Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال چهارم، شماره اول، بهار و تابستان ۱۴۰۵، صفحات ۱۴۵- ۱۲۹

ارائه مدلی ترکیبی و کم‌پارامتر مبتنی بر معماری ترنسفورمر و GRU برای تحلیل احساسات متون فارسی

محسن نورائی[۱]، حمیدرضا غفاری[۲*]، فاطمه زریسفی کرمانی[۳]

[۱] دانشجوی دکتری، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد فردوس، خراسان جنوبی، ایران

[۲] دانشیار، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد فردوس، خراسان جنوبی، ایران

[۳] استادیار، دانشکده ریاضی و علوم کامپیوتر، دانشگاه جیرفت، کرمان، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۹/۲۵

تاریخ پذیرش:
۱۴۰۴/۱۲/۲۶

کلیدواژه‌ها:
پردازش زبان طبیعی، تحلیل احساسات، کاهش پیچیدگی محاسباتی، مدل ترنسفورمر

نویسنده‌مسئول:
hghaffari@ferdowsiau.ac.ir

چکیده

تحلیل احساسات به‌عنوان یکی از زیرمجموعه‌های طبقه‌بندی متن، ابزاری حیاتی در حوزه‌هایی مانند مدیریت اطلاعات، تحلیل داده و بهبود عملکرد است. پیچیدگی متون فارسی، به دلیل ویژگی‌هایی نظیر طعنه، چندمعنایی و ترکیب زبان‌ها، چالش‌های منحصربه‌فردی را برای شناسایی احساسات ایجاد می‌کند. در این پژوهش، مدلی جدید معرفی‌شده که با ترکیب GRU و ParsBERT، ویژگی‌های محلی و سراسری متن را استخراج می‌کند. در این پژوهش تمرکز اصلی بر کاهش پیچیدگی محاسباتی و زمان آموزش با حفظ سطح مطلوب دقت بوده است، تا مدل بتواند در محیط‌های با منابع محدود نیز قابلیت استفاده داشته باشد. این مدل با کاهش لایه‌های رمزگذار و افزودن GRU، پارامترها و زمان آموزش را بهبود می‌بخشد. آزمایش‌ها روی مجموعه داده‌های اسنپ‌فود و طاقچه نشان داد که این مدل نسبت به مدل‌های پایه دقت بالاتری دارد و آموزش آن نیز سریع‌تر است. مدل ارائه‌شده در مجموعه داده طاقچه به ۷۶/۷ و در اسنپ‌فود به ۸۶/۱۶ در معیار F1 دست یافت، درحالی‌که تعداد پارامترهای آن کاهش چشمگیری داشت. این نتایج نشان می‌دهد که مدل ارائه‌شده برای تحلیل احساسات متون فارسی، کارآمد و بهینه است.

: 10.22034/ABMIR.2026.24111.1200

E-ISSN: 2821-2037

/The Author 2026. Published by Yazd University This is an open access article under the CC BY 4.0 License (https://creativecommons.org/licenses/by/4.0/).

۱- مقدمه

طبقه‌بندی متون به‌عنوان یکی از کاربردهای اصلی پردازش زبان طبیعی (NLP)، نقش حیاتی در سازماندهی و تحلیل داده‌های متنی در حوزه‌های مختلف ایفا می‌کند. این فرآیند شامل دریافت متن به‌عنوان ورودی و طبقه‌بندی آن به یک یا چند دسته با استفاده از الگوریتم‌های یادگیری ماشین یا قواعد برنامه‌ریزی‌شده است. اهمیت طبقه‌بندی متون در حوزه‌های مختلفی مانند مدیریت اخبار، تحلیل اجتماعی، پشتیبانی از مشتریان، تحلیل داده‌ها، تشخیص هرزنامه‌ها و بهبود رتبه‌بندی موتورهای جستجو به‌وضوح قابل‌مشاهده است. برای مثال، در حوزه مدیریت اخبار، طبقه‌بندی متون به سازماندهی اخبار بر اساس موضوعات مختلف کمک می‌کند و دسترسی کاربران به اطلاعات موردنیاز را تسهیل می‌نماید. در حوزه تحلیل اجتماعی، این فناوری به درک احساسات و نظرات عمومی درباره موضوعات مختلف کمک می‌کند و می‌تواند در تصمیم‌گیری‌های استراتژیک مورداستفاده قرار گیرد.

یکی از کاربردهای مهم طبقه‌بندی متون، تحلیل احساسات است که هدف آن استخراج نظرات، احساسات و تمایلات کاربران از متون موجود در وب‌سایت‌ها و شبکه‌های اجتماعی است. این فناوری به‌طور گسترده‌ای در کسب‌وکارها برای نظارت بر نظرات مشتریان، اندازه‌گیری اعتبار برند و درک رفتار مصرف‌کنندگان مورداستفاده قرار می‌گیرد. بااین‌حال، تحلیل احساسات در زبان‌های کم‌منبع مانند فارسی، به دلیل ویژگی‌های زبانی منحصربه‌فرد، با چالش‌های خاصی همراه است. این چالش‌ها شامل وجود کلمات طعنه‌آمیز، عبارات چندمعنایی، ترکیب‌های زبانی و شیوه‌های بیانی پیچیده در زبان فارسی می‌شود. علاوه بر این، تحقیقات موجود در این حوزه عمدتاً بر روی زبان انگلیسی متمرکز شده‌اند و منابع محدودی برای زبان فارسی وجود دارد. این موضوع نیاز به توسعه مدل‌های خاصی را که بتوانند به‌طور مؤثر با چالش‌های زبان فارسی مقابله کنند، بیش‌ازپیش آشکار می‌سازد.

رویکردهای مختلفی برای طبقه‌بندی متن مورداستفاده قرارگرفته می‌شوند که به‌طور کلی می‌توان این رویکردها را در دو گروه تقسیم کرد:

۱-۱ روش‌های مبتنی بر واژگان

در این رویکرد، ویژگی‌های موردنیاز برای طبقه‌بندی متن استخراج می‌شوند و سپس از الگوریتم‌های داده‌کاوی برای طبقه‌بندی استفاده می‌شود. این ویژگی‌ها می‌توانند شامل تعداد کلمات، تعداد تکرار کلمات، تعداد جملات، استفاده از واژگان خاص و غیره باشند. برای این منظور می‌توان روش‌های TF-IDF[1] و روش‌های تعبیه‌سازی کلمات مثل Word2Vec، GLOVE و FASTTEXT را نام برد که ورودی‌های متنی را وزن‌دهی می‌کنند و آن‌ها را به فضای برداری منتقل می‌کنند. در این فضا، کلمات با بردارهای عددی نشان داده می‌شوند و می‌توان از ویژگی‌های این بردارها برای طبقه‌بندی استفاده کرد. لازم به ذکر است که برای آموزش برخی از این مدل‌های تعبیه کلمات نیز از شبکه‌های عصبی مصنوعی استفاده شده است.

۱-۲ روش‌های مبتنی بر یادگیری عمیق

این رویکرد از شبکه‌های عصبی عمیق برای یادگیری و استخراج ویژگی‌های پیچیده و ارتباطات نهفته در داده‌ها استفاده می‌کند. با استفاده از این روش، می‌توان مدل‌هایی برای تشخیص الگوهای پیچیده در متن‌ها و طبقه‌بندی آن‌ها آموزش داد. این روش‌ها از شبکه‌های عصبی عمیق مانند شبکه‌های عصبی بازگشتی (RNN)[2]، شبکه‌های عصبی پیچشی (CNN)[3] و مدل‌های ترنسفورمر استفاده می‌کنند.

ترکیب این رویکردها و استفاده از الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی عمیق می‌تواند به افزایش دقت و کارایی بالاتر در طبقه‌بندی متن منجر شود. همچنین، استفاده از روش‌های پیش‌پردازش متن مانند رمزگذاری و تصحیح خطاها و استفاده از

[1] Term Frequency-Inverse Document Frequency

[2] Recurrent neural networks

[3] Convolutional neural networks

واژگان و مدل‌های زبانی پیش‌آموزش داده شده نیز می‌تواند به بهبود عملکرد طبقه‌بندی متن کمک کند.

در سال‌های اخیر، مدل‌های مبتنی بر معماری‌های ترنسفورمر مانند BERT[1] و پارس‌برت به دلیل توانایی بالا در استخراج ویژگی‌های پیچیده از متون، به‌طور گسترده‌ای در کارهای پردازش زبان طبیعی مورداستفاده قرارگرفته‌اند. بااین‌حال، این مدل‌ها به دلیل تعداد زیاد پارامترها و لایه‌های عمیق، نیازمند منابع محاسباتی قابل توجه و زمان آموزش طولانی هستند. علاوه بر این، مدل‌های ترنسفورمر عمدتاً بر استخراج ویژگی‌های کلی متن تمرکز دارند و ممکن است در استخراج ویژگی‌های محلی (مانند وابستگی‌های کوتاه‌مدت میان کلمات) ضعیف عمل کنند. از سوی دیگر، مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی مانند GRU[2] و LSTM[3]، اگرچه در استخراج ویژگی‌های محلی متن توانمند هستند، اما در یادگیری وابستگی‌های طولانی‌مدت و ویژگی‌های سراسری متن عملکرد ضعیف‌تری دارند.

نوآوری اصلی این پژوهش در ارائه یک معماری ترکیبی و بهینه‌شده برای تحلیل احساسات متون فارسی نهفته است؛ این معماری جدید تلاش می‌کند میان دقت مدل‌های سنگین مبتنی بر ترنسفورمر و نیازهای عملی سیستم‌های واقعی—که اغلب با محدودیت منابع مواجه هستند—تعادل برقرار کند. در این راستا، نخستین گام مهم، بازطراحی ساختار ParsBERT و کاهش عمق آن از ۱۲ لایه به ۵ لایه بوده است. با توجه به بازطراحی انجام‌شده، این کاهش لایه‌ها نه تنها منجر به افت محسوس عملکرد نشده، بلکه با کاهش تعداد پارامترها و افزایش سرعت آموزش، مدل را به گزینه‌ای قابل‌استفاده‌تر در محیط‌های کم‌منبع تبدیل کرده است. سبک‌سازی ParsBERT به این شکل، اقدامی نوآورانه محسوب می‌شود زیرا برخلاف تلاش‌های مشابه که معمولاً به فشرده‌سازی مبتنی بر کوانتیزاسیون یا تقطیر دانش محدود می‌شوند، در این پژوهش ساختار ترنسفورمر بازطراحی شده و عمق آن بهینه‌سازی شده است.

