مقاله پژوهشی

رویکرد ارزیابی هیجان نوین جهت مراقبت از سرطان مبتنی بر مدل‌های زبانی بزرگ

نفیسه فارغ‌زاده، محمد قبادی، پریسا رحمانی و مهدی بازرگانی

چکیده: پژوهش حاضر، در راستایرفع دغدغه و مواجهه با چالش کمبود ابزارهای تخصصی بومی در حوزه تحلیل هیجانات بیماران سرطانی و با هدف ارائه یک سیستم نوین برای ارزیابی هیجانات این بیماران آغاز شده است. هدف اصلی این مطالعه، طراحی و پیاده‌سازی یک سیستم ترکیبی مبتنی بر یادگیری عمیق برای شناسایی حضور یک یا چند هیجان به صورت همزمان از میان شش دسته هیجانی (شادی، غم، خشم، ترس، امید، ناامیدی)، در متون فارسی شبکه‌های اجتماعی است. این رویکرد نوین امکان درک بهتر پیچیدگی و هم‌زمانی هیجانات را فراهم می‌نماید. به این منظور، مجموعه‌ای شامل ۱۰،۰۰۰ پست مرتبط با سرطان از پلتفرم‌های اجتماعی برچسب‌گذاری شد. مدل پیشنهادی که از تلفیق هوشمندانه مدل زبان بزرگ بومیParsBERT با شبکه Bi-GRU بهره می‌برد، برای این وظیفه تنظیم دقیق گردید. نتایج ارزیابی، دستاورد قابل توجه این پژوهش را به وضوح نشان می‌دهد؛ روش پیشنهادی به دقت ۸۶٫۸٪ و معیار F1-Score میانگین ۸۴٪ دست یافت. سیستم پیشنهادی، بهبودی معادل ۳٫۲٪ در دقت نسبت به مدل‌های پایه نشان داد. سیستم در تشخیص هیجانات پرتکرار مانند غم با معیار F1-Score معادل ۹۲٪ و ناامیدی با معیار F1-Score معادل ۹۰٪، عملکردی موثر و ابزاری خودکار و قدرتمند برای پایش هیجانات بیماران سرطانی و مداخلات روان درمانی موثرتر ارائه می‌دهد.

***کلیدواژه:* پردازش زبان طبیعی، تحلیل هیجانات، سرطان، مدل‌های زبان بزرگ، یادگیری عمیق.**

۱- مقدمه

اخیراً با گسترش فناوری‌های دیجیتال و رونق رسانه‌های اجتماعی، پیوسته حجم عظیمی از داده‌های متنی به‌صورت آنلاین تولید می‌شود که تحلیل آن‌ها می‌تواند دریچه‌ای نوین به بازنمایی تجارب و هیجانات انسانی باشد. تحلیل هیجان و احساسات[1] به‌عنوان یکی از شاخه‌های کلیدی حوزه پردازش زبان طبیعی[2] (NLP) در دو دهه اخیر رشد فزاینده‌ای داشته است. این حوزه علمی به دنبال شناسایی، طبقه‌بندی و درک احساسات بیان‌شده در متون انسانی است و در ابتدا تمرکز اصلی آن بر روی متون عمومی مانند نقد فیلم‌ها، بررسی محصولات و شبکه‌های اجتماعی بود [۱]. در دهه‌های اخیر، علاقه و توسعه فناوری‌های تحلیل احساسات رشد چشمگیری داشته است، به‌ویژه به دلیل اهمیت فزاینده رسانه‌های اجتماعی در بیان افکار و هیجانات. یافته‌های حاصل از تحلیل احساسات، می‌تواند در حوزه نوظهور روایت درمانی[3] که به داستان‌ها و روایت‌های بیماران در فرایند درمان می‌پردازد، سهم بسزایی ایفا کند. با بهره‌گیری از ابزارهای مختلف و متنوع تحلیل هیجانات، متخصصان سلامت قادر خواهند بود الگوهای پنهان روایت و هیجان بیماران را تشخیص داده و نگاه انسانی‌تری نسبت به بیماری و بیماران فراهم آورند. اما به‌تدریج با ظهور حوزه‌های جدیدی مانند سلامت دیجیتال، تحلیل هیجان به یکی از ابزارهای حیاتی برای درک وضعیت روانی بیماران تبدیل شد [۲].

تحقیقات متعدد نشان داده‌اند که هیجانات افراد، به‌ویژه بیماران خاص و مزمن، به‌طور مستقیم بر کیفیت زندگی و حتی اثربخشی درمان‌های پزشکی تأثیرگذار هستند [۳]. بیماران سرطانی به دلیل شرایط ویژه بیماری، فشارهای جسمی و روانی، و چالش‌های درمانی، بیش از سایر گروه‌ها در معرض نوسانات هیجانی قرار دارند [۴]. در نتیجه، تحلیل هیجان در این گروه از بیماران نه‌تنها جنبه پژوهشی دارد، بلکه کاربرد مستقیم بالینی و اجتماعی نیزیافته است.

حوزه سلامت به‌ویژه در حوزه بیماری‌های مزمن مانند سرطان، از منابع غنی داده‌های بدون ساختار متنی تشکیل شده است که از رسانه‌های اجتماعی، انجمن‌های آنلاین، بازخورد بیماران و پرونده‌های الکترونیکی سلامت تهیه می‌شوند. روش‌های سنتی و دستی برای تحلیل این حجم عظیم داده‌ها ناکارآمد و گاهی ناکافی بوده‌اند، از این رو تحلیل احساسات با ارائه راهکارهای خودکار و مقیاس‌پذیر به تحول در تحلیل تجربیات بیماران کمک کرده و امکان تصمیم‌گیری بهتری را برای ارائه‌دهندگان خدمات و سیاست‌گذاران فراهم می‌آورد [۵].

در زمینه تحلیل احساسات متون فارسی، چالش‌های قابل‌توجهی وجود دارد. زبان فارسی دارای ویژگی‌های صرفی و نحوی پیچیده‌ای است، از جمله استفادهٔ گسترده از افعال مرکب، کنایه‌ها و استعاره‌ها، که تحلیل خودکار احساسات را دشوارتر از زبان‌هایی همچون انگلیسی می‌کند. افزون بر این، کمبود داده‌های برچسب‌گذاری‌شده و ابزارهای تخصصیNLP برای فارسی، به ویژه در حوزه‌های حساس سلامت روان، توسعه سیستم‌های کارآمد تحلیل احساسات را با محدودیت‌های جدی مواجه کرده است.

این مقاله در تاریخ ۲۹ مهر ماه ۱۴۰۴ دریافت و در تاریخ ۲۸ بهمن ماه ۱۴۰۴ بازنگری شد.

نفیسه فارغ‌زاده (نویسنده مسئول)، دانشکده کامپیوتر، واحد خدابنده، دانشگاه آزاد اسلامی، زنجان، ایران، (email: fareghzadeh@iau.ac.ir).

محمد قبادی، دانشکده کامپیوتر، واحد الکترونیک، دانشگاه آزاد اسلامی، تهران، ایران، (email: mohammad.ghobadi@iau.ac.ir).

پریسا رحمانی، دانشکده کامپیوتر، واحدپردیس، دانشگاه آزاداسلامی، تهران، ایران، (email: parisa.rahmani@iau.ac.ir).

مهدی بازرگانی، دانشکده کامپیوتر، واحد خدابنده، دانشگاه آزاد اسلامی، زنجان، ایران، (email: mbzirn@iau.ac.ir).

