di
5 tokens: problems watch what you eat
```

شکل ۱۴: حذف توکن‌های با طول کمتر از ۲ در نظرات کاربران بخش داده‌های آزمایشی

همانطور که در دو شکل بالا مشاهده می‌شود تعداد توکن‌های جملات پس از حذف کلمات با طول کمتر از ۲ به شدت کاهش می‌یابد. به عنوان مثال نظر اول در داده‌های آزمایش دارای ۹۰ توکن بوده که پس از حذف کلمات با طول کمتر از ۲ به ۵۳ کاهش یافته است. در گام بعد از مرحله آماده سازی متون، کلمات بی ارزش که دارای بار معنایی نیستند مثل ضمایر ملکی از جملات کاربران حذف می‌شود. این گام نیز از جمله گام‌هایی در مرحله پیش پردازش است که باعث می‌شود طول

- لایه ورودی (input): این لایه وظیفه دریافت ورودی را برعهده دارد و در مدل پیشنهادی ورودی‌هایی با ابعاد $[1 \times 50 \times 50]$ را دریافت می‌کند.
- لایه تاشو (fold): لایه‌ای ایجاد می‌کند که ابعاد داده‌ها و ترتیب ورودی آن را با هم ترکیب می‌کند. بطور دقیق این لایه دسته‌ای از دنباله‌های تصویر را به دسته‌ای از تصاویر تبدیل می‌کند. از یک لایه تاشو برای انجام عملیات کانولوشن در توالی‌های تصویر که مستقل هستند استفاده می‌شود.
- لایه کانولوشن (conve): این لایه از داده‌های ورودی ویژگی استخراج می‌کند. در مدل پیشنهادی تعداد فیلترهای این لایه برابر با ۲۰ و ابعاد فیلترها برابر با 5×5 در نظر گرفته شد.
- لایه نرمال‌سازی دسته‌ای (bn): این نوع لایه هر کانال را در یک دسته کوچک از داده‌ها نرمال‌سازی می‌کند. این لایه در کاهش حساسیت به تغییرات درون داده‌ها مفید است. برای سرعت بخشیدن به آموزش شبکه عصبی کانولوشن و کاهش حساسیت به مقاردهای اولیه شبکه، در معماری پیشنهادی از لایه نرمال‌سازی دسته‌ای بین لایه کانولوشن و لایه تابع فعال‌ساز غیرخطی ReLU استفاده شد.



شکل ۱۹: معماری پیشنهادی و پارامترهای و وزن‌های هر لایه از آن

- لایه تابع فعال‌ساز غیر خطی ReLU: یک لایه واحد خطی اصلاح شده ایجاد می‌کند. این لایه همانطور که در فصل قبل شرح داده شد فرآیند نرمال‌سازی خروجی را انجام می‌دهد. این نوع لایه یک عملیات آستانه ساده را انجام می‌دهد که در آن هر مقدار ورودی کمتر از صفر روی صفر تنظیم می‌شود.
- لایه گشودن (unfold): این لایه ابعاد داده‌ها و ترتیب ورودی آن را به حالت اولیه بازبایی می‌کند.
- لایه مسطح‌سازی (flatten): یک لایه مسطح، ابعاد فضایی ورودی را در بُعد کانال جمع می‌کند. این لایه در معماری پیشنهادی برای تبدیل ماتریس اعداد (جمله‌هایی که تبدیل به ماتریس دو بعدی شدند) به بردارهای قابل پردازش برای شبکه عصبی بازگشتی است.
- لایه واحد بازگشتی دروازه‌ای (GRU): یک لایه بازگشتی با واحد پنهان GRU در شبکه درج می‌کند. این لایه به عنوان ورودی تعداد واحد پنهان را دریافت می‌کند که در مدل پیشنهادی برابر با ۲۵۶ در نظر گرفته شده است.
- لایه حذف تصادفی (dr): در طول آموزش، لایه حذف، به طور تصادفی عناصر ورودی را با احتمال ۰.۵ صفر می‌کند. این لایه در معماری پیشنهادی برای جلوگیری از مشکل بیش برآزش درج شده است.
- لایه GRU دوم است. در معماری پیشنهادی یک شبکه بسیار عمیق تعریف شده است که در آن از دو لایه GRU استفاده شده است.