نوآوری دیگر پژوهش، نحوه ترکیب خروجی پارس‌برت با شبکه GRU است. در بسیاری از کارهای پیشین، تنها از بردار[CLS] به‌عنوان نماینده کل جمله استفاده می‌شود؛ درحالی‌که در این مطالعه، برخلاف رویکردهای رایج، کل توالی خروجی ParsBERT وارد GRU شده است. این تصمیم باعث شده است که GRU بتواند الگوهای ترتیبی و وابستگی‌های کوتاه‌برد را بهتر مدل‌سازی کند و به نوعی خلأ ناشی از کاهش عمق ترنسفورمر را جبران نماید. این شیوه ترکیب باعث می‌شود مدل به‌طور هم‌زمان از مزایای استخراج ویژگی‌های سراسری توسط ترنسفورمر و ویژگی‌های محلی توسط GRU بهره‌مند شود؛ قابلیتی که در بسیاری از معماری‌های موجود برای زبان فارسی کمتر مورد توجه قرارگرفته است.

در کنار این موارد، هدف‌گذاری کلی پژوهش نیز جنبه‌ای نوآورانه دارد. برخلاف بخش قابل‌توجهی از تحقیقات موجود که صرفاً بر بهبود دقت تمرکز کرده‌اند، در این پژوهش تلاش شده است مدل طوری طراحی شود که هم برای کاربردهای علمی و هم برای محیط‌های عملیاتی و صنعتی قابل‌استفاده باشد. درواقع، هدف ارائه مدلی است که بدون نیاز به سخت‌افزارهای قدرتمند، عملکردی نزدیک به مدل‌های بزرگ ارائه دهد و درعین‌حال سریع، سبک و مقرون‌به‌صرفه باشد. نتایج تجربی نیز نشان می‌دهد که این هدف محقق شده و مدل پیشنهادی توانسته است با وجود کاهش تعداد پارامترها، عملکردی رقابتی و پایدار ارائه دهد.

درمجموع، نوآوری این پژوهش را می‌توان در بازطراحی ساختار ParsBERT برای کاهش هزینه محاسباتی، استفاده کامل و غیرمتعارف از خروجی توالی ترنسفورمر در ترکیب با GRU و دستیابی به تعادلی کارآمد میان دقت، سرعت و کارایی عملی خلاصه کرد. این ویژگی‌ها مدل پیشنهادی را به رویکردی متفاوت و قابل‌اتکا برای تحلیل احساسات فارسی تبدیل می‌کند.

ساختار این مقاله به شرح زیر است: در بخش ۲، کارهای مرتبط در حوزه تحلیل احساسات و مدل‌های مبتنی بر ترنسفورمر بررسی می‌شود. در بخش ۳، روش‌شناسی و معماری مدل پیشنهادی به‌تفصیل شرح داده می‌شود. در بخش ٤، مجموعه داده‌ها و تنظیمات آزمایش‌ها ارائه می‌شود. در بخش ۵، نتایج و تحلیل‌ها

[1] Bidirectional Encoder Representations from Transformers
[2] Gated Recurrent Unit
[3] Long Short-Term Memory

ارائه‌شده و درنهایت، در بخش ٦، بحث و در بخش ۷، نتیجه‌گیری و پیشنهادهایی برای کارهای آینده ارائه می‌شود.

۲- کارهای مرتبط

در سال‌های اخیر، تحقیقات متعددی در زمینه تحلیل احساسات به زبان‌های مختلف، به‌ویژه انگلیسی، انجام‌شده است. بااین‌حال، مطالعات محدودی در زبان‌های کم‌منبع، مانند فارسی، به‌منظور شناسایی حالات عاطفی صورت گرفته است.

تحلیل احساسات از نخستین حوزه‌های مورد توجه در پردازش زبان طبیعی بوده و در طول چند دهه گذشته از رویکردهای سنتی مبتنی بر ویژگی تا معماری‌های پیچیده یادگیری عمیق و مدل‌های مبتنی بر ترنسفورمر تکامل‌یافته است. نخستین تلاش‌ها برای طبقه‌بندی متن به دهه ۱۹۵۰ بازمی‌گردد. جایی که لُهن در سال ۱۹۵۸ مفاهیم اولیه طبقه‌بندی آماری متن را معرفی کرد [۱]. این رویکردهای کلاسیک عمدتاً بر استخراج ویژگی‌های دستی استوار بودند و نمایشی سطحی از متن ارائه می‌کردند. روش‌های مبتنی بر واژگان و مدل‌های برداری همچون TF-IDF ، GloVe،Word2Vec و FastText به‌مرورزمان به‌عنوان راهکارهایی موثرتر معرفی شدند و توانستند نمایش برداری پایدارتری از واژگان ارائه دهند [۷،۱۲]. بااین‌حال، این روش‌ها همچنان قادر به مدل‌سازی روابط عمیق معنایی و وابستگی‌های بلندبرد نبودند.

با پیشرفت یادگیری عمیق، مدل‌هایی نظیر CNN، LSTM و GRU به‌صورت گسترده در تحلیل احساسات مورداستفاده قرار گرفتند. در [۲]، یک معماری ترکیبی CNN–LSTM برای تحلیل احساسات فارسی پیشنهاد شد که طی آن ویژگی‌های محلی توسط CNNاستخراج و وابستگی‌های زمانی توسط LSTM مدل‌سازی شدند. در [٤] نیز چندین مدل ترکیبی مبتنی بر یادگیری عمیق ارائه شد که نشان دادند استفاده از ترکیب CNN، RNN و SVM می‌تواند نتایج پایدارتری نسبت به استفاده منفرد از یک مدل تولید کند. همچنین در [٥]، مدل ACNN-TL یک شبکه کانولوشنال مبتنی بر توجه همراه با یادگیری انتقال ارائه شد که عملکرد مناسبی در تحلیل احساسات نشان داد. رویکردهای دیگری نیز ارائه‌شده‌اند که به‌صورت اختصاصی بر تحلیل احساسات جنبه‌محور تمرکز داشته‌اند. مانند مدل توجه سلسله‌مراتبی (HAM) در [٦] که برای استخراج ویژگی‌های جنبه و تحلیل قطبیت مورداستفاده قرار گرفت.

در مسیر بهبود کیفیت بازنمایی‌های برداری، رویکردهای توسعه‌یافته‌ای نظیر تعبیه‌های بهبودیافته کلمات [7] (IWV) و ترکیب CNN-BiLSTM-Attention [8] مطرح شدند که هدف آن‌ها ارتقای قابلیت درک مدل از ویژگی‌های سطح کلمه و جمله بود. تحقیقات دیگری به بهبود تحلیل احساسات با استفاده از روش‌های ترکیبی در حوزه‌های غیرفارسی اشاره دارند. ازجمله در [۹]، روشی مبتنی بر BERT برای تحلیل نظرات مرتبط با سهام چینی ارائه شد و در [۱۰] استفاده از SCL (یادگیری مکاتبات ساختاری) برای بهبود انطباق حوزه پیشنهاد گردید. همچنین مطالعات خارجی متعددی به بررسی تأثیر معماری‌های عمیق بر تحلیل احساسات در شبکه‌های اجتماعی پرداخته‌اند.[26–23]

با ظهور مدل‌های مبتنی بر ترنسفورمر، رویکرد تحلیل احساسات وارد مرحله جدیدی شد. مدل BERT با استفاده از سازوکار توجه دوسویه در مرحله پیش‌آموزش [۱۱] توانست نمایش‌های باکیفیت‌تری تولید کند و استفاده از آن در کاربردهای متعدد NLP تثبیت شد. در حوزه زبان فارسی، مهم‌ترین پیشرفت در این حوزه معرفی ParsBERT توسط فرهانی و همکاران [۱۲] بود که با جمع‌آوری یک مجموعه‌داده بزرگ فارسی و آموزش یک مدل ترنسفورمر بومی، کیفیت تحلیل احساسات، طبقه‌بندی متن و شناسایی موجودیت‌های نامدار را به‌طور قابل‌توجهی ارتقا داد. بااین‌حال، مدل‌های مبتنی بر ترنسفورمر همچنان با مشکلات مهمی مواجه هستند؛ ازجمله تعداد پارامترهای بالا، مصرف حافظه زیاد، و نیاز به سخت‌افزار قدرتمند. به همین دلیل تحقیقات متعددی به سمت توسعه روش‌های سبک‌سازی ترنسفورمرها سوق یافته است. در [١٦] و [١٧] راهکارهایی همچون هرس پویا، تقلیل محاسبات و کاهش نیاز حافظه در مرحله تنظیم دقیق ارائه‌شده‌اند. ازجمله روش TokenTune [17] که با انتخاب زیرمجموعه‌ای از توکن‌ها در گذر بازگشتی، مصرف حافظه را به‌طور معناداری کاهش می‌دهد. پژوهش‌های مرتبط با تحلیل احساسات فارسی نیز در کنار توسعه معماری‌های پیچیده، به مسائل کاربردی نظیر کمبود

داده و دشواری یادگیری مدل‌ها در شرایط منابع محدود پرداخته‌اند. به‌عنوان مثال در [۲۲]، روشی برای داده‌افزایی فارسی پیشنهاد شد که با افزایش تنوع داده آموزشی به بهبود عملکرد مدل‌های طبقه‌بندی کمک می‌کند. مطالعات دیگر نیز به ترکیب روش‌های قانون‌محور، یادگیری عمیق و مدل‌های زبانی پرداخته‌اند؛ برای مثال در [۱۸]، ترکیبی از Naive Bayes، قوانین سفارشی و BERT برای تحلیل احساسات توییت‌های فارسی ارائه‌شده است. پژوهش [۱۹] نیز با افزودن یک لایه طبقه‌بند بر روی خروجی BERT، راهکار بهینه‌سازی مشخصی برای تحلیل احساسات ارائه داده است.

مطالعات اخیر در زمینه پردازش زبان طبیعی کارآمد، بر کاهش هزینه محاسباتی مدل‌های ترنسفورمر بزرگ از پیش آموزش‌دیده از طریق دو مسیر اصلی متمرکز شده‌اند: (۱) بهینه‌سازی معماری و (۲) فشرده‌سازی پس از آموزش.