1. Sentiment and Emotion Analysis
2. Natural Language Processing
3. Narrative Medicine

ظهور مدل‌های زبانی بزرگ[1] (LLMs) مانند BERT[2] و GPT[3] تحول بزرگی در زمینه پردازش زبان طبیعی ایجاد کرده است. این مدل‌ها قادرند با آموزش بر روی حجم بالایی از داده‌های متنی، معانی عمیق و روابط پیچیده بین کلمات و جملات را درک کنند. در حوزه زبان فارسی نیز توسعه مدل‌هایی مانند ParsBERT، برپایه معماری ترنسفورمر و آموزش بر روی مجموعه‌های گسترده متنی، توانمندی چشمگیری در درک زبان فارسی و تحلیل‌های زبانی دارد. اما تحلیل دقیق هیجانات، به ویژه در متن‌های پیچیده بیماران، نیازمند درک ترتیبی و وابستگی‌های زمانی نیز هست؛ لذا ترکیب مدل‌های ترنسفورمری با شبکه‌های عصبی بازگشتی مانند Bi-GRU[4]، چارچوب مؤثری برای شناسایی هیجانات به شمار می‌رود [۶].

تحقیقات پیشین در این حوزه، عمدتاً بر حوزه داده‌های انگلیسی زبان متمرکز بوده‌اند. با این حال، هنوز پژوهشی که طیف گسترده‌ای از هیجانات مانند شادی، غم، خشم، ترس، امید و ناامیدی را در داده‌های غیررسمی بومی‌سازی‌شده فارسی‌زبان تحلیل کند، وجود ندارد.پژوهش حاضر در ادامه همین خلأ، با طراحی و پیاده‌سازی مدلی ترکیبی، درصدد شناسایی شش دسته هیجانی کلیدی شامل احساسات شادی، غم، خشم، ترس، امید، ناامیدی، در متون شبکه‌های اجتماعی بیماران سرطانی فارسی‌زبان است. دسته‌بندی دقیق این هیجانات می‌تواند مبنایی برای مداخلات روان‌شناختی هدفمند و تصمیم‌گیری‌های بهتر در حوزه سلامت به شمار رود و گامی ارزنده در توسعه ابزارهایNLP برای زبان فارسی باشد. لذا مهم ترین دستاوردهای پژوهش حاضر به شرح زیر می‌باشد:

- طراحی و پیاده‌سازی سیستم ترکیبی و بومی‌سازی‌شده ارزیابی هیجانات به‌منظور تحلیل احساسات چندگانه، در متون شبکه‌های اجتماعی بیماران سرطانی فارسی‌زبان.
- ترکیب رویکردهای ParsBERT و Bi-GRU برای استخراج نمایش معنایی عمیق و وابستگی‌های ترتیبی، با بهبود معنادار نسبت به مدل‌های پایه.
- ارائه چارچوب تحلیل چندبرچسبی هیجانات شامل تحلیل شش دسته هیجانی کلیدی در متون فارسی بیماران سرطانی، که بازنمایی واقع‌بینانه‌تری از هم‌زمانی احساسات فراهم می‌نماید.
- گزارش شاخص‌های کمّی ارزیابی (Accuracy, Macro-F۱) همراه با تحلیل ماتریس درهم‌ریختگی و پایایی برچسب‌گذاری.
- قابلیت استفاده سیستم پیشنهادی در پایش روانی بیماران و پشتیبانی از مداخلات هدفمند در حوزه پزشکی و سلامت در بستر شبکه‌های اجتماعی فارسی.

در ادامه مقاله، ابتدا در بخش دوم تاریخچه تحقیق و مطالعات مرتبط کلیدی را بررسی و با دیدگاه پیشنهادی این پژوهش مقایسه می‌شود. سپس در بخش سوم ابعاد سیستم پیشنهادی و فازهای اجرایی آن را تشریح می‌گردد. در بخش چهارم پیاده‌سازی و ارزیابی سیستم پیشنهادی و بحث و مقایسه با رویکردهای مرتبط انجام شده است. در نهایت نیز در بخش پنجم نتیجه‌گیری و راهکارهای آتی پژوهش بیان شده است.

۲- تحقیقات مرتبط

۲-۱ تاریخچه و رویکردهای کلاسیک

تحلیل احساسات به‌عنوان شاخه‌ای مهم از پردازش زبان طبیعی، بر شناسایی، استخراج و طبقه‌بندی خودکار احساسات، نظرات و نگرش‌های بیان شده در متون تمرکز دارد. بیماران سرطانی به دلیل شرایط ویژه بیماری، فشارهای جسمی و روانی، و چالش‌های درمانی، بیش از سایر گروه‌ها در معرض نوسانات هیجانی قرار دارند. در نتیجه، تحلیل هیجان در این گروه از بیماران نه‌تنها جنبه پژوهشی دارد، بلکه کاربرد مستقیم بالینی و اجتماعی نیزیافته است.برای رسیدن به این هدف، رویکردهای مختلفی در طول زمان پیشنهاد شده‌اند. از روش‌های کلاسیک و سنتی اولیه مبتنی بر لغت‌نامه‌های احساسی و قواعد دستی گرفته تا مدل‌های یادگیری ماشین و اخیراً شبکه‌های عصبی عمیق و مدل‌های زبان بزرگ، هرکدام نقش مهمی در پیشرفت این حوزه ایفا کرده‌اند. در ادامه، ابتدا به روش‌های کلاسیک و سپس به الگوریتم‌های یادگیری ماشین سنتی پرداخته می‌شود.

نخستین نسل از سیستم‌های تحلیل هیجان عمدتاً مبتنی بر فرهنگ‌لغت‌ها و واژگان احساسی بودند. در این رویکرد، مجموعه‌ای از واژه‌ها همراه با برچسب‌های هیجانی (مثبت، منفی، خنثی یا هیجان‌های خاص) تهیه می‌شد و سپس متون با مقایسه واژه‌هایشان با این لغت‌نامه تحلیل می‌گردیدند [۵]. در حوزه سلامت، استفاده از این رویکرد بسیار رایج بود، چرا که در مراحل اولیه پایش داده‌های بالینی محدود و ساده بودند. اما با گذشت زمان و افزایش پیچیدگی داده‌های متنی بیماران، محدودیت‌های روش‌های لغت‌نامه‌ای آشکار شد [۶].

رویکرد مبتنی بر قاعده نیزیک روش ساختاریافته برای تشخیص هیجانات در متن است که از قواعد زبانی و معنایی برای شناسایی بار احساسی جملات استفاده می‌کند. این فرآیند با پیش‌پردازش متن از جمله توکن‌سازی، حذف کلمات توقف و برچسب‌زنی دستوری آغاز می‌شود [۷]. سپس، قواعدی بر پایه واژگان احساسی و ساختارهای زبانی (مانند نفی، تشدید یا عطف) تعریف می‌شوند تا بار احساسی کلمات و عبارات با توجه به زمینه تعیین شود. این روش اغلب از منابع واژگانی احساسی مانند WordNet-Affect [۵] یا SentiWordNet [۶] بهره می‌برد و می‌تواند با افزودن سرنخ‌هایی مانند شکلک‌ها یا اصطلاحات تخصصی، دقت خود را در متون غیررسمی افزایش دهد. مزیت اصلی این روش، شفافیت و قابلیت تفسیر بالا است، اما در تشخیص کنایه یا احساسات پیچیده ضعیف عمل می‌کند و به‌طور گسترده‌ای به دقت قواعد و جامعیت واژگان وابسته است .به‌عنوان مثال، مطالعه سیل و همکاران [۷] یک سیستم تشخیص و تحلیل احساس در سطح جمله ارائه داده‌اند که از قواعد معنایی و تحلیل وابستگی برای شناسایی احساسات پایه (مانند خشم، شادی، غم) استفاده می‌کند و نتایج قابل قبولی در داده‌های انگلیسی به دست آورده است.