- لایه تمام متصل (fc): یک لایه کاملاً متصل ایجاد می‌کند. این لایه به عنوان ورودی، تعداد خروجی مسئله (تعداد کلاس) را دریافت می‌کند. در معماری پیشنهادی، تحلیل احساس در سه کلاس مثبت، منفی و خنثی انجام می‌شود در

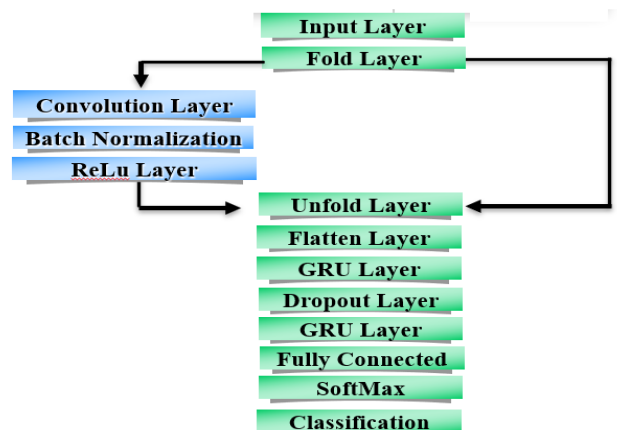
مربوط به یک کلمه از جمله اول را نشان می‌دهد. با توجه به محدودیت فضا از ۳۰۰ بعد تولیدی شبکه FastText فقط ۹ بعد اول برای هر کلمه نشان داده شده است. همانطور که در این شکل نشان داده شده ماتریس خروجی برای نظر اول دارای ابعاد ۷ در ۳۰۰ است چون نظر اول دارای ۷ توکن یا کلمه است. شبکه تعبیه کلمات FastText برای کلمه ۵ در این جمله که کلمه Bystolic است نتوانسته ویژگی تولید کند در مدل پیشنهادی این بردارها از ماتریس جملات بطور کامل مقدار صفر می‌گیرند و سپس ماتریس جملات به ابعاد ۵۰ در ۵۰ تغییر سایز داده می‌شوند.

documentstrain(1, 1).Vocabulary									
1	2	3	4	5	6	7	8	9	
1 side	effect	take	combination	bystolic	fish	oil			
7x300 single									
1	2	3	4	5	6	7	8	9	
0.0012	-0.0224	-0.0703	0.0499	-0.0315	0.0281	0.0200	-0.1395	-0.1581	
0.1587	-0.0427	0.0563	-0.0378	0.0762	-0.0069	0.0174	-0.0527	0.0079	
-0.0019	0.0302	-0.0063	0.0662	0.0179	-0.0362	-0.0382	-0.0638	0.0354	
-0.0294	-0.0030	-0.0120	0.0765	-0.0603	0.0175	-0.0228	0.0252	8.0000e-04	
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
-0.1016	0.0367	-0.1085	0.0941	-0.0793	-0.0824	0.1755	-0.0363	-0.0188	
-0.0204	0.0720	-0.0203	-0.0230	0.0224	-0.0932	0.1693	-0.1876	-0.1572	
7x300 single									
1	2	3	4	5	6	7	8	9	
0.0012	-0.0224	-0.0703	0.0499	-0.0315	0.0281	0.0200	-0.1395	-0.1581	
0.1587	-0.0427	0.0563	-0.0378	0.0762	-0.0069	0.0174	-0.0527	0.0079	
-0.0019	0.0302	-0.0063	0.0662	0.0179	-0.0362	-0.0382	-0.0638	0.0354	
-0.0294	-0.0030	-0.0120	0.0765	-0.0603	0.0175	-0.0228	0.0252	8.0000e-04	
0	0	0	0	0	0	0	0	0	
-0.1016	0.0367	-0.1085	0.0941	-0.0793	-0.0824	0.1755	-0.0363	-0.0188	
-0.0204	0.0720	-0.0203	-0.0230	0.0224	-0.0932	0.1693	-0.1876	-0.1572	
50x50 single									
1	2	3	4	5	6	7	8	9	
-0.0184	-0.0602	-0.0119	-6.9646e-04	-0.0237	0.0048	0.0371	-0.0421	-0.0788	
-0.0174	-0.0597	-0.0127	-6.5500e-04	-0.0232	0.0042	0.0369	-0.0412	-0.0792	
-0.0155	-0.0587	-0.0143	-5.7209e-04	-0.0221	0.0029	0.0364	-0.0394	-0.0802	
-0.0126	-0.0572	-0.0166	-4.4772e-04	-0.0205	0.0010	0.0356	-0.0366	-0.0815	
-0.0078	-0.0551	-0.0207	-3.0613e-04	-0.0177	-0.0022	0.0345	-0.0319	-0.0837	

شکل ۱۷: ماتریس خروجی شبکه تعبیه کلمات FastText برای نظر اول در بخش داده‌های آموزشی