بهینه‌سازی معماری در پی بازطراحی شبکه اصلی است تا به تعادل بهتری بین دقت و کارایی دست یابد. نمونه‌های شاخص نظیر DistilBERT که عمق BERT را با استفاده از تقطیر دانش کاهش می‌دهد و TinyBERT که هم تعداد لایه‌ها و هم ابعاد پنهان را فشرده‌تر می‌کند وALBERT که اشتراک‌گذاری پارامترها و تعبیه‌های عامل‌بندی شده را معرفی می‌کند تا مصرف حافظه را به‌طور چشمگیری کاهش دهد می‌باشند. این مدل‌ها نشان داده‌اند که ترنسفورمرهای کم‌عمق یا فشرده می‌توانند بخش عمده‌ای از قدرت بازنمایی BERT کامل را حفظ کنند و درعین‌حال هزینه‌های استنتاج و آموزش را به میزان قابل‌توجهی کاهش دهند.

دسته دیگری از پژوهش‌ها، بر تکنیک‌های فشرده‌سازی پس از آموزش مانند تقطیر دانش، هرس وزن و کمّی‌سازی با دقت پایین متمرکز است. این روش‌ها هدفشان انتقال دانش از یک مدل معلم بزرگ به یک مدل دانش‌آموز کوچک‌تر یا کاهش دقت عددی برای کاستن از حجم حافظه و تأخیر است، بدون آنکه معماری اصلی تغییر کند.

در مقابل روش‌های فوق، کار ما از راهبرد بهینه‌سازی در سطح معماری پیروی می‌کند. به‌جای اعمال فشرده‌سازی پس از آموزش، ما به‌طور صریح ساختار مدل را با کاهش عمق ParsBERT از پیش آموزش‌دیده بازطراحی می‌کنیم و افت ظرفیت مدل‌سازی ترتیبی ناشی از آن را از طریق یک ماژول بازگشتی سبک‌وزن جبران می‌نماییم. این رویکرد به ما امکان می‌دهد اطلاعات زمینه‌ای را از دنباله کامل توکن حفظ کنیم، درحالی‌که تعداد کل پارامترها و زمان آموزش را پایین نگه می‌داریم. بنابراین، روش ما با روش‌های تقطیر و کمّی‌سازی، رابطه مکمل دارد و می‌تواند در کارهای آتی از آن‌ها بهره‌مند شود.

بررسی مجموعه این پژوهش‌ها نشان می‌دهد که گرچه مدل‌های ترنسفورمر دقت بسیار بالایی ارائه می‌دهند، اما در کاربردهای واقعی به‌ویژه در زبان فارسی که منابع پردازشی و داده‌ای محدودتر است به دلیل هزینه محاسباتی و حافظه‌ای سنگین، استفاده از آن‌ها همواره عملی نیست. علاوه براین، بخش قابل‌توجهی از پژوهش‌ها یا تنها از توکن [CLS] به‌عنوان نمایش جمله استفاده کرده‌اند [۱۹]، یا صرفاً رویکردهای فشرده‌سازی مدل را بررسی کرده‌اند [۱۶،۱۷]، بدون آن که به بازطراحی ساختاری خود معماری پارس‌برت (مانند کاهش عمق رمزگذار) توجه کنند. همچنین مطالعات اندکی وجود دارد که نشان دهند چگونه می‌توان ضعف مدل‌های ترنسفورمر در استخراج ویژگی‌های محلی را با استفاده از شبکه‌های بازگشتی جبران کرد و هم‌زمان هزینه محاسباتی مدل را کاهش داد. ازاین‌رو، جای خالی پژوهشی که بتواند مدلی سبک، دقت‌محور و مناسب برای محیط‌های کم‌منبع ارائه دهد، کاملاً محسوس است.

بنابراین، پژوهش حاضر با بازطراحی ساختاری ParsBERT از طریق کاهش تعداد لایه‌های رمزگذار، انتقال کل توالی خروجی این مدل به لایه‌هایGRU و استفاده هم‌زمان از ویژگی‌های محلی و سراسری، تلاش می‌کند خلأهای موجود در کارهای پیشین را برطرف کرده و راهکاری عملی برای تحلیل احساسات فارسی ارائه دهد.

۳- روش‌شناسی

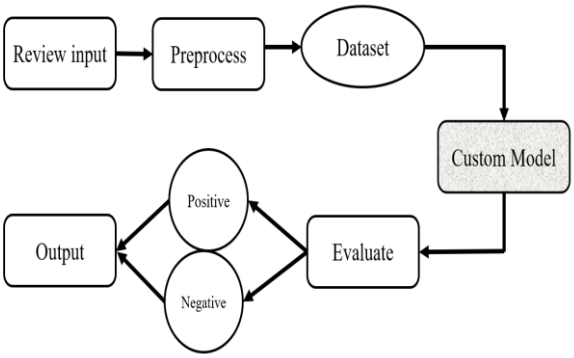
در این پژوهش یک معماری ترکیبی برای تحلیل احساسات در متون فارسی ارائه‌شده است که از ترکیب توانایی‌های مدل زبانی پارس‌برت با شبکه بازگشتی GRU بهره می‌گیرد. ایده اصلی این معماری بر آن استوار است که پارس‌برت قادر به استخراج ویژگی‌های عمیق و معنایی از کل متن است، درحالی‌که GRU

برای مدل‌سازی وابستگی‌های محلی و ترتیبی مناسب‌تر عمل می‌کند.

در مرحله نخست، ورودی‌های متنی پس از توکن‌سازی و آماده‌سازی به مدل پارس‌برت داده می‌شوند و خروجی توالی آخر به دست می‌آید. برخلاف بسیاری از کارهای پیشین که صرفاً از بردار [CLS] استفاده می‌کنند، در این پژوهش کل توالی خروجی به لایه‌های GRU منتقل می‌شود. بدین ترتیب GRU می‌تواند الگوهای زمانی و وابستگی‌های محلی میان توکن‌ها را مدل‌سازی کند.

برای افزایش توانایی مدل در یادگیری روابط عمیق‌تر، از دو لایه GRU دوجهته به‌صورت متوالی استفاده شده است. خروجی GRU دوم سپس با روش میانگین‌گیری بر روی بُعد توالی فشرده‌سازی شده و یک بردار نماینده برای کل جمله ساخته می‌شود. این بردار پس از عبور از یک لایه Dropout و چند لایه تمام‌متصل وارد طبقه‌بند نهایی می‌گردد.

معماری پیشنهادی علاوه بر افزایش توانایی مدل در استخراج هم‌زمان ویژگی‌های سراسری و محلی متن، تعداد لایه‌های ترنسفورمر را نیز کاهش داده است. این موضوع موجب کاهش چشمگیر پارامترها و زمان آموزش می‌شود و امکان استفاده از مدل را در محیط‌های با منابع محاسباتی محدود فراهم می‌سازد. جزئیات فنی مربوط به طراحی، پیاده‌سازی و بهینه‌سازی مدل، به همراه منطق استفاده از روش‌های یادگیری عمیق، در بخش‌های بعدی مقاله به‌طور جامع توضیح داده شده است.



شکل (۱): معماری کلی فرآیند تحلیل احساسات.

۳-۱ پیش‌پردازش مجموعه داده

همان‌طور که در شکل شماره (۱) مشاهده می‌شود، اولین گام پیش‌پردازش داده‌های ورودی است. نظرات کاربران در فضای اجتماعی معمولاً دارای ویژگی‌هایی هستند که چالش‌های متعددی را برای آموزش مدل‌های یادگیری عمیق ایجاد می‌کنند. این ویژگی‌ها شامل استفاده از کلمات عامیانه و غیررسمی، وجود غلط‌های املایی، استفاده از ایموجی‌ها و شکلک‌های مختلف و حتی تکرارهای غیرمعمول در حروف و کاراکترها است. چنین مواردی می‌توانند فرآیند تحلیل و پردازش داده‌ها را پیچیده و زمان‌بر کنند.

برای مقابله با این چالش‌ها، ضروری است که پیش از ارائه داده‌ها به مدل، فرآیند تمیزسازی و پیش‌پردازش انجام شود. برخی از اقدامات مؤثر در این راستا عبارت‌اند از: حذف فاصله‌های غیرضروری میان کلمات برای استانداردسازی داده‌های متنی، حذف ایموجی‌ها و شکلک‌ها، حذف اعداد، برچسب‌ها، آدرس‌های ایمیل و لینک‌های HTTP که معمولاً در تحلیل احساسات و سایر وظایف پردازش زبان طبیعی ضروری نیستند و اصلاح کلمات دارای تکرار غیرمعمول در حروف و کاراکترها (مانند "عاااااللیییی" به "عالی") برای بهبود کیفیت داده‌ها.

این اقدامات، با هدف استانداردسازی در داده‌های متنی، نه‌تنها به بهبود دقت مدل‌های یادگیری عمیق کمک می‌کنند، بلکه پیچیدگی محاسباتی را نیز کاهش می‌دهند.

۳-۲ مدل BERT

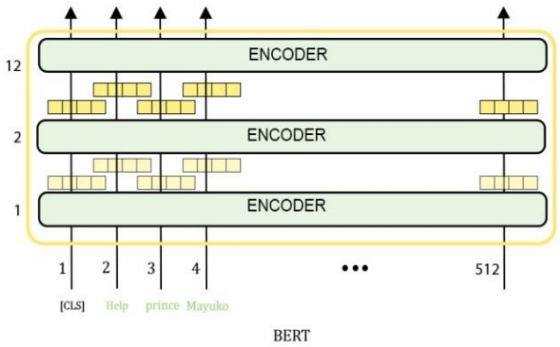
مدل BERT [۲۰] یک مدل زبانی پیش‌آموزش دیده است که توسط تیم پژوهشی شرکت گوگل در سال ۲۰۱۸ توسعه داده شده است و به‌عنوان یکی از مدل‌های پیش‌آموزش دیده موفق در حوزه پردازش زبان طبیعی شناخته می‌شود. مدل BERT مبتنی بر معماری ترنسفورمر و مکانیزم توجه است که با مدل‌سازی هم‌زمان وابستگی‌های بین کلمات، نمایش برداری غنی و زمینه‌محور از توکن‌ها تولید می‌کند. این مدل در دو مرحله پیش‌آموزش و تنظیم نهایی آموزش می‌یابد. در مرحله پیش‌آموزش، با استفاده از روش (MLM)[1]، بخشی از توکن‌ها به‌صورت تصادفی پنهان‌شده و مدل

[1] Masked Language Modeling

با پیش‌بینی آن‌ها، بازنمایی معنایی دقیقی از متن می‌آموزد. در مرحله تنظیم نهایی، مدل با داده‌های برچسب‌دار برای وظایف خاصی مانند طبقه‌بندی متن آموزش داده می‌شود.