با رشد داده‌های متنی و افزایش توان پردازشی، پژوهشگران به سمت الگوریتم‌هایادگیری ماشین حرکت کردند. در این رویکرد، به‌جای اتکا به قواعد دستی، سیستم‌ها از روی داده‌های برچسب‌گذاری‌شده آموزش می‌دیدند تا بتوانند الگوهای پنهان را کشف کنند.

این تغییر رویکردیک جهش اساسی محسوب می‌شد؛ چرا که امکان پردازش داده‌های بزرگ‌تر و پیچیده‌تر را فراهم آورد. از آن زمان، الگوریتم‌های مختلفی در این حوزه به کار گرفته شدند که مهم‌ترین آن‌ها

1. Large Language Models
2. Bidirectional Encoder Representations from Transformers
3. Generative Pre-Trained Transformer
4. Bi-Directional Gated Recurrent Unit

الگوریتم ماشین بردار پشتیبان[1] (SVM) [۸]، الگوریتم بیز ساده[2] [۹]، الگوریتم جنگل تصادفی [۱۰] تا [۱۳] و الگوریتم نزدیکترین همسایه[3] (KNN) [۱۴] تا [۱۶] هستند. به‌عنوان نمونه، مطالعه ویکارسا و طاهر در پژوهش [۹] از بیز ساده برای طبقه‌بندی هیجانات کاربران توییتر استفاده کرده و دقت قابل‌قبولی در شناسایی احساسات پایه (مانند شادی، خشم و ناراحتی) گزارش کرده‌اند. این نشان می‌دهد که علیرغم سادگی، بیز ساده می‌تواند در محیط‌های پویای آنلاین ابزاری کارآمد برای تحلیل سریع احساسات باشد. هریک از روش های ذکر شده دارای نقاط قوت و ضعفی هستند که در ادامه بررسی می‌گردد. همچنین، مطالعه دیگر [۱۴] بر روی تالارهای گفتگوی آنلاین بیماران مبتلا به سرطان تخمدان نشان داده است که با تحلیل متون می‌توان پر تکرارترین علائم و نگرانی‌های بیماران مانند درد، تهوع و اضطراب را شناسایی و دسته‌بندی کرد.

۲-۲ نقاط قوت و ضعف روش‌های کلاسیک

الگوریتم‌های یادگیری ماشین کلاسیک و سنتی در زمان خود یک پیشرفت بزرگ نسبت به رویکردهای کلاسیکِ مبتنی بر قواعد به شمار می‌آمدند، زیرا قادر بودند الگوهای آماری را مستقیماً از داده‌ها و بدون نیاز به تعریف قوانین دستی استخراج کنند. با این حال، این الگوریتم‌ها با محدودیت‌های ذاتی مواجه بودند که کارایی آن‌ها را، به‌ویژه در تحلیل داده‌های متنی پیچیده، به چالش می‌کشید.

مهم‌ترین این محدودیت‌ها عبارت بودند از:

- نیاز به مهندسی ویژگی: کارایی این مدل‌ها به شدت به فرآیند زمان‌بر و تخصصیِ استخراج و انتخاب دستی ویژگی‌های مناسب از متن وابسته بود.
- ضعف در درک روابط پیچیده: این الگوریتم‌ها در درک معنای عمیق، مفاهیم انتزاعی و روابط طولانی‌مدت میان کلمات در یک متن طولانی، توانایی محدودی داشتند.
- حساسیت به داده‌ها: عملکرد آن‌ها به کیفیت داده‌های ورودی بسیار حساس بود و نویزیا داده‌های نامتعادل می‌توانست به سادگی نتایج را تحت تأثیر قرار دهد.

این چالش‌ها و مشکلات مسیر را برای ظهور مدل‌های یادگیری عمیق[4] (DL) هموار ساخت. این نسل نوین از مدل‌ها، با بهره‌گیری از معماری‌های شبکه‌ای چندلایه، قادر به یادگیری خودکار بازنمایی‌ها و ویژگی‌های سلسله‌مراتبی مستقیماً از داده‌های خام هستند و نیاز به مهندسی ویژگی دستی را به حداقل می‌رسانند [۱۷].

در نتیجه، می‌توان گفت هرچند الگوریتم‌های سنتی نقطه شروع ارزشمندی در تحلیل احساسات بیماران بوده‌اند، اما محدودیت‌های بنیادی آن‌ها در درک ظرایف زبانی و هیجانات چندلایه، به‌ویژه در زمینه حساسی مانند تجربیات بیماران سرطانی، آشکار بود. این ناتوانی، ضرورت حرکت جامعه پژوهشی به سمت پارادایم‌های پیشرفته‌تریادگیری عمیق و مدل‌های زبان بزرگ را برای دستیابی به تحلیل‌های دقیق‌تر و عمیق‌تر ایجاب می‌کرد.

۲-۳ روش‌های نوین مبتنی بر یادگیری عمیق

ظهور یادگیری عمیق در دهه ۲۰۱۰، یک تغییر پارادایم اساسی در حوزه پردازش زبان طبیعی ایجاد کرد. برخلاف مدل‌هاییادگیری ماشین سنتی که به فرآیند پرهزینه و تخصصی مهندسی ویژگی دستی متکی بودند، شبکه‌های عصبی عمیق قادرند بازنمایی‌های سلسله‌مراتبی و ویژگی‌های پیچیده را به‌طور خودکار و مستقیم از داده‌های خام استخراج کنند [۱۷]. این قابلیت، به‌ویژه برای تحلیل هیجان که پدیده‌ای چندوجهی، وابسته به زمینه و ذاتاً مبهم است، اهمیتی بنیادین دارد.

شبکه‌های عصبی کانولوشنال[5] (CNN) به دلیل توانایی در شناسایی الگوهای محلی در داده‌های متنی، به‌سرعت در NLP نیز به کار گرفته شدند. CNN با استفاده از فیلترهای کانولوشن، می‌تواند n-gramهای کلیدی و ویژگی‌های معنایی را استخراج کند، که این ویژگی آن را برای وظایفی مانند طبقه‌بندی قطبیت احساسات در متون کوتاه بسیار کارآمد می‌سازد [۱۸] و [۱۹]. شبکه‌های عصبی بازگشتی نیز (RNN) برای این منظور طراحی شدند، اما با مشکل محو شدن گرادیان مواجه بودند. برای حل این مشکل، معماری‌های پیشرفته‌تری مانند حافظه طولانی کوتاه‌مدت[6] (LSTM) معرفی شدند [۲۰]. همچنین نسخه ساده‌تر و بهینه‌تر آن، یعنی واحد بازگشتی دروازه‌دار (GRU)، با پارامترهای کمتر به عملکردی مشابه دست یافت و به دلیل سرعت بالاتر، به گزینه‌ای محبوب در این حوزه تبدیل شد. برای بهره‌گیری همزمان از مزایای هر دو رویکرد، معماری‌های ترکیبی توسعه یافتند. در این مدل‌ها، معمولاً یک لایه CNN برای استخراج ویژگی‌های محلی از متن به کار می‌رود و خروجی آن به یک لایهRNN (مانند LSTM یا GRU) سپرده می‌شود تا وابستگی‌های ترتیبی و طولانی‌مدتمیان این ویژگی‌ها را مدل‌سازی کند. این رویکرد ترکیبی در تحلیل داده‌های پیچیده‌تر مانند یادداشت‌های روزانه بیماران، عملکرد برتری از خود نشان داده است [۲۱].