۳-۴ تحلیل احساس با استفاده از معماری عمیق پیشنهادی

پس از تشکیل ماتریس جملات با ابعاد ۵۰ در ۵۰ در آخرین گام از مدل معماری ترکیبی عمیق پیشنهادی تعریف می‌شود و با استفاده از داده‌های بخش آموزش، فرآیند یادگیری آن تکمیل می‌شود. معماری ترکیبی عمیق پیشنهادی در شکل ۱۸ و پارامترهای هر لایه در شکل ۱۹ نشان داده شده است. این دو شکل توسط متلب تولید شده است. معماری پیشنهاد شده در دسته معماری‌های بسیار عمیق قرار می‌گیرد به این دلیل که در بخش بازگشتی از دو لایه GRU با یک لایه حذف تصادفی در بین آن‌ها تشکیل شده است. این معماری از ۱۳ لایه به شرح زیر تشکیل شده است:



شکل ۱۸: معماری عمیق ترکیبی پیشنهادی

جدول (۱-۳) مثالی از ارزیابی مدل

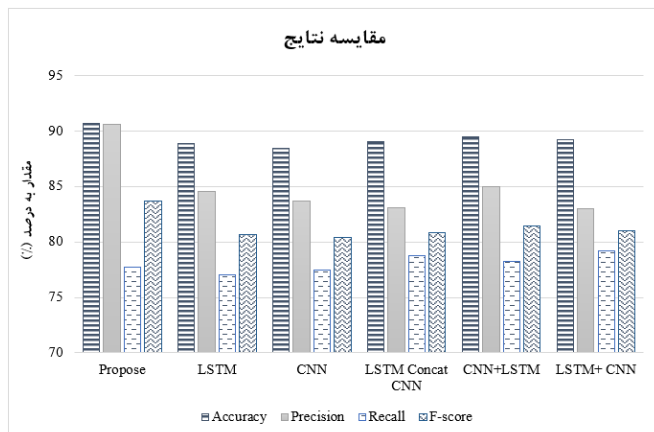
Id	Predicted Class	Real Class	Success or Failure
1	Positive	Positive	✓
2	Positive	Positive	✓
3	Positive	Positive	✓
4	Negative	Positive	×
5	Negative	Negative	✓
6	Negative	Negative	✓
7	Positive	Positive	✓
8	Neutral	Neutral	✓
9	Positive	Positive	✓
10	Neutral	Neutral	✓

جدول (۲-۳) ماتریس درهم ریختگی مثال فرضی

		Actual Class			
		Positive	Negative	Neutral	
Predict Class	Positive	True Positive 5	False Positive 0	False Positive 0	Precision= 5/5=1
	Negative	False Negative 1	True Negative 2	False Negative 0	Precision= 2/3=0.66
	Neutral	False Neutral 0	False Neutral 0	True Neutral 2	Precision= 2/2=1
		Recall=5/6=0.83	Recall=2/2=1	Recall=2/2=1	Accuracy=5+2+2/10=0.9

۴-۴ مقایسه نتایج روش‌ها

مقایسه‌ای بین مدل پیشنهادی و ۵ روش ارائه شده در مطالعه کلن رویز و همکاران [۷] انجام شد. نتایج این آزمایش در شکل (۴-۴) و جدول (۴-۴) گزارش شده است. در این آزمایش کلیه معیارهای ارزیابی که در روش پایه استفاده شده شامل صحت، دقت، فراخوانی و میانگین هارمونیک مقایسه و بررسی شده است.



شکل ۲۱: مقایسه نتایج روش‌ها

جدول (۴-۴) نتایج مقایسه روش‌ها

	Accuracy	Precision	Recall	F-score
Propose	90.7	90.59	77.74	83.676
LSTM	88.82	84.55	77.05	80.62
CNN	88.43	83.64	77.42	80.41
LSTM Concat CNN	89.02	83.08	78.73	80.85
CNN+LSTM	89.43	84.98	78.22	81.46
LSTM+CNN	89.19	83.01	79.16	81.03

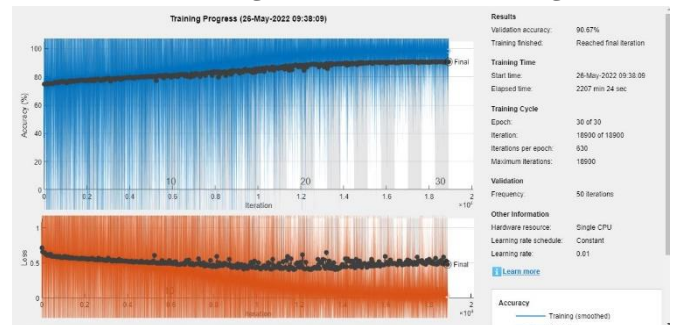
نتایج آزمایش مقایسه‌ای به خوبی نشان می‌دهد که مدل ما در مقایسه با ۵ مدل ارائه شده در روش کلن و همکاران بهتر عمل نموده است. یکی از عوامل موثر در موفقیت مدل ما در مقایسه با مدل‌های دیگر نحوه ترکیب دو شبکه عصبی عمیق در

نتیجه تعداد خروجی برای این لایه ۳ تعریف شده است که در شکل (۳-۱۷) مشاهده می‌شود. یک لایه کاملاً متصل، ورودی را در یک ماتریس ضرب می‌کند و سپس با بردار بایاس جمع می‌کند.