ویژگی متمایز مدل BERT، درک دوسویه متن است؛ به‌طوری‌که برخلاف مدل‌های یک‌جهته، بافت کلمات را به‌طور هم‌زمان از هر دو جهت در نظر می‌گیرد و همین امر موجب بهبود عملکرد آن در وظایف مختلف پردازش زبان طبیعی می‌شود. لایه توجه در این مدل با تخصیص وزن به توکن‌ها، میزان اهمیت هر کلمه را در بافت جمله تعیین کرده و امکان استخراج روابط معنایی دقیق‌تر را فراهم می‌سازد.

لایه توجه یکی از اجزای کلیدی مدل BERT است که نقشی حیاتی در پردازش زبان طبیعی ایفا می‌کند. وظیفه اصلی این لایه، تمرکز بر بخش‌های معنادار متن و درک بهتر روابط میان کلمات است. این قابلیت، به مدل کمک می‌کند تا نمایش‌های برداری دقیق‌تری از جملات تولید کند و معنای متن را به‌طور کامل‌تر درک کند.



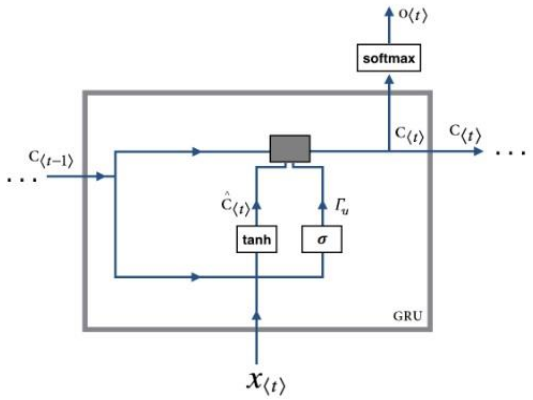
شکل (۲): لایه‌های رمزگذار مدل BERT

همان‌طور که در شکل (۲) مشاهده می‌کنید، در مدل BERT جملات ورودی ابتدا به چندین توکن تقسیم‌شده و هر توکن به یک بردار کلمه تبدیل می‌شود. لایه توجه با استفاده از این بردارها، وزن‌هایی را برای هر توکن محاسبه می‌کند. این وزن‌ها نشان‌دهنده میزان اهمیت هر توکن در بافت جمله هستند. به‌عبارت‌دیگر، لایه توجه مشخص می‌کند که کدام توکن‌ها باید در مراحل بعدی پردازش، توجه بیشتری دریافت کنند.

مدل پارس‌برت [۱۲] یک نسخه از مدل BERT است که به‌طور ویژه برای زبان فارسی آموزش داده شده و با در نظر گرفتن ویژگی‌های ساختاری این زبان، ازجمله تغییرات صرفی، ترکیب حروف و نیم‌فاصله‌ها، عملکرد بهتری نسبت به مدل‌های عمومی چندزبانه در پردازش متون فارسی ارائه می‌دهد.

۳-۳ لایه‌های GRU

GRU یکی از شبکه‌های عصبی بازگشتی برای مدل‌سازی داده‌های دنباله‌ای مانند متن و گفتار است که با هدف کاهش پیچیدگی محاسباتی نسبت به LSTM و حفظ توانایی یادگیری وابستگی‌های بلندمدت ارائه‌شده است. همان‌طور که در شکل (۳) مشاهده می‌کنید، این مدل از دو دروازه اصلی، یعنی دروازه بازنشانی[1] و دروازه به‌روزرسانی[2]، برای کنترل جریان اطلاعات استفاده می‌کند؛ به‌گونه‌ای که دروازه بازنشانی میزان تأثیر اطلاعات گذشته و دروازه به‌روزرسانی میزان حفظ اطلاعات پیشین و ورود اطلاعات جدید را تعیین می‌کند.



شکل (۳): معماری لایه‌های GRU

ساختار ساده‌تر GRU منجر به کاهش تعداد پارامترها و افزایش سرعت آموزش می‌شود و در بسیاری از کاربردها عملکردی قابل‌مقایسه یا حتی بهتر از LSTM، به‌ویژه در شرایط داده و منابع محاسباتی محدود، ارائه می‌دهد. به دلیل این ویژگی‌ها،GRU به‌طور گسترده در حوزه‌هایی مانند پردازش زبان طبیعی، ترجمه ماشینی، تحلیل سری‌های زمانی و تشخیص گفتار مورداستفاده قرار می‌گیرد.

[1] Reset Gate

[2] Update Gate

۳-٤ لایه کلاسبند

لایه کلاسبند در مدل پیشنهادی وظیفه تبدیل ویژگی‌های استخراج‌شده توسط معماری ارائه‌شده به برچسب‌های احساسی را بر عهده دارد. این لایه از چندین بخش تشکیل‌شده است که هر یک نقش خاصی در بهبود عملکرد مدل ایفا می‌کنند. اجزای اصلی این لایه به شرح زیر است:

۳-۴-۱ لایه خطی ابتدایی

این لایه یک تبدیل خطی را با ورودی به‌اندازه دو برابر اندازه ویژگی‌های خروجی مدل پارس‌برت، (hidden_size*2) و خروجی به‌اندازه ۲۵٦ انجام می‌دهد. هدف از این تبدیل خطی، کاهش ابعاد ویژگی‌های ورودی و استخراج اطلاعات مهم از آن‌ها است.

۳-۴-۲ تابع فعال‌سازی

پس از لایه خطی، از تابع فعال‌سازی ReLU[1] استفاده می‌شود. این تابع ورودی‌های منفی را به صفر تبدیل کرده و ورودی‌های مثبت را بدون تغییر باقی می‌گذارد. این عمل باعث افزایش عدم خطی بودن مدل و توانایی آن در یادگیری روابط پیچیده می‌شود.

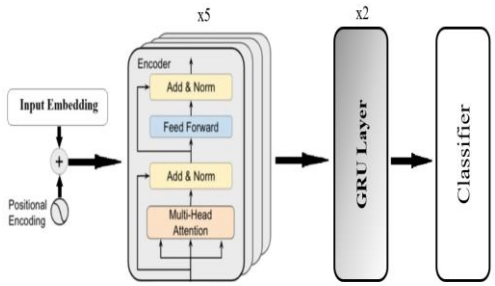
۳-۴-۳ لایه Dropout

برای جلوگیری از بیش‌برازش، از لایه Dropout با احتمال ۰/۲ استفاده شده است. این لایه به‌صورت تصادفی برخی از ویژگی‌ها را به صفر می‌رساند و با کاهش تعداد پارامترهای قابل یادگیری، تعمیم‌پذیری مدل را افزایش می‌دهد.

۳-۴-۴ لایه خطی نهایی

در این مرحله، یک لایه خطی دیگر با ورودی به‌اندازه ۲۵٦ (خروجی لایه قبلی) و خروجی به‌اندازه تعداد برچسب‌های احساسی (config.num_labels) به کار گرفته می‌شود. خروجی این لایه احتمال‌های مرتبط با برچسب‌های احساسی است که بر اساس ویژگی‌های استخراج‌شده توسط مدل پارس‌برت و تبدیل‌های خطی محاسبه می‌شود. این معماری کلاسبند به مدل امکان می‌دهد تا بر اساس ویژگی‌های استخراج‌شده از متن، برچسب احساسی متن را با دقت بالایی پیش‌بینی کند. به‌طور خلاصه، معماری مدل مبتنی بر پارس‌برت و GRU ارائه‌شده در این مقاله دارای دو نکته جدید به شرح زیر است:

انتخاب مدل‌های یادگیری عمیق مناسب جهت ترکیب با یکدیگر. مدل‌های مبتنی بر ترنسفورمر، مانند پارس‌برت، توانایی بالایی در استخراج ویژگی‌های سراسری از متن دارند. از سوی دیگر، شبکه‌های عصبی کانولوشنی و بازگشتی (مانند GRU) قابلیت بسیار خوبی در استخراج ویژگی‌های محلی دارند. با ترکیب این مدل‌ها، می‌توان از مزایای استخراج ویژگی‌های محلی و سراسری به‌طور هم‌زمان بهره‌مند شد. این ترکیب منجر به بهبود دقت و عملکرد مدل در تحلیل احساسات می‌شود. معماری کلی مدل ارائه‌شده در شکل (٤) قابل‌مشاهده است.



شکل (٤): معماری مدل ارائه‌شده

همان‌طور که در شکل (٤) مشاهده می‌شود، در مدل ترکیبی پیشنهادی، خروجی کامل توالی حاصل از پارس‌برت مستقیماً به لایه‌های GRU داده می‌شود. برخلاف برخی رویکردها که تنها از بردار [CLS] استفاده می‌کنند، ما تمامی بردارهای توکن را به GRU انتقال می‌دهیم تا این شبکه بتواند وابستگی‌های ترتیبی و محلی را روی نمایش‌های معنایی استخراج‌شده توسط پارس‌برت مدل‌سازی کند. به‌طور مشخص، دو لایه GRU دوجهته به‌صورت متوالی به کار گرفته‌شده‌اند؛ به‌این‌ترتیب که خروجی اولین GRU به‌عنوان ورودی GRU دوم عمل می‌کند. سپس، برای فشرده‌سازی خروجی نهایی GRU، عملگر میانگین‌گیری بر روی بُعد توالی اعمال‌شده و یک بردار نماینده از کل جمله به دست می‌آید. این بردار پس از گذر از یک لایه Dropout و چند لایه تمام‌متصل وارد

[1] Rectified Linear Unit

طبقه‌بند نهایی می‌شود. این طراحی امکان استخراج هم‌زمان ویژگی‌های کلی سراسری و محلی متن را فراهم می‌سازد.

مدل‌های مبتنی بر ترنسفورمر به دلیل تعداد زیاد پارامترها، پیچیدگی محاسباتی بالایی دارند. برای کاهش این پیچیدگی، تغییراتی در لایه‌ها و معماری مدل پارس‌برت اعمال‌شده است. این تغییرات علاوه بر کاهش تعداد پارامترها، زمان آموزش مدل را نیز به میزان قابل‌توجهی کاهش داده است. این بهبودها بدون افت دقت، کارایی مدل را در پردازش زبان طبیعی افزایش داده است.