پتانسیل بالای مدل‌هاییادگیری عمیق در درک ظرایف زبانی، منجر به کاربرد گسترده آن‌ها در تحلیل داده‌های متنی بیماران سرطانی شده است. برای نمونه، کرانل و همکاران در ۲۰۱۶ با تحلیل توییت‌های مرتبط با سرطان، توانستند با استفاده از این مدل‌ها الگوهای هیجانی بیماران را در مراحل مختلف درمان شناسایی کنند [۲۲]. پژوهش‌های جدیدتر نیز بر تشخیص هیجانات دقیق‌تر و چندبعدی متمرکز شده‌اند. در همین راستا، سوسی و کاراگی با معرفی مجموعه داده تخصصی CancerEmo، نشان دادند که مدل‌های مبتنی بر رویکردیادگیری عمیق قادر به تشخیص طیف وسیع‌تری از هیجانات مانند خشم، ترس، و امید در متون تولیدشده توسط بیماران هستند و افق‌های جدیدی را برای درک عمیق‌تر تجربیات آن‌ها گشوده‌اند [۲۳].

با پیشرفت معماری ترنسفورمر، نسل جدیدی از مدل‌ها با عنوان مدل‌های زبان بزرگ ظهور کردند که نمایانگر پیشرفته‌ترین دستاوردها در حوزه پردازش زبان طبیعی هستند. این مدل‌ها با آموزش بر روی حجم بی‌سابقه‌ای از داده‌های متنی، توانایی درک و تولید زبان طبیعی را به سطحی نزدیک به انسان رسانده‌اند.نقطه عطف حقیقی در پردازش زبان طبیعی، معرفی معماری ترنسفورمر توسط واسوانی و همکاران بود [۲۴]. این معماری با کنار گذاشتن ساختارهای بازگشتی و معرفی سازوکار «توجه»، انقلابی در مدل‌سازی زبان ایجاد کرد. مکانیزم توجه به مدل اجازه می‌دهد تا اهمیت هر کلمه را در ارتباط با تمام کلمات دیگر در یک متن، بدون توجه به فاصله آن‌ها، ارزیابی کند. این قابلیت، محدودیت اصلی شبکه‌های RNN در به خاطر سپردن وابستگی‌های طولانی‌مدت را

1. Support Vector Machine
2. Naïve Bayes
3. K-Nearest Neighbors
4. Deep Learning
5. Convolutional Neural Network yes
6. Long Short-Term Memory

برطرف نمود و به سنگ بنای توسعه مدل‌های زبان بزرگ تبدیل شد.

در حوزه پژوهش و بر پایه معماری ترنسفورمر، دو خانواده اصلی از LLMها ظهور کردند. مدل BERT با استفاده از یک رویکرد مبتنی بر رمزگذار و مکانیزمیادگیری دوسویه، به درک عمیقی از زمینه در هر دو جهت چپ و راست یک کلمه دست یافت [۲۵] و [۲۶]. در مقابل، خانواده GPT با معماری مبتنی بر رمزگشا و رویکرد خودبازگشتی، توانایی بی‌نظیری در تولید متن‌های منسجم و طبیعی از خود نشان دادند و استانداردهای جدیدی را در وظایف تحلیل زبان تعریف کردند [۲۷].

همچنین کاربرد LLMها در تحلیل هیجان بیماران سرطانی نیز در پژوهش‌های اخیر مورد توجه قرار گرفته است. برای مثال، در مطالعه‌ای برجسته، ونگ و همکاران مدلی به نام CALLM[1] را معرفی کردند که به‌طور خاص برای تحلیل محتوای متنی بیماران سرطانی طراحی شده است. این مدل با تحلیل یادداشت‌های روزانه بیماران، توانست هیجانات پیچیده‌ای مانند غم، ترس، و امید را با دقتی قابل‌توجه شناسایی کند. با این حال، دقت گزارش‌شده (۷۳٪) نشان می‌دهد که این حوزه همچنان یک زمینه تحقیقاتی فعال و نیازمند بهبود است [۲۸].

علی‌رغم توانمندی‌های بی‌سابقه، LLMها در کاربردهای بالینی با محدودیت‌های مهمی مانند نیاز به داده‌های برچسب‌دار انبوه، هزینه محاسباتی و انرژی بسیار بالا، عملکرد ضعیف‌تر در زبان‌های کم‌منبع و چالش‌های اخلاقی و حریم خصوصی روبرو هستند [۲۹].

درحوزه پژوهش نیز پژوهش‌های بین‌المللی در زمینه تحلیل هیجان بیماران سرطانی، گستره وسیعی از منابع داده و رویکردهای تحلیلی را در بر می‌گیرد که می‌توان آن‌ها را در چند حوزه اصلی دسته‌بندی کرد.یکی از جریان‌های اصلی پژوهش، بر تحلیل داده‌های تولیدشده توسط بیماران در شبکه‌های اجتماعی، به‌ویژه توییتر، متمرکز بوده است. این پلتفرم‌ها به دلیل دسترسی آسان و حجم بالای داده‌های عمومی، بستری مناسب برای مطالعات در مقیاس بزرگ فراهم می‌کنند. در مقابل داده‌های عمومی شبکه‌های اجتماعی، گروه دیگری از مطالعات بر داده‌های بالینی موجود در پرونده‌های الکترونیک سلامت و یادداشت‌های پزشکان متمرکز شده‌اند [۳۰]. این داده‌ها، اگرچه دسترسی به آن‌ها دشوارتر است، اما از اعتبار بالینی بیشتری برخوردارند. پژوهش‌ها در این حوزه نشان داده‌اند که تحلیل هیجانات ابرازشده در یادداشت‌های بیماران مبتلا به سرطان ریه، می‌تواندیک عامل پیش‌بینی‌کننده برای کیفیت زندگیو پاسخ آن‌ها به درمان باشد [۳۱]. همچنین یک روند نوظهور در این حوزه، حرکت به سمت تحلیل‌های چندوجهی است. با این حال، یک محدودیت فراگیر در ادبیات موجود، تمرکز غالب بر روی داده‌های انگلیسی‌زبان است [۳۲]. این امر سبب شده تا زبان‌های کم‌منبع مانند فارسی، با وجود جمعیت قابل‌توجه بیماران، در حاشیه باقی بمانند. در نتیجه، نیاز به پژوهش‌های بومی که بتوانند از این ظرفیت‌های فناورانه برای تحلیل تجربیات بیماران فارسی‌زبان بهره‌برداری کنند، بیش از پیش احساس می‌شود.زبان فارسی در حوزه پردازش زبان طبیعی در دسته زبان‌های کم‌منبع طبقه‌بندی می‌شود. این امر بدان معناست که در مقایسه با زبان‌های پرداده مانند انگلیسی، دسترسی به پیکره‌های متنی برچسب‌دار در مقیاس بزرگ و ابزارهای پردازشی پیشرفته برای فارسی محدود است [۳۳]. علاوه بر کمبود منابع، چالش‌های ساختاری زبان فارسی، از جمله پیچیدگی‌های صرفی، فراوانی افعال مرکب، و شکاف قابل‌توجه میان زبان نوشتاری معیار و گونه‌های گفتاری غیررسمی، فرآیندهای تحلیل خودکار را پیچیده‌تر می‌سازد.