• لایه SoftMax: این لایه برای مسائل طبقه‌بندی چند کلاسی تعریف می‌شود؛ که برای هر نمونه داده احتمال عضویت در هر کلاس را محاسبه می‌کند و مجموع احتمال‌های محاسبه شده برای هر داده برابر با ۱ خواهد بود.

• لایه طبقه‌بندی: آخرین لایه در معماری شبکه پیشنهادی است که در آن کلاس داده‌ها مشخص می‌شود و هر داده به کلاسی تخصیص می‌یابد که در آن کلاس بیشترین احتمال عضویت را داشته باشد.

پس از تعریف معماری فوق و تشکیل لایه‌های شبکه پیشنهادی، فرآیند یادگیری شبکه فوق با استفاده از داده‌های بخش آموزش انجام می‌شود. در مدل پیشنهادی برای فرآیند یادگیری شبکه از تابع بهینه‌ساز نزول گرادیان تصادفی با حرکت (SGDM) استفاده شده است و تعداد گردش شبکه ۵۰ و سایز دسته‌های کوچک داده نیز برابر با ۲۵۶ تعریف گردید. در شکل ۲۰ نمودار مرحله یادگیری شبکه پیشنهادی بر روی داده‌های آموزشی نشان داده شده است. با توجه به اینکه مدل پیشنهادی در دسته مدل‌های با ناظر است در مرحله یادگیری برچسب احساس داده‌ها همراه با داده‌ها به شبکه ارائه می‌شود اما در فاز ارزیابی برچسب داده‌ها به شبکه ارائه نمی‌شود تا شبکه بتواند آن‌ها را پیش‌بینی کند.



شکل ۲۰: نمودار مرحله یادگیری شبکه پیشنهادی

شکل بالا از دو نمودار تشکیل شده است. نمودار بالایی صحت مدل و نمودار پایینی خطای مدل است. نقطه‌های تیره وضعیت پیش‌بینی شبکه در داده‌های ارزیابی (Validation) را نشان می‌دهند. اگر فرآیند یادگیری در شبکه به درستی انجام شود، خطای شبکه در طول فرآیند یادگیری کاهش می‌یابد. در هر گردش از فرآیند یادگیری شبکه، خروجی تولید شده توسط شبکه با کلاس حقیقی داده‌ها مقایسه می‌شود و مقدار خطای عملکرد شبکه از مقایسه دو مقدار مذکور حاصل می‌شود. مقدار خطا به شبکه بازگردانده می‌شود تا شبکه براساس آن وزن‌های هر لایه را به نحوی برورسانی کند که در طول یادگیری خطا کاهش یابد. همانطور که در شکل مشاهده می‌شود صحت مدل در داده‌های ارزیابی برابر با ۹۰٫۶۷٪ است. پس از پایان فرآیند یادگیری شبکه داده‌های آزمایش بدون برچسب احساس آن‌ها به شبکه ارائه می‌شود تا شبکه آموزش دیده و آماده، کلاس احساس داده‌های جدید را پیش‌بینی کند. خروجی این بخش کلاس‌های مثبت، منفی و خنثی برای داده‌های بخش آزمایش است.

- [11] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352-2449, 2017.
- [12] Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *International Journal of Forecasting*, vol. 37, no. 1, pp. ۳۸۸-۴۲۷, ۲۰۲۱.
- [13] K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint pp. 1-15*, 2014.
- [14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint pp. 1-9*, 2014.
- [15] J. Carrillo-de-Albornoz, J. Rodriguez Vidal, and L. Plaza, "Feature engineering for sentiment analysis in e-health forums," *PloS one*, vol. 13, no. 11, pp. 1-25, 2018.
- [16] S. Yadav, A. Ekbal, S. Saha, and P. Bhattacharyya, "Medical sentiment analysis using social media: towards building a patient assisted system," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, ۲۰۱۸, pp. ۲۷۹۰-۲۷۹۷.