٤- آزمایش‌ها و نتایج

برای ارزیابی و تأیید اعتبار روش پیشنهادی، آزمایش‌هایی بر روی مجموعه داده‌های اسنپ‌فود و طاقچه انجام‌شده است. فرآیند آموزش مدل روی محیط Google Colab صورت پذیرفت که دارای پردازنده گرافیکی NVIDIA Tesla T4 با ١٦ گیگابایت حافظه گرافیکی، ٣٢ گیگابایت RAM و پردازنده مرکزی مجازی شده از نوع Intel Xeon بود. این بستر محاسباتی برای آموزش مدل در مقیاس متوسط و در زمان قابل‌قبول مناسب ارزیابی شد. با توجه به منابع محدود، طراحی مدل به‌گونه‌ای انجام شد که با کاهش تعداد پارامترها و استفاده از لایه‌های سبک‌تر، مصرف حافظه به حداقل برسد. در این قسمت نتایج حاصل از این آزمایش‌ها همراه با تحلیل عملکرد روش پیشنهادی ارائه می‌شود تا کارایی و دقت مدل در پردازش و تحلیل احساسات مورد ارزیابی قرار گیرد. جزئیات روش آزمایش و معیارهای ارزیابی در ادامه بحث شده و مقایسه‌ای جامع میان نتایج مدل پیشنهادی و سایر روش‌های موجود ارائه خواهد شد.

٤-١ معیارهای ارزیابی

در این پژوهش، معیارهای مختلفی نظیر دقت، صحت، حساسیت و معیار F1 برای ارزیابی عملکرد مدل ارائه‌شده استفاده شده است. این معیارها کمک می‌کنند تا قدرت و دقت مدل را در دسته‌بندی داده‌های مختلف سنجیده شوند. دقت، میزان پیش‌بینی‌های صحیح مدل را نسبت به کل پیش‌بینی‌ها محاسبه می‌کند:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (١)$$

صحت، نشان‌دهنده نسبت پیش‌بینی‌های مثبت درست به تمام پیش‌بینی‌های مثبت است و به شکل زیر تعریف می‌شود:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (٢)$$

حساسیت، نشان‌دهنده نسبت پیش‌بینی‌های صحیح مثبت به کل نمونه‌های مثبت است:

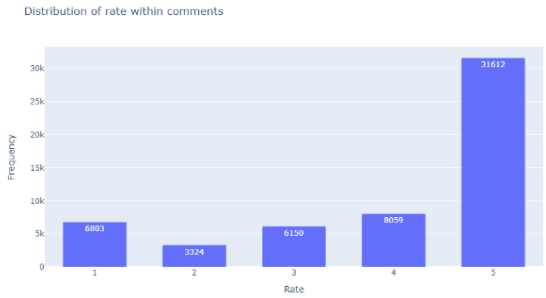
$$\text{Recall} = \frac{TP}{TP + FN} \quad (٣)$$

معیار F1، میانگین هارمونیک صحت و حساسیت است که نشان‌دهنده تعادل خوبی بین این دو معیار است و به شکل زیر محاسبه می‌شود:

$$F1 = \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \times 2 \quad (٤)$$

٤-٢ معرفی مجموعه داده‌ها

در این پژوهش برای حصول اطمینان از عملکرد مدل ارائه داده شده، از دو مجموعه داده معروف در زبان فارسی به نام‌های طاقچه و اسنپ‌فود استفاده شده است. مجموعه داده طاقچه حاوی بالغ بر ٦٧٦٠٠ ردیف ازنظرات مشتریان درباره محصولات مختلف در فروشگاه اینترنتی کتاب است که در سال ٢٠١٩ از طریق خزش وب ایجادشده است. در این مجموعه داده نظرات مشتریان در رابطه با هر محصول از ١ تا ٥ رتبه‌بندی شده‌اند.



شکل (٥): مجموعه داده طاقچه

با توجه به کار تحلیل احساسات، در این پژوهش مقادیر برچسب‌های کمتر یا مساوی ٣ را به‌عنوان نظر منفی و مقادیر بالاتر از ٣ را به‌عنوان نظرات مثبت تبدیل کرده و آن را به‌صورت باینری در نظر گرفته‌ایم و همان‌طور که در شکل زیر مشاهده می‌شود این مجموعه داده متعادل نیست که می‌توان با استفاده از روش‌های مختلف آن را متعادل کرد. طول بزرگ‌ترین جمله در این مجموعه

داده ۲۵۸۸ توکن و ۸۰/۳۲ درصد جملات بیشتر از ۳ کلمه وکمتر از ۱۲۸ کلمه دارند. مجموعه داده دیگری که برای آموزش مدل پیشنهادی مورداستفاده قرارگرفته است اسنپ‌فود است که حاوی ۷۰۰۰۰ نظرات کاربران شرکت اسنپ‌فود با دو برچسب مثبت و منفی است و ۹۹/۴۳ درصد جملات بیشتر از ۳ و کمتر از ۱۲۸ کلمه دارند.

هرچند مجموعه داده طاقچه ازنظر تنوع موضوعی و زبانی ارزشمند است، اما محدودیت اصلی آن، توزیع برچسب‌های مثبت و منفی در این مجموعه نامتوازن است. برای رفع این مشکل، در این پژوهش از تکنیک افزایش داده مبتنی بر جابجایی کلمات استفاده شده است که باعث بهبود تعادل میان کلاس‌ها گردید.

جدول (۱): مقادیر مجموعه داده‌ها

۳ < کلمات < ۱۲۸	تعداد نمونه	برچسب	مجموعه داده
۸۰/۳۲٪	۶۸۰۳	۱	طاقچه
	۳۳۲۴	۲	
	۶۱۵۰	۳	
	۸۰۵۹	۴	
	۳۱۶۱۲	۵	
۹۹/۴۳٪	۳۵۰۰۰	مثبت	اسنپ‌فود
	۳۵۰۰۰	منفی	

همچنین استفاده از دو مجموعه داده طاقچه و اسنپ‌فود باعث شد که مدل پیشنهادی نه تنها روی متونی با زبان رسمی‌تر و ساختارمند (نقد کتاب‌ها) آموزش ببیند، بلکه با داده‌هایی غیررسمی، محاوره‌ای و روزمره (نظرات کاربران درباره خدمات غذایی) نیز مواجه شود. به‌این‌ترتیب، مدل توانست در شرایط متنوع‌تری آموزش ببیند و برای استفاده عملی در محیط‌های واقعی مناسب‌تر شود.

در هر دو مجموعه داده برای هر احساس، بیشتر پیام‌ها حدود ۱۵ کلمه هستند و طولانی‌ترین توییت‌ها بسیار کمتر از مقدار ۷۶۸ که حداکثر اندازه زمینه پارس‌برت است می‌باشند. متن‌هایی که طولانی‌تر از اندازه بافت مدل هستند، باید کوتاه شوند که اگر متن کوتاه‌شده حاوی اطلاعات مهمی باشد، می‌تواند منجر به کاهش عملکرد شود. در این مورد، به نظر می‌رسد که مشکلی نخواهد بود.

۴-۳ تنظیمات پارامترهای آموزش

در این پژوهش از بهینه‌ساز AdamW استفاده شده است، زیرا این روش با تصحیح کاهش وزن[1] در مقایسه با Adam[2] معمولی، نتایج پایدارتری در مدل‌های پردازش زبان طبیعی ارائه می‌دهد. همچنین حداکثر طول دنباله ورودی را ۱۲۸ و اندازه دسته ۲۴ انتخاب‌شده است، زیرا مقدارهای کوچک‌تر مانند ۱۶ منجر به افزایش نویز گرادیان‌ها و کاهش پایداری آموزش شدند، درحالی‌که مقادیر بالاتر مانند ۶۴ نیاز به حافظه پردازشی بیشتر داشتند که اجرای مدل را در برخی سخت‌افزارها محدود می‌کرد. تعداد دوره‌های آموزشی ۳ و نرخ یادگیری اولیه 1e-۵ تنظیم‌شده است زیرا مدل پیشنهادی شامل یک بخش از پیش‌آموزش‌یافته پارس‌برت و بخش‌های جدید GRU و کلاسبند نهایی است. استفاده از نرخ یادگیری پایین در مدل‌های از پیش‌آموزش‌یافته، باعث جلوگیری از تخریب وزن‌های از قبل یادگرفته‌شده می‌شود و به مدل اجازه می‌دهد تا بهینه‌سازی تدریجی و پایدارتری داشته باشد. نرخ‌های بالاتر مانند 1e-۴ در آزمایش‌های اولیه منجر به بی‌ثباتی در همگرایی و افت عملکرد مدل شد، درحالی‌که نرخ‌های بسیار پایین‌تر مانند 1e-۶ باعث کندی در یادگیری و نیاز به تعداد بیشتری از دوره‌های آموزشی شد. همچنین از یک زمان‌بند پویا برای تنظیم خودکار میزان نرخ یادگیری در طول فرآیند آموزش مدل پیشنهادی و از CrossEntropyLoss به‌عنوان تابع محاسبه خطا استفاده شده است.

۴-۴ معماری مدل

در این پژوهش، مدل پارس‌برت به‌عنوان مدل پایه مورداستفاده قرارگرفته است. این مدل شامل ۱۲ لایه رمزگذار است که وظیفه استخراج ویژگی‌های متن ورودی را بر عهده دارند. در راستای بهینه‌سازی عملکرد مدل، تعداد لایه‌های رمزگذار به ۵ لایه کاهش‌یافته است. همچنین، از تابع فعال‌ساز GELU_NEW در لایه‌های رمزگذار استفاده شده و مقدار پارامتر hidden_dropout_prob به ۰/۲ افزایش داده شده است. علاوه بر این تغییرات، دو لایه GRU به مدل اضافه‌شده است تا قابلیت

[1] Weight Decay Correction

[2] Adaptive Moment Estimation

استخراج ویژگی‌های محلی متن نیز فراهم شود. درنهایت، یک لایه کلاسبند برای پیش‌بینی کلاس مربوطه در نظر گرفته‌شده است.

همان‌طور که گفته شد مدل پارس‌برت به دلیل تعداد پارامترهای زیاد، از پیچیدگی محاسباتی بالایی برخوردار است که منجر به افزایش زمان و هزینه آموزش می‌شود. بااین‌حال، این مدل به دلیل استفاده از مکانیزم (WWM)[1]، عملکرد مطلوبی در پردازش کلمات کوتاه و جملات ناقص دارد. در این مکانیزم، به جای ماسک کردن هر توکن به‌صورت جداگانه، کل کلمه ماسک می‌شود. این روش باعث می‌شود که مدل درک بهتری از مفهوم کلی کلمات در جمله داشته باشد و درنتیجه، دقت بالاتری در وظایف پردازش زبان طبیعی ارائه دهد. از سوی دیگر، مدل‌های یادگیری عمیق سنتی مانند CNN، LSTM و GRU اگرچه در استخراج ویژگی‌های محلی متن عملکرد مناسبی دارند، اما فاقد قابلیت بهره‌گیری از مکانیزم توجه هستند. به‌منظور کاهش پیچیدگی محاسباتی مدل پارس‌برت و بهره‌گیری از مزایای هر دو دسته مدل‌های یادگیری عمیق، در این پژوهش تغییراتی در معماری لایه‌های مدل پارس‌برت اعمال‌شده و با مدل GRU ترکیب‌شده است. همچنین، نوع کلاسبند تغییریافته تا عملکرد مدل بهینه شود. در ادامه، مدل پیشنهادی با معماری‌های مختلف و مدل‌های یادگیری متفاوت پیاده‌سازی و مقایسه شده است. نتایج حاصل از آزمایش‌ها نشان‌دهنده بهبود عملکرد مدل پیشنهادی و کاهش پیچیدگی محاسباتی آن در مقایسه با مدل‌های پایه است.