مطالعات اولیه در زمینه تحلیل هیجان فارسی، مسیر را برای پژوهش‌های آتی هموار کردند. از جمله تلاش‌های پیشگام، پژوهش روشنفکر و همکاران بود که با به‌کارگیری مدل‌های یادگیری عمیق بر روی داده‌های عمومی، برتری این رویکردها را نسبت به روش‌های سنتی در زبان فارسی نشان داد [۳۴]. در پاسخ به کمبود داده، پژوهشگران اقدام به توسعه مجموعه داده‌های عمومی برای تحلیل احساسات کردند که از آن جمله می‌توان به ArmanEmo و Persian Sentiment Treebank اشاره نمود. با این حال، علی‌رغم این پیشرفت‌ها، فقدان داده‌های تخصصی در حوزه سلامت همچنان به عنوان یک شکاف تحقیقاتی باقی مانده است [۳۵]. در زمینه تحلیل هیجان بیماران در ایران، ادبیات پژوهشی بسیار محدودی وجود دارد. یکی از مطالعات مرتبط، پژوهش یزدانی و همکاران است که با تحلیل داده‌های جمع‌آوری‌شده از بیماران بستری، احساسات آن‌ها را با دقت بالایی (حدود ۹۰٪) به سه قطبیت مثبت، منفی و خنثی طبقه‌بندی کرد [۳۶]. با وجود نتایج امیدوارکننده، این پژوهش با چند محدودیت کلیدی مانند محدودیت در دانه‌بندی هیجانات، دامنه کاربرد محدود و عدم پوشش بیماران خاص مواجه بود. سایر پژوهش‌ها نیز عمدتاً بر روی تحلیل داده‌های عمومی کاربران ایرانی در شبکه‌های اجتماعی متمرکز بوده‌اند و به طور خاص به جمعیت بیماران نپرداخته‌اند [۳۷] تا [۳۹].

۲-٤ شکاف‌های تحقیقاتی و محدودیت‌های کلیدی در ادبیات پژوهش

با وجود پیشرفت‌های حاصل شده، مرور انتقادی ادبیات پژوهش نشان‌دهنده چندین چالش بنیادین و شکاف تحقیقاتی است که مسیر پژوهش‌های آینده، به‌ویژه برای زبان فارسی را مشخص می‌کند.بر اساس مرور جامع ادبیات پژوهش، می‌توان شکاف‌های تحقیقاتی کلیدی را که انگیزه اصلی این پژوهش را تشکیل می‌دهند، به شرح زیر جمع‌بندی نمود:

- شکاف منابع داده و تمرکز بر زبان انگلیسی: مهم‌ترین و بنیادین‌ترین شکاف، عدم توازن شدید منابع میان زبان انگلیسی و زبان‌های کم‌منبع مانند فارسی است. در حالی که پژوهش‌های پیشرفته بر داده‌های غنی انگلیسی متمرکز هستند، زبان فارسی در حاشیه قرار گرفته است [۴۰]. این به فقدان کامل یک پیکره تخصصی، عمومی و برچسب‌گذاری‌شده برای تحلیل هیجانات بیماران سرطانی فارسی‌زبان منجر می‌گردد.
- محدودیت در عمق تحلیل هیجانات: ادبیات موجود، چه در سطح بین‌المللی و چه در مطالعات محدود فارسی، عمدتاً بر تحلیل قطبیت احساسات در سطح درشت‌دانه (مثبت، منفی، خنثی) متمرکز است. این رویکرد سطحی، قادر به ثبت و تحلیل طیف پیچیده و چندبعدی هیجانات بالینی مهمی مانند امید، ترس، خشم و ناامیدی نیست که نقشی حیاتی در تجربه بیماران سرطانی ایفا می‌کنند [۴۱].
- بهره‌برداری ناکافی از مدل‌های زبان بزرگ پیشرفته: با وجود توسعه و در دسترس بودن مدل‌های زبانی بزرگ و قدرتمند برای زبان فارسی، مانند ParsBERT، کاربرد این مدل‌های پیشرفته در حوزه حساس تحلیل هیجان بیماران، بسیار محدود بوده و پتانسیل کامل آن‌ها در این زمینه کشف نشده است.
- فقدان نوآوری در معماری مدل‌ها: اکثر مطالعات انجام‌شده به استفاده از یک معماری واحد (مانند CNN یا LSTM به‌تنهایی) اکتفا کرده‌اند. معماری‌های ترکیبی که بتوانند از مزایای مکمل مدل‌های

1. Context-Aware Language Model for Medicine

جدول۱: مقایسه سیستماتیک دیدگاه پیشنهادی پژوهش حاضر با تحقیقات مرتبط کلیدی.

پژوهش	نوع مدل	دامنه کاربرد	زبان داده	نوع مسئله	نقاط قوت	محدودیت‌ها نسبت به روش پیشنهادی
ونگ و همکاران [۲۸]	LLM تخصصی پزشکی	تحلیل محتوای متنی بیماران سرطانی	انگلیسی	تک برچسبی احساس/هیجان	مدل تخصصی دامنه سرطان، داده بالینی عمیق	عدم بررسی زبان‌های کم‌منبع؛ عدم چندبرچسبی
گو و همکاران [۲۹]	BioBERT / ClinicalBERT	تحلیل احساسات و استخراج موجودیت در متون بالینی	انگلیسی	عمدتاً تک‌برچسبی	بهره‌گیری از پیش‌آموزش پزشکی، دقت بالا در متون بالینی	عدم تمرکز بر هیجانات چندبرچسبی بیماران سرطانی
لی و همکاران [۳۱]	Transformer چندوظیفه‌ای	تحلیل احساسات در سلامت دیجیتال	انگلیسی	عمدتاً تک‌برچسبی	چارچوب چندوظیفه‌ای، مقیاس‌پذیر	عدم تمرکز دامنه‌ای بر سرطان
روشنفکر و همکاران [۳۴]	DL	تحلیل احساس فارسی	فارسی	داده عمومی	نشان دادن برتری DL در فارسی	عدم تمرکز در حوزه پزشکی و سلامت
یزدانی و همکاران [۱۰]	ML/DL	بیماران بستری	فارسی	داده بیمارستانی	دقت بالا	عدم تحلیل هیجانات ریزدانه
دیدگاه پیشنهادی پژوهش حاضر (ParsBERT + Bi-GRU,) (۲۰۲۵)	LLMبومی RNN + ترکیبی	تحلیل هیجانات بیماران سرطانی در شبکه‌های اجتماعی	فارسی	چندبرچسبی و تحلیل هیجانات چندبعدی	بومی‌سازی برای زبان فارسی، مناسب زبان کم‌منبع، دقت بالا، تحلیل هم‌زمان چند هیجان، کاربرد مستقیم در حوزه پزشکی و سلامت دیجیتال	

مختلف به‌طور همزمان بهره‌مند شوند، یک حوزه کمتر بررسی‌شده و پرظرفیت برای نوآوری است.

به‌عنوان جمع‌بندی، جدول ۱ دیدگاه پیشنهادی پژوهش حاضر را با تحقیقات کلیدی مرتبط که از منظر نوع مسئله، دامنه کاربرد، زبان تحلیل، نوع داده و مدل قابلیت تحلیل سیستماتیک دارند مقایسه می‌نماید.