- [17] Z. Min, "Drugs reviews sentiment analysis using weakly supervised model," in *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2019: IEEE, pp. ۳۳۲-۳۳۶.
- [18] H. Peng, Y. Ma, S. Poria, Y. Li, and E. Cambria, "Phonetic-enriched Text Representation for Chinese Sentiment Analysis with Reinforcement Learning," *arXiv preprint* 2019.
- [19] C. Song, X.-K. Wang, P.-f. Cheng, J.-q. Wang, and L. Li, "SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis," *Knowledge-Based Systems*, p. 105572, 2020.
- [20] B. Liu, "Text sentiment analysis based on CBOW model and deep learning in big data environment," *Journal of Ambient Intelligence Humanized Computing*, vol. 11, no. 2, pp. 451-458, 2020.

یک معماری مشترک است. در مدل پیشنهادی یک شبکه بسیار عمیق ارائه شده که با درج یک لایه حذف تصادفی بین دو لایه GRU بازگشتی باعث شده فرآیند یادگیری در مدل به نحو بهتری انجام شود. از طرف دیگر استفاده از شبکه تعبیه کلمات پیش آموزش دیده FastText از دیگر عوامل موثر در نتایج شبکه پیشنهادی است. شبکه تعبیه کلمات پیش آموزش دیده بالغ بر ۱۶ بلیون توکن دارد که می تواند به خوبی از کلمات مختلف ویژگی معنایی مناسب استخراج کند. همچنین مرحله پیش پردازش نیز در موفقیت مدل ما موثر بوده است. کلیه عوامل فوق باعث شده مدل بتواند به نتایج امیدوار کننده ای در تحلیل احساس در صنعت داروسازی دست یابد.

۵-جمع بندی و نتیجه گیری

در این تحقیق نتایج مدل معرفی گردید. برای ارزیابی مدل از پایگاه داده drug review استفاده شد که بالغ بر ۲۱۰۰۰۰ داده دارد. آزمایش ها در این فصل در دو دسته اصلی شامل چند گروه انجام شدند. در دسته اول آزمون های ارزیابی انجام شد. آزمون های ارزیابی سه گروه آزمایش بودند که در آن ها تعداد واحد پنهان لایه بازگشتی، تعداد فیلتر لایه کانولوشن و سائز داده های ورودی شبکه بررسی شدند. نتایج این آزمایش ها نشان داد که مدل در داده هایی با ابعاد ۵۰ در ۵۰ و با تعداد ۲۰ فیلتر در لایه کانولوشن و تعداد ۲۵۶ سلول پنهان در لایه بازگشتی GRU می تواند به بهترین نتایج دست یابد. پس از آن در دسته دوم از آزمایش ها مقایسه ای بین نتایج مدل پیشنهاد شده و ۵ مدل ارائه شده در روش پایه کلن و همکاران انجام شد. نتایج این آزمایش نشان داد که بیشترین موفقیت مدل در معیار دقت و پس از آن در صحت و میانگین f و کمترین موفقیت مدل در معیار فراخوانی است. در واقع مدل در معیار فراخوانی در مقایسه با سه روش پایه ضعیف تر عمل کرده است. اما در مقابل در معیار دقت به ۵ تا ۷٫۵٪ و در صحت بین ۱٫۲ تا ۲٫۲ در مقایسه با روش های پیشین بهبود داشته است.

۶-مراجع

- [1] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, pp. 1-26, 2021.
- [2] C. Colón-Ruiz and I. Segura-Bedmar, "Comparing deep learning architectures for sentiment analysis on drug reviews," *Journal of Biomedical Informatics*, vol. 110, pp. 1-11, 2020.
- [3] F. K. Khattak, S. Jebblee, C. Pou-Prom, M. Abdalla, C. Meaney, and F. Rudzicz, "A survey of word embeddings for clinical text," *Journal of Biomedical Informatics: X*, vol. 4, pp. 1-18, 2019.
- [4] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *Neural and Evolutionary Computing*, pp. 1-11, 2015.
- [5] Z. Shi, L. He, K. Suzuki, T. Nakamura, and H. Itoh, "Survey on neural networks used for medical image processing," *International journal of computational science*, vol. 3, no. 1, pp. 1-16, 2009.
- [6] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Information Fusion*, vol. 42, pp. 146-157, 2018.
- [7] Y. Li, Z. Hao, and H. Lei, "Survey of convolutional neural network," *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508-2515, 2016.
- [8] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11-26, 2017.
- [9] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, 2015.
- [10] M. Z. Alom et al., "The history began from alexnet: A comprehensive survey on deep learning approaches," *arXiv preprint arXiv:1803.01164*, pp. 1-11, 2018.