٤-٥ آزمایش حساسیت تعداد لایه‌های رمزگذار

به‌منظور بررسی میزان تأثیر تعداد لایه‌های رمزگذار ParsBERT بر عملکرد و هزینه محاسباتی مدل، یک آزمایش حساسیت طراحی شد. هدف از این آزمایش تعیین تعداد لایه‌هایی است که بهترین تعادل میان دقت، سرعت آموزش و تعداد پارامترها را فراهم می‌کند. برای این منظور، مدل با ساختار پیشنهادی پژوهش در چهار پیکربندی مختلف شامل ۳، ٤، ٥ و ٦ لایه رمزگذار مورد ارزیابی قرار گرفت.

در هر پیکربندی، سایر اجزای مدل ثابت نگه داشته شد تا تأثیر تعداد لایه‌ها به‌صورت مستقل بررسی شود. نتایج به‌دست‌آمده برای مجموعه‌داده طاقچه و اسنپ‌فود در جدول زیر گزارش‌شده است.

جدول (۲): نتایج آزمایش حساسیت تعداد لایه‌های رمزگذار

تعداد لایه رمزگذار	تعداد پارامترها	زمان آموزش	F1 (طاقچه)	F1 (اسنپ‌فود)
۳	بسیار کم	بسیار سریع	۰/۷٤۲	۰/۸٤۲
٤	کم	سریع	۰/۷۵۸	۰/۸۵۵
٥	کاهش‌یافته	بهینه‌ترین	۰/۷٦۷	۰/۸٦۱
٦	بالاتر	زمان بیشتر	۰/۷٦٦	۰/۸٦۲

نتایج جدول نشان می‌دهد که تعداد لایه‌های رمزگذار تأثیر مستقیمی بر عملکرد مدل و زمان آموزش دارد. در معماری‌های با ۳ و ٤ لایه، به دلیل کاهش بیش‌ازحد عمق شبکه، توانایی مدل در استخراج روابط سراسری متن کاهش‌یافته و عملکرد F1 کمی افت کرده است. با افزایش تعداد لایه‌ها به ٥، مدل توانسته است نمایش‌های معنایی مناسب‌تری تولید کرده و بهترین توازن میان دقت و هزینه محاسباتی را ارائه دهد.

افزایش لایه‌ها به ٦ نیز منجر به بهبود بسیار جزئی دقت شده است، اما افزایش زمان آموزش و تعداد پارامترها نسبت به مدل ٥ لایه، این انتخاب را ازنظر عملی توجیه‌پذیر نمی‌سازد. بنابراین، معماری ٥ لایه به‌عنوان گزینه بهینه انتخاب شده است، چراکه ضمن حفظ عملکرد رقابتی، هزینه محاسباتی را کاهش داده و زمان آموزش را در سطح مطلوبی نگه می‌دارد.

٤-٦ آزمایش حساسیت تعداد لایه‌های BiGRU

همان‌طور که در جدول (۳) مشاهده می‌شود، تعداد لایه‌های GRU به‌صورت تجربی با مقایسه ساختارهای یک‌لایه، دولایه و سه‌لایه تعیین‌شده است. نتایج نشان می‌دهد یک لایه ظرفیت کافی برای مدل‌سازی الگوهای ترتیبی سطح‌بالا را ندارد، درحالی‌که سه لایه تنها بهبود ناچیزی ایجاد کرده و هم‌زمان موجب افزایش بیش‌برازش و ناپایداری آموزش می‌شود. پیکربندی دولایه بهترین

[1] Whole Word Masking

توازن میان قدرت بازنمایی، تعمیم‌پذیری و هزینه محاسباتی را فراهم کرده است.

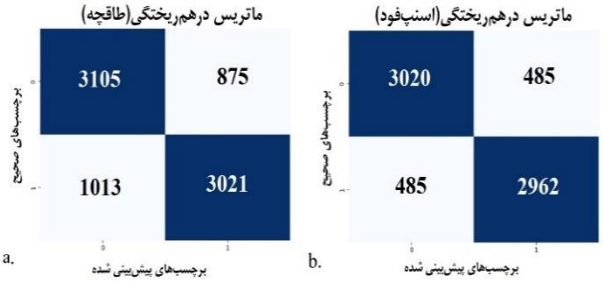
جدول (۳): نتایج آزمایش حساسیت تعداد لایه‌های BiGRU

تعداد لایه BiGRU	میانگین F1 ± انحراف معیار (طاقچه)	میانگین F1 ± انحراف معیار (اسنپ‌فود)
۱	۰/۷۳۶ ± ۰/۰۰۵	۰/۸۵۳ ± ۰/۰۰۶
۲	۰/۷۶۷ ± ۰/۰۰۳	۰/۸۶۱ ± ۰/۰۰۴
۳	۰/۷۴۷ ± ۰/۰۰۴	۰/۸۶۰ ± ۰/۰۰۷

۵- تجزیه‌وتحلیل نتایج

۵-۱ ماتریس درهم‌ریختگی

ماتریس سردرگمی ابزاری ارزشمند برای ارزیابی عملکرد مدل‌های طبقه‌بندی است. با استفاده از این ماتریس، پیش‌بینی‌های مدل را می‌توان مستقیماً با مقادیر واقعی مقایسه کرد. شکل (٦) ماتریس سردرگمی برای هر دو مجموعه داده استفاده شده را نشان می‌دهد که در زیر به‌تفصیل مورد تجزیه‌وتحلیل قرارگرفته است.

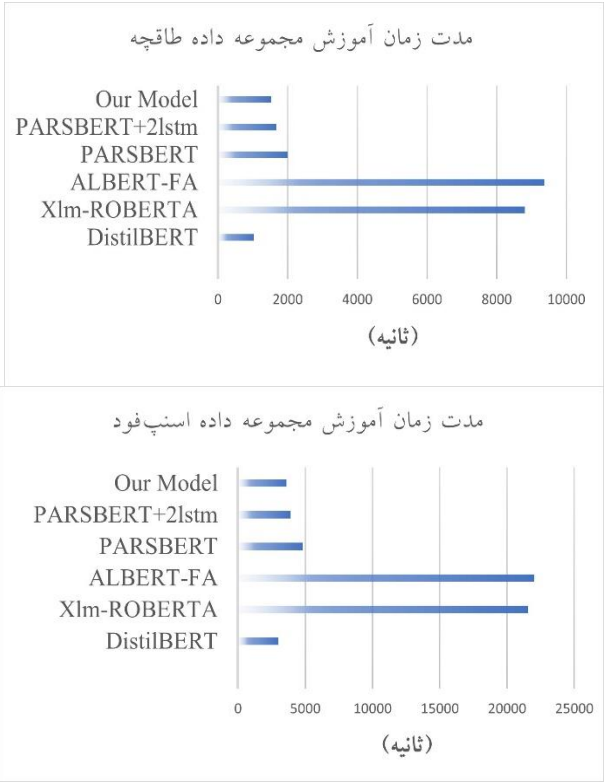


شکل (٦): ماتریس درهم‌ریختگی

این ماتریس سردرگمی نشان می‌دهد که مدل پیشنهادی به‌دقت طبقه‌بندی بسیار بالایی دست می‌یابد. تعداد نمونه‌های طبقه‌بندی‌شده نادرست (مثبت کاذب و منفی کاذب) حداقل است که بر استحکام و قابلیت اطمینان مدل در شناسایی صحیح نمونه‌ها در هر دو مجموعه داده تأکید می‌کند. این عملکرد استثنایی تضمین می‌کند که این مدل برای کاربردهای دنیای واقعی مناسب است.

۵-۲ مدت زمان آموزش مدل

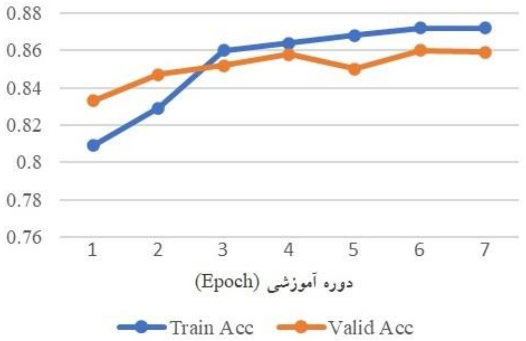
در شکل (۷) مقایسه‌ای از مدت زمان آموزش مدل توسط دو مجموعه داده مختلف نمایش داده شده است. همان‌طور که مشاهده می‌شود مدت زمان آموزش مدل ارائه‌شده کاهش چشمگیری نسبت به مدل پایه دارد.



شکل (۷): مدت زمان آموزش بر روی مدل‌های مختلف

۵-۳ مقایسه دقت آموزش و دقت اعتبارسنجی

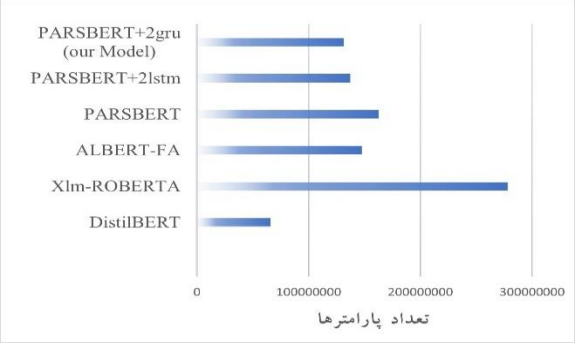
در شکل (۸) مقایسه دقت آموزش با دقت داده‌های اعتبارسنجی بر روی مدل ارائه‌شده نمایش داده شده است. همان‌طور که مشاهده می‌شود دقت داده‌های اعتبارسنجی با شیبی مناسب همراه با دقت داده‌های آموزشی رشد می‌کند که نمایانگر این است که مدل ارائه‌شده دچار فرابرازش نشده است.



شکل (۸): مقایسه دقت آموزش و دقت اعتبارسنجی

۵-٤ تعداد پارامترهای مدل

در شکل (۹) تعداد پارامترهای مدل‌های مختلف با یکدیگر مقایسه شده است. همان‌طور که مشاهده می‌شود تعداد پارامترهای آموزشی مدل ارائه‌شده کاهش چشمگیری نسبت به مدل پایه دارد.