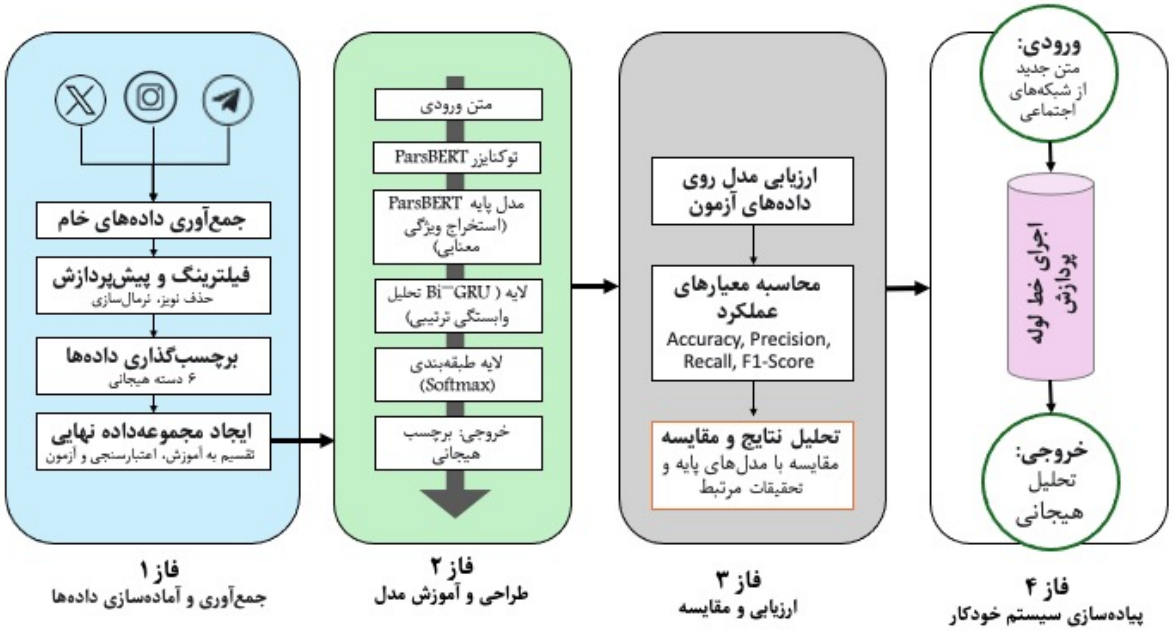
بر اساس بررسی و مقایسه سیستماتیک تحقیقات مرتبط کلیدی، پژوهش حاضر در پاسخ مستقیم به چالش‌های موجود و شکاف‌های شناسایی‌شده در ادبیات، خود را به عنوان یک پژوهش بنیادین و کاربردی در حوزه تحلیل هیجان بیماران سرطانی فارسی‌زبان معرفی می‌کند.برخلاف جریان غالب پژوهش‌های بین‌المللی، این تحقیق به طور انحصاری بر روی جمعیت بیماران سرطانی فارسی‌زبان متمرکز شده و تلاش می‌کند تا شکاف عمیق منابع و پژوهش را برای این زبان کم‌داده پر کند.ابزار توسعه‌یافته می‌تواند به عنوان یک سیستم پایش غیر تهاجمی در اختیار متخصصان سلامت روان، انکولوژیست‌ها و سیاست‌گذاران قرار گیرد تا با درک بهتر هیجانات بیماران، مداخلات روان‌شناختی پیشگیرانه و به‌موقع را طراحی و اجرا نمایند.

در بخش بعدی ابعاد کارکردی سیستم پیشنهادی، معماری، ماژول‌های عملیاتی و فازهای اجرائی مربوطه را به‌صورت مبسوط بیان می‌گردد.

۳- سیستم پیشنهادی

چارچوب عملیاتی سیستم پیشنهادی در چهار فاز اصلی اجرا گردید: ۱) جمع‌آوری و آماده‌سازی داده‌ها، ۲) طراحی و آموزش مدل، ۳) ارزیابی و ۴) پیاده‌سازی سیستم. همچنین، معماری سیستم پیشنهادی به‌صورت یک چارچوب چندلایه و مرحله‌ای طراحی شده است که جریان پردازش از داده خام تا استقرار سامانه را به‌صورت نظام‌مند مدیریت می‌کند. مدل سیستم پیشنهادی این پژوهش براساس یک معماری ترکیبی است که با هدف بهره‌گیری همزمان از قدرت نمایش‌های معنایی عمیق و مدل‌سازی وابستگی‌های ترتیبی طراحی شده است. در این چارچوب، ابتدا متن‌های ورودی به مدل زبانی پیش‌آموزش‌دیده ParsBERT داده می‌شوند تا بازنمایی‌های زمینه‌محور استخراج گردد؛ این بازنمایی‌ها با استفاده از سازوکار Self-Attention قادرند وابستگی‌های بلندبرد و روابط معنایی پیچیده میان واژگان را در سطح جمله مدل کنند و دانش زبانی عمومی و تخصصی زبان فارسی را به فضای برداری منتقل نمایند. سپس خروجی برداری ParsBERT به یک لایه Bi-GRU منتقل می‌شود تا وابستگی‌های ترتیبی و تغییرات تدریجی هیجانی در توالی متن به‌صورت دوطرفه مدل‌سازی گردد؛ این مرحله به‌ویژه در متون روایی بیماران که شدت یا نوع هیجان در طول جمله تغییر می‌کند، نقش تکمیلی مهمی ایفا می‌کند. بدین ترتیب، ParsBERT مسئول استخراج دانش معنایی عمیق و بافت‌محور است، در حالی‌که Bi-GRU ساختار توالی‌محور و پویایی هیجانات را پالایش و تقویت می‌کند. در نهایت، بردارهای غنی‌شده به لایه خروجی چندبرچسبی با تابع فعال‌سازی سیگموید متصل می‌شوند تا امکان پیش‌بینی هم‌زمان چند هیجان فراهم گردد. این ترکیب هدفمند موجب افزایش تعمیم‌پذیری مدل، بهبود تفکیک هیجانات هم‌پوشان و ارتقای شاخص‌های ارزیابی کمی نسبت به استفاده منفرد از هر یک از رویکردها شده است.معماری سیستم پیشنهادی در شکل ۱ نمایش داده ناهمگنی نوشتاری زبان فارسی ایفا می‌کند. در لایه دوم، لایه نمایش و بازنمایی معنایی متن قرار دارد که در آن داده‌های متنی به بردارهای عددی با استفاده از مدل زبانی پیش‌آموزش‌دیده) مبتنی بر معماری ترانسفورمر و سازگار با زبان فارسی (تبدیل می‌شوند؛ در این مرحله وابستگی‌های بافتی و روابط معنایی عمیق بین واژگان استخراج شده و ویژگی‌های زمینه‌محور تولید می‌گردد. لایه سوم، هسته مدل یادگیری و طبقه‌بندی چندبرچسبی هیجانات است که شامل لایه‌های تنظیم دوباره[1] و لایه خروجی با تابع فعال‌سازی سیگموید برای پیش‌بینی هم‌زمان چند هیجان می‌باشد؛ در این بخش، بهینه‌سازی مدل از طریق تابع زیان مناسب چندبرچسبی و تنظیم ابرپارامترها انجام شده است. در ادامه، لایه ارزیابی و تحلیل عملکرد قرار دارد که معیارهایی نظیر صحّت[2]، دقّت[3]، بازخوانی[4] و Macro-F۱ و تحلیل ماتریس درهم‌ریختگی را برای سنجش

1. Fine-Tuning
2. Accuracy
3. Precision
4. Recall



شکل ۱: چارچوب عملیاتی سیستم پیشنهادی.