شکل (۹): تعداد پارامتر مدل‌های مختلف

۵-۵ معیار F1

در جدول (٤) نتایج معیار F1 بر روی دو مجموعه داده مختلف مقایسه شده است. همان‌طور که مشاهده می‌کنید معیار F1 به‌دست‌آمده توسط مدل ارائه‌شده در این پژوهش، بر روی مجموعه داده طاقچه بالاترین مقدار را در مقایسه با سایر مدل‌ها به دست آورده است. در خصوص مجموعه داده اسنپ‌فود نیز نتیجه‌ای تقریباً معادل با مدل پایه پارس‌برت به‌دست‌آمده که با توجه به تعداد پارامتر کمتر و مدت زمان آموزش کمتر، برتری مدل ارائه‌شده را در مقایسه با سایر مدل‌ها اثبات می‌کند.

۵-٦ معیار ROC-AUC

منحنی ROC و معیار AUC را می‌توان برای ارزیابی عملکرد کلی مدل و توانایی آن در تمایز بین کلاس‌های مثبت و منفی به کار گرفت. معیار AUC ارزیابی جامعی از عملکرد مدل در تمام سطوح آستانه تصمیم‌گیری ارائه می‌دهد و چشم‌انداز کلی‌تری از کیفیت طبقه‌بندی مدل به دست می‌دهد. همان‌گونه که در جدول (۵) مشاهده می‌شود، مدل پیشنهادی در مجموعه اسنپ‌فود به مقدار AUC برابر با ۰/۸٦۰۵ دست‌یافته است که به‌طور معناداری از مدل‌های سبک نظیر DistilBERT و ALBERT-FA بالاتر بوده و با وجود کاهش چشمگیر تعداد پارامترها، تنها فاصله بسیار اندکی با ParsBERT دارد. این نتیجه نشان می‌دهد که معماری پیشنهادی، ضمن حفظ توان تفکیک بالا، ازنظر کارایی محاسباتی نیز بسیار بهینه‌تر است.

در مجموعه طاقچه که ازنظر زبانی و توزیع احساسات چالش‌برانگیزتر است، برتری مدل پیشنهادی آشکارتر می‌شود؛ به‌طوری‌که مقدار AUC برابر با ۰/۷٦٤۷ به‌دست آمده که از تمامی مدل‌های مقایسه‌شده، ازجمله ParsBERT پایه و نسخه مبتنی بر LSTM بیشتر است. این امر نشان می‌دهد که استفاده از ساختار BiGRU بر روی خروجی‌های توکن‌محور ترنسفورمر کم‌عمق، توان مدل را در تشخیص الگوهای ترتیبی و وابستگی‌های دوطرفه به‌طور مؤثری افزایش داده است.

جدول (۵): ارزیابی مدل ارائه‌شده بر روی مجموعه‌داده‌های طاقچه و اسنپ‌فود

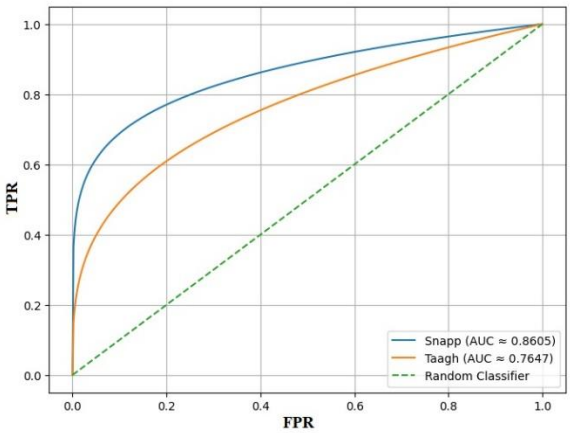
مدل	AUC (طاقچه)	AUC (اسنپ‌فود)
DistilBERT	۰/۷۳٤	۰/۸۰۷
XLM-RoBERTa	۰/۷٤۹	۰/۸۱۷
ALBERT-FA	۰/۷۲۳	۰/۷۹۲
ParsBERT	۰/۷٦۲	۰/۸۷۲
LSTM۲ParsBERT +	۰/۷۲۵	۰/۸۵٤
BiGRU۲ParsBERT + (مدل پیشنهادی)	۰/۷٦٤۷	۰/۸٦۰۵

هر نقطه روی منحنی ROC نشان‌دهنده یک جفت مقدار نرخ مثبت واقعی (TPR) و نرخ مثبت کاذب (FPR) است. خط مورب از نقطه (۰, ۰) تا (۱, ۱) نشان‌دهنده حدس تصادفی است؛ هر چه منحنی ROC از این خط دورتر باشد، عملکرد مدل بهتر است. مساحت زیر منحنی (AUC) سطح زیر منحنی ROC را کمّی‌سازی کرده و توانایی مدل در تفکیک کلاس‌های مثبت و منفی را نشان می‌دهد. مقدار AUC برابر با ۰/۵ نشان‌دهنده حدس تصادفی است، درحالی‌که AUC برابر با ۱ بیانگر تفکیک کامل کلاس‌ها است. درمجموع، نتایج ROC–AUC تأیید می‌کنند که مدل پیشنهادی، در مقایسه با سایر معماری‌ها، بهترین توازن میان دقت تفکیک، پایداری تصمیم‌گیری و پیچیدگی محاسباتی را ارائه می‌دهد و

به‌ویژه در سناریوهای با منابع محدود، جایگزین مناسبی برای مدل‌های حجیم ترنسفورمری محسوب می‌شود.

جدول (٤): مقایسه نتایج مدل ارائه‌شده

مدل	تعداد پارامترها	مدت زمان آموزش (ثانیه)		معیار F1	
		اسنپ‌فود	طاقچه	اسنپ‌فود	طاقچه
DistilBERT	٦٦,٠٠٠,٠٠٠	٢٩٨٣/٤٢٨٥٣٧	١٠٢٤/٧٢٨٤١٥	٠/٧١٩٣	٠/٦٨٢٥
Xlm-ROBERTA	٢٧٨,٠٤٥,١٨٦	٢١٥٧٩/٧١٢٤٥١	٨٨١٥/٧٤٨٢٣٥	٠/٨٤٣١	٠/٧٥٣٣
ALBERT-FA	١٤٧,٤٨٢,٨٨٢	٢٢٠٣٥/٦١٧٢٤٥	٩٣٦٦/٣٠١٤٦٧	٠/٨٣٩٥	٠/٧٣٧١
PARSBERT	١٦٢,٨٤٢,٨٨٢	٤٧٩٩/٤٩٣٠٤٧	٢٠٠٥/٩٠٦٧٥٥	٠/٨٦٥٦	٠/٧٥٣١
PARSBERT+2lstm	١٣٧,٢٣٧,٧٦٢	٣٩٠٧/١٦١٧٢٤	١٦٧٠/٦٧٠٠٥	٠/٨٦٠٨	٠/٧٦٢٣
PARSBERT+2gru (our Model)	١٣١,٣٣٣,٣٧٨	٣٦٠٩/٣٢٤٤	١٥٢٨/٤١٦٢٤٤	٠/٨٦١٦	٠/٧٦٧



شکل (١٠): منحنی ROC مدل پیشنهادی

٦- بحث

نتایج آزمایش‌ها نشان می‌دهد که مدل پیشنهادی توانسته است با کاهش قابل‌توجه تعداد پارامترها و زمان آموزش، عملکردی رقابتی در تحلیل احساسات متون فارسی ارائه دهد. مدل درمجموعه‌داده طاقچه بهترین مقدار F1 را کسب کرده است که دلیل آن ساختار رسمی‌تر و منسجم‌تر این داده‌هاست؛ ویژگی‌هایی که با نمایش‌های معنایی پارس‌برت سازگاری بیشتری دارند. در مقابل، داده‌های اسنپ‌فود دارای سبک محاوره‌ای، غلط‌های املایی و تنوع زبانی بالاتری هستند که موجب کاهش جزئی دقت همه مدل‌ها شده است. بااین‌حال، مدل پیشنهادی عملکردی نزدیک به پارس‌برت پایه ارائه کرده و این موضوع نشان‌دهنده پایداری مدل در شرایط داده‌های نویزی است.

کاهش تعداد لایه‌های رمزگذار پارس‌برت از ۱۲ به ۵، در کنار استفاده از GRU، باعث کاهش چشمگیر پارامترها و زمان آموزش شده است. مدل پیشنهادی حدود ۲۰–۲۵٪ سریع‌تر آموزش می‌بیند، درحالی‌که عملکرد آن در سطح رقابتی باقی مانده است. این ویژگی مدل را برای کاربردهای عملی و دستگاه‌های با منابع محدود مناسب می‌سازد. اگرچه در داده‌های اسنپ‌فود مقدار F1 اندکی کمتر از مدل پایه است، اما نسبت به کاهش قابل توجه پیچیدگی محاسباتی، این افت کوچک قابل‌قبول و منطقی است. مدل پیشنهادی بهترین نقطه تعادل میان سرعت، مصرف منابع و دقت را ارائه می‌دهد و برای سیستم‌های واقعی که پاسخ‌گویی سریع اهمیت دارد، گزینه‌ای کارآمد محسوب می‌شود.

اندازه بردار پنهان GRU با مقادیر ۱۲۸، ۲۵٦ و ۵۱۲ بررسی شده است. مقدار ۱۲۸ منجر به گلوگاه اطلاعاتی و افت دقت شده، درحالی‌که مقدار ۵۱۲ افزایش قابل‌توجهی در تعداد پارامترها ایجاد کرده بدون آن‌که بهبود معناداری در عملکرد حاصل شود. مقدار

۲۵۶ به‌عنوان نقطه تعادل میان ظرفیت بازنمایی و پیچیدگی محاسباتی انتخاب شده است. علاوه بر آن، برای کنترل بیش‌برازش، نرخ‌های Dropout برابر با ۰/۱، ۰/۲ و ۰/۳ ارزیابی شده‌اند. نرخ ۰/۱ منجر به بیش‌برازش خفیف، و نرخ ۰/۳ موجب کاهش ظرفیت مدل و کم‌برازش شده است. مقدار ۰/۲ بهترین عملکرد و پایداری را در هر دو مجموعه‌داده فراهم کرده و به‌عنوان مقدار بهینه انتخاب شده است.