کمی عملکرد مدل محاسبه می‌کند و الگوهای خطا را استخراج می‌نماید. نهایتاً، در لایه چهارم، لایه پیاده‌سازی و کاربرد عملی سیستم طراحی شده است که مدل آموزش‌دیده را در قالب یک سامانه قابل استفاده برای تحلیل هیجانات متنی در سناریوهای واقعی (مانند پایش روان‌شناختی دیجیتال) مستقر می‌کند. این ساختار لایه‌ای باعث تفکیک مسئولیت‌ها، افزایش قابلیت توسعه‌پذیری و امکان تحلیل دقیق هر مرحله از فرایند پردازش شده و انسجام مفهومی بین مراحل جمع‌آوری داده تا بهره‌برداری کاربردی را تضمین می‌کند.

به‌منظور طراحی و پیاده‌سازی مدل تحلیل هیجان (شامل احساسات شادی، غم، خشم، ترس، امید، ناامیدی) بیماران سرطانی فارسی‌زبان، نخستین گام گردآوری داده‌های متنی بود. با توجه به محدودیت شدید در منابع آماده، این پژوهش ناچار به ایجادیک پیکره اختصاصی شد. داده‌ها از سه منبع اصلی استخراج گردیدند:

- شبکه اجتماعی ایکس (توییتر سابق): بیماران سرطانی و خانواده‌های آنان معمولاً تجربیات شخصی، احساسات و روند درمان خود را در قالب توییت‌های کوتاه منتشر می‌کنند. این داده‌ها به دلیل خودانگیخته بودن و استفاده از زبان محاوره‌ای، بازتاب دقیقی از هیجانات واقعی بیماران محسوب می‌شوند.
- اینستاگرام: بسیاری از بیمارانیا خانواده‌ها در قالب کپشن و استوری تجربه‌های خود را با دیگران به اشتراک می‌گذارند. جمع‌آوری این داده‌ها با استفاده از خزنده‌های وب و کلیدواژه‌های مرتبط با سرطان انجام شد.
- کانال‌ها و گروه‌های تلگرام: در ایران، تلگرام بستری فعال برای تبادل تجربه‌های بیماران است. پیام‌های منتشرشده در این فضا حاوی حجم زیادی از داده‌های هیجانی و حمایتی است.

در مجموع، حدود ۵۰ هزار جمله خام گردآوری شد که پس از فرایند پاک‌سازی و حذف داده‌های نامرتبط، نزدیک به ۳۰ هزار جمله معتبر باقی ماند. این فرآیند برچسب‌گذاری به گونه‌ای طراحی شد که برچسب‌زن‌ها می‌توانستند تمام هیجانات قابل تشخیص در یک متن را انتخاب کنند. به این ترتیب، هر نمونه داده به جاییک برچسب واحد، با یک بردار باینری به طول شش نمایش داده شد که در آن هر عنصر، بیانگر وجود یا عدم وجود یک هیجان خاص است. برای ارزیابی توافق میان برچسب‌زن‌ها نیز از معیارهای مناسب برای داده‌های چندبرچسبی مانند شاخص کاپای فلیس استفاده شد. همچنین، توافق میان برچسب‌زن‌ها با شاخص کاپای کوهن محاسبه شد که برابر با ۰/۸۲ بود و نشان‌دهنده همبستگی بالا است.

لازم به ذکر است، در فرآیند برچسب‌گذاری داده‌های متنی برای تشخیص هیجانات بیماران، یکی ازمهم‌ترین چالش‌ها، اطمینان از پایایی برچسب‌ها و میزان توافق میان چند برچسب‌زن انسانی است. از آنجا که در این پژوهش هر نمونه می‌تواند به‌صورت چند برچسبی دارای چندهیجان هم‌زمان باشد، ارزیابی میزان توافق برچسب‌زن‌ها اهمیت حیاتی دارد؛ زیرا کیفیت داده‌های آموزشی مستقیماً بر عملکرد مدل ترکیبی پیشنهادی (ParsBERT + Bi-GRU) اثر می‌گذارد. برای سنجش این توافق، از شاخص کاپای فلیس استفاده شده است که تعمیم یافته رویکرد کاپای کوهن برای بیش از دو برچسب‌زن است. تعریف رسمی شاخص کاپای فلیس به‌صورت زیر است:

فرض کنید که برای N نمونه متنی، هر کدام توسط n برچسب‌زن به یکی از k دسته هیجانی نسبت داده شده‌اند. اگر n_{ij} تعداد برچسب زن‌هایی باشد که نمونه iام را در کلاس jام قرار داده‌اند، آنگاه میزان توافق مشاهده شده برای نمونه i به‌صورت (۱) محاسبه می‌شود

$$P_i = ۱ - n_{ij} \sum_{j=۱}^{k} \frac{۱}{n(n-۱)} \qquad (۱)$$

سپس میانگین توافق مشاهده‌شده بر روی تمام نمونه‌ها برابر است با

$$p^{-} = p_i \sum_{i=۱}^{N} \frac{۱}{N} \qquad (۲)$$

همچنین میزان توافق مورد انتظار بر اساس شانس با فرض توزیع برچسب‌ها به‌صورت زیرتعریف می‌شود. ابتدا لازم است نسبت کلی هر کلاس بر اساس (۳) محاسبه شود

$$P_j = n_{ij} \sum_{i=۱}^{N} \frac{۱}{N_n} \qquad (۳)$$

و سپس توافق مورد انتظار

$$p_e^- = \sum_{j=1}^{K} (p_j)^2 \quad (۴)$$

در نهایت شاخص کاپای فلیس به‌صورت (۵) تعریف می‌شود

$$K = \frac{p^- - p_e^-}{1 - p_e^-} \quad (۵)$$

تفسیر مقادیر K به‌صورت استاندارد انجام می‌شود: مقادیر نزدیک به ۱ نشان‌دهنده توافق بسیار بالا، مقادیر بین ۰٫۶ تا ۰٫۸ نشان‌دهنده توافق خوب، و مقادیر کمتر از ۰٫۴ بیانگر توافق ضعیف هستند. در مقاله حاضر، مقدار $K \approx ۰٫۸۲$ گزارش شده است که طبق این تفسیر، نشان‌دهنده توافق عالی میان برچسب‌زن‌ها است. این نتیجه به‌صورت تجربی نشان می‌دهد که تعریف شش دسته هیجانی (شادی، غم، خشم، ترس، امید و ناامیدی) برای متون فارسی بیماران سرطانی به خوبی برای انسان‌ها قابل تمایز بوده و داده‌های برچسب‌خورده از پایایی کافی برخوردارند. ارتباط این شاخص با اعتبار نتایج مدل پیشنهادی بدین‌صورت است که هر چه پایایی برچسب‌ها بیشتر باشد، نویز برچسب‌گذاری کاهش یافته و مدل یادگیری عمیق قادر خواهد بود الگوهای معنایی پایدارتری را بیاموزد. به بیان دیگر، مقدار بالای κ مستقیماً به معنای افزایش اعتبار درونی مجموعه داده آموزشی است و توضیح می‌دهدکه چرا مدل ترکیبی ParsBERT + Bi-GRU توانسته است به مقادیر بالای معیار-F۱ و دقت کلی دست‌یابد. بنابراین، گزارش و تبیین شاخص کاپای فلیس در این مقاله نه تنها یک الزام روش شناختی است، بلکه پشتوانه‌ای برای اعتماد به نتایج تجربی ارائه شده محسوب می‌شود.