هم‌چنین ساختارهای GRU، LSTM و نسخه‌های دوسویه آن‌ها را با یکدیگر مقایسه کردیم. نتایج نشان می‌دهد LSTM با وجود پیچیدگی و تعداد پارامتر بیشتر، بهبود معناداری نسبت به GRU ایجاد نمی‌کند. در مقابل، استفاده از ساختار دوسویه (BiGRU) به دلیل دسترسی هم‌زمان به وابستگی‌های پیشرو و پسرو، دقت بالاتری نسبت به نسخه‌های یک‌سویه فراهم کرده است. ازاین‌رو، BiGRU به‌عنوان بهترین گزینه ازنظر توازن میان دقت و هزینه محاسباتی انتخاب شده است.

برای بررسی معناداری آماری بهبودهای حاصل، آزمون t زوجی بر روی مقادیر F1 در ۵ اجرای مستقل انجام شد. نتایج نشان داد اختلاف عملکرد مدل پیشنهادی و ParsBERT در مجموعه اسنپ‌فود با مقدار $p = ۰/۰۳$ ازنظر آماری معنادار است ($p < 0/05$)، درحالی‌که در مجموعه طاقچه این اختلاف حتی معنادارتر نیز بوده است.

داده‌افزایی اعمال‌شده بر مجموعه‌داده طاقچه توانسته است عدم تعادل برچسب‌ها را تا حدی کاهش دهد. در اسنپ‌فود نیز ساختار نویزی داده‌ها توجیه‌کننده نزدیکی عملکرد مدل پیشنهادی به پارس‌برت پایه است. با وجود نتایج مطلوب، محدودیت‌هایی همچون قدیمی بودن داده‌ها و ارزیابی مدل تنها بر روی وظیفه تحلیل احساسات وجود دارد که می‌تواند در تحقیقات آتی برطرف شود.

۷- نتیجه‌گیری

نتایج تجربی نشان دادند که مدل پیشنهادی ضمن کاهش تعداد پارامترها و زمان آموزش، عملکردی رقابتی در تحلیل احساسات متون فارسی ارائه می‌دهد. اگرچه پارس‌برت اصلی در برخی موارد دقت بالاتری دارد، اما رویکرد ما با ایجاد تعادل میان دقت و کارایی محاسباتی، راهکاری عملی برای محیط‌های کم‌منبع محسوب می‌شود.

نوآوری این پژوهش در طراحی ساختاری و نحوه یکپارچه‌سازی آن‌ها است، نه در معرفی یک بلوک عصبی جدید. به‌طور خاص، در این کار عمق ParsBERT به‌صورت هدفمند کاهش داده شده و به‌جای استفاده از نمایش تک‌برداری[CLS]، کل توالی خروجی توکن‌ها به یک BiGRU تزریق شده است تا وابستگی‌های ترتیبی که در نسخه کم‌عمق ترنسفورمر تضعیف می‌شوند، بازیابی گردند. این بازطراحی مسیر جریان اطلاعات و جبران افت ظرفیت مدل کم‌عمق از طریق یک ماژول ترتیبی سبک، یک راهکار معماری جدید برای ایجاد توازن میان دقت و کارایی محاسباتی محسوب می‌شود و در کارهای پیشین برای تحلیل احساسات فارسی به این شکل بررسی نشده است. همچنین با کاهش لایه‌های ترنسفورمر و به‌کارگیری تکنیک‌های داده‌افزایی، محدودیت‌های مربوط به نامتوازن بودن داده‌ها تا حد زیادی برطرف شد.

بااین‌حال، پژوهش حاضر نیز محدودیت‌هایی دارد. نخست آن‌که ممکن است داده‌های مورداستفاده (مانند مجموعه داده طاقچه)، تمام ویژگی‌های زبان معاصر فارسی را پوشش ندهند. دوم، تمرکز مطالعه تنها بر روی تحلیل احساسات بوده و برای سایروظایف پردازش زبان طبیعی نیازمند بررسی‌های بیشتر است.

به‌عنوان مسیرهای آینده، می‌توان از این رویکرد برای سایر زبان‌های کم‌منبع یا وظایفی همچون خلاصه‌سازی و پرسش و پاسخ استفاده کرد. همچنین مقایسه دقیق‌تر ازنظر مصرف حافظه، سرعت آموزش و کارایی روی دستگاه‌های واقعی (مانند تلفن همراه) می‌تواند مسیر ارزشمندی برای تحقیقات بعدی باشد.

References

[1] H. P. Luhn, "The automatic creation of literature abstracts," IBM Journal of Research and Development, vol. 2, no. 2, pp. 159–165,Apr. 1958.

[2] Z. B. Nezhad and M. A. Deihimi, "A combined deep learning model for Persian sentiment analysis," IIUM Engineering Journal, vol. 20, no. 1, pp. 129–139, 2019.

[3] S. A. H. Minoofam, A. Bastanfard, and M. R. Keyvanpour, "TrCLA: A transfer learning approach to reduce negative transfer for cellular learning automata," IEEE Transactions on Neural Networks and Learning Systems, 2021.

[4] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, "Hybrid deep learning models for sentiment analysis," Complexity, vol. 2021.

[5] H. Sadr and M. Nazari Soleimandarabi, "ACNN-TL: Attention-based convolutional neural network coupling with transfer learning and contextualized word representation for enhancing the performance of sentiment classification," The Journal of Supercomputing, vol. 78, no. 7, pp. 10149–10175, 2022.

[6] Lakizadeh and Z. Zinaty, "A novel hierarchical attention-based method for aspect-level sentiment classification," Journal of AI and Data Mining, vol. 9, no. 1, pp. 87–97, 2021.

[7] S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pretrained word embeddings," Expert Systems with Applications, vol. 117, pp. 139–147, 2019.

[8] Q. Zhang, Z. Zhang, M. Yang, and L. Zhu, "Exploring coevolution of emotional contagion and behavior for microblog sentiment analysis: A deep learning architecture," Complexity, vol. 2021, 2021.

[9] M. Li, L. Chen, J. Zhao, and Q. Li, "Sentiment analysis of Chinese stock reviews based on BERT model," Applied Intelligence, vol. 51, no. 7, pp. 5016–5024, 2021.

[10] M. B. Dastgheib, S. Koleini, and F. Rasti, "The application of deep learning in Persian documents sentiment analysis," International Journal of Information Science and Management(IJISM),vol.18,no.1,pp.115,2020.

[11] Keramatfar, H. Amirkhani, and A. Jalaly Bidgoly, "Multi-thread hierarchical deep model for context-aware sentiment analysis," Journal of Information Science, p. 0165551521990617, 2021.

[12] M. Farahani, M. Gharachorloo, M. Farahani, and M. Manthouri, "ParsBERT: Transformer-based model for Persian language understanding," Neural Processing Letters, vol. 53, no. 6, pp. 3831–3847, 2021.

[13] S. S. Sadeghi, H. Khotanlou, and M. R. Mahand, "Automatic Persian text emotion detection using cognitive linguistic and deep learning," Journal of AI and Data Mining, vol. 9, no. 2, pp. 169–179, 2021.

[14] D. Xue et al., "Deep learning-based personality recognition from text posts of online social networks," Applied Intelligence, vol. 48, no. 11, pp. 4232–4246, 2018.

[15] D. Xu, Z. Tian, R. Lai, X. Kong, Z. Tan, and W. Shi, "Deep learning based emotion analysis of microblog texts," Information Fusion, vol. 64, pp. 1–11, 2020.

[16] V. Lorenzo, M. N. Franco, A. Pappalardo, and R. Prevete, "Dynamic hard pruning of neural networks at the edge of the internet," Journal of Network and Computer Applications, vol. 200, p. 103330, 2022.

[17] N. Park, X. Liu, G. Yang, and A. Simoulin, "Memory-efficient fine-tuning of transformers via token selection," in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, 2024.

[18] Y. Z. Vakili, A. Fallah, and S. Zakeri, "Enhancing sentiment analysis of Persian tweets: A transformer-based approach," in 2024 10th International Conference on Web Research(ICWR),Tehran,Iran,2024,pp.226230.

[19] L. Mathew and B. V. R., "Efficient classification techniques in sentiment analysis using transformers," in Advances in Electronics, Communication and Computing, Singapore: Springer, 2022, pp. 795–803.

[20] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

[21] Karimi, S., & Jafarinejad, F. (2024). Aspect-Based Sentiment Analysis using the Attentional Encoder Network. Signal and Data Processing, 20(4), 121-128.

[22] Mir, M., & Noferesti, S. (2024). Using Data Augmentation Techniques for Sentiment Analysis of Users' Opinions on Reopening of Schools During the Covid-19 Epidemic. Signal and Data Processing, 21(2), 3-14.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 4, No. 1, Spring & Summer 2026, pp. 129 - 145

A Hybrid and Low-Parameter Model Based on Transformer and GRU Architectures for Persian Sentiment Analysis

Mohsen Nooraee[1], Hamidreza Ghaffari[2*], Fatemeh Zarisfi Kermani[3]

[1] PhD student, Department of Computer Engineering, Islamic Azad University, Ferdows Branch, South Khorasan, Iran

[2] Associate Professor, Department of Computer Engineering, Islamic Azad University, Ferdows Branch, South Khorasan, Iran

[3] Assistant Professor, Department of Mathematics, Faculty of Science, University of Jiroft, Kerman, Iran

Article Information

Original Research Paper

Received:
2025 December 16

Accepted:
2026 March 17

Keywords:
Natural Language Processing, Sentiment Analysis, Computational Complexity Reduction, Transformer Model

Corresponding Author*:
hghaffari@ferdowsiau.ac.ir

Abstract

Sentiment analysis, as a subset of text classification, is a vital tool in areas such as information management, data analysis, and performance improvement. The complexity of Persian texts, due to features such as sarcasm, polysemy, and language combinations, poses unique challenges for sentiment recognition. In this study, a new model is introduced that extracts local and global features of the text by combining GRU and ParsBERT. The main focus of this study was to reduce the computational complexity and training time while maintaining the desired level of accuracy, so that the model can be used in resource-constrained environments. This model improves the parameters and training time by reducing the encoder layers and adding GRU. Experiments on the Snapfood and Taqcheh datasets showed that this model has higher accuracy than the baseline models and its training is also faster. The proposed model achieved F1 score of 76.76 in the niche dataset and 16.86 in the Snapfood dataset, while the number of parameters was significantly reduced. These results indicate that the proposed model is efficient and optimal for sentiment analysis of Persian texts.

: 10.22034/ABMIR.2026.24111.1200

E-ISSN: 2821-2037 /The Author 2026. Published by Yazd University This is an open access article under the CC BY 4.0 License (https://creativecommons.org/licenses/by/4.0/).