در گام بعدی توسعه، متون به زبان فارسی گردآوری‌شده به دلیل ماهیت غیررسمی، نیازمند مراحل گسترده پیش‌پردازش بودند. مراحل کلیدی پیش‌پردازش عبارتند از:

- حذف نویز: شامل حذف لینک‌ها، ایموجی‌های غیرمرتبط، شماره تلفن‌ها و نشانی‌های اینترنتی.
- یکنواخت‌سازی نوشتار: در فارسی،یک واژه ممکن است به اشکال مختلف نوشته شود (مانند «می‌روم» و «میرم»). این موارد با استفاده از کتابخانه Hazm استانداردسازی شدند.
- توکن‌سازی و جداسازی جملات: جملات به واحدهای واژگانی تقسیم شدند تا امکان پردازش توسط مدل فراهم شود.
- حذف توقف‌واژه‌ها: کلماتی مانند «از»، «به» و «که» حذف شدند مگر در مواردی که نقش معنایی در هیجان ایفا می‌کردند.
- ریشه‌یابی و Lemmatization: واژه‌ها به شکل پایه خود بازگردانده شدند تا تنوع صرفی باعث افزایش غیر واقعی ابعاد داده نشود.
- تبدیل ایموجی‌ها به نشانه‌های متنی: بسیاری از بیماران از ایموجی‌ها برای ابراز هیجان استفاده می‌کنند. این موارد به معادل متنی خود (مثلاً ⟵ شادی) تبدیل شدند.

این پیش‌پردازش باعث شد داده‌ها کیفیت بالاتری پیدا کنند و مدل بتواند روابط معنایی دقیق‌تری بیاموزد.مدل‌های زبان بزرگ مانند BERT نشان داده‌اند که در درک معنای عمیق زبان بسیار موفق هستند. در حوزه فارسی نیزParsBERT که بر پایهBERT طراحی شده و بر روی متون فارسی آموزش دیده است، یکی از قدرتمندترین مدل‌ها به شمار می‌رود.از سوی دیگر، شبکه‌های بازگشتی دوطرفه با واحدهای GRU (Bi-GRU) توانایی بالایی در مدل‌سازی وابستگی‌های زمانی دارند. ترکیب این دو رویکرد امکان بهره‌گیری هم‌زمان از مزیت درک معنایی ترنسفورمرها و مدل‌سازی توالی در RNNها را فراهم می‌کند.

معماری سیستم پیشنهادی یک معماری لایه‌ای جامع شامل لایه‌های عملیاتی به شرح زیر می‌باشد:

- لایه ورودی: داده‌های متنی پس از توکن‌سازی، به توکن‌های WordPiece مورد استفاده در ParsBERT تبدیل شدند.
- لایه‌های ParsBERT: مدل ParsBERT با ۱۲ لایه ترنسفورمر، ویژگی‌های معنایی عمیقی از متن استخراج کرد. خروجی این لایه یک بردار متراکم با ابعاد ۷۶۸ بود.
- لایه Bi-GRU: بردارهای خروجی ParsBERT به لایه بازگشتی دوطرفه GRU داده شدند تا وابستگی‌های زمانی به‌ویژه در متون طولانی‌تر مدل‌سازی شوند.
- لایه Attention: یک لایه توجه افزوده شد تا بتواند بخش‌های مهم‌تر متن از نظر هیجانی را وزن‌دهی کند.
- لایه تمام‌متصل Dense: خروجی لایه‌های قبلی به یک لایه Dense با شش نورون متصل شد که هرکدام نمایندهیکی از هیجانات بودند.
- لایه Sigmoid: لایه خروجی Softmax با یک تابع فعال‌سازیSigmoid جایگزین شد. برخلاف Softmax که احتمالات را به گونه‌ای تنظیم می‌کند که جمع آن‌ها یک شود (مناسب برای انتخاب یک گزینه)، Sigmoid احتمال حضور هر هیجان را به صورت مستقل بین۰ و ۱ محاسبه می‌کند. این ویژگی به مدل اجازه می‌دهد تا به طور همزمان چندین هیجان را برایبک متن پیش‌بینی کند.

مشخصات فنی معماری و سیستم پیشنهادی در پیاده‌سازی به شرح زیر می‌باشد:

- اندازه دسته[1]: ۳۲
- نرخ یادگیری اولیه: ۲×۱۰^{-۵}
- تعداد دوره‌های آموزشی[2]: ۱۰
- بهینه‌ساز: AdamW
- تابع هزینه[3]: BCE، این تابع هزینه برای مسائل طبقه‌بندی باینری یا چندبرچسبی طراحی شده است و خطا را برای هر برچسب هیجانی به صورت جداگانه محاسبه می‌کند که کاملاً با معماری جدید مدل سازگار است.
- محیط اجرا: سیستم سخت افزاری دارای کارت گرافیک Nvidia RTX۳۰۶۰ با حافظه GB ۱۲ بود.

بمنظور تحلیل و ارزیابی عملکرد مدل، از مجموعه‌ای از شاخص‌های استاندارد استفاده شد:

- دقّت: نسبت جملات درست پیش‌بینی‌شده به کل جملات.
- دقّت طبقه‌بندی: نسبت نمونه‌های درست در هر دسته به کل نمونه‌های پیش‌بینی‌شده در آن دسته.
- بازخوانی: نسبت نمونه‌های درست شناسایی‌شده در هر دسته به کل نمونه‌های واقعی آن دسته.
- میانگین F۱: میانگین هارمونیک دقّت و بازخوانی که تعادل میان آن‌ها را نشان می‌دهد.
- ماتریس درهم‌ریختگی: برای بررسی خطاهای مدل در دسته‌بندی هیجانات خاص.

1. Batch Size
2. Epochs
3. Binary Cross-Entropy Loss

جدول۲: نتایج ارزیابی مدل پیشنهادی بر روی داده‌های آزمایشی.

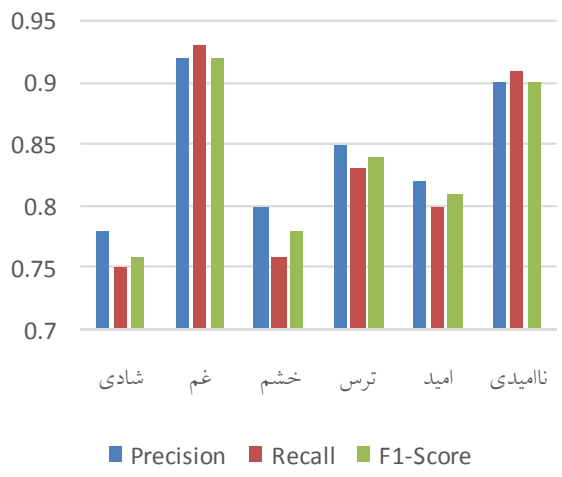
دسته هیجانی	دقّت	بازخوانی	معیار-F۱
شادی	۰٫۷۸	۰٫۷۵	۰٫۷۶
غم	۰٫۹۲	۰٫۹۳	۰٫۹۲
خشم	۰٫۸۰	۰٫۷۶	۰٫۷۸
ترس	۰٫۸۵	۰٫۸۳	۰٫۸۴
امید	۰٫۸۲	۰٫۸۰	۰٫۸۱
ناامیدی	۰٫۸۲	۰٫۸۰	۰٫۸۱
میانگین	۰٫۸۵	۰٫۸۳	۰٫۸۴

لازم به ذکر است، در این پژوهش، با توجه به ماهیت چندبرچسبی مسئله و همچنین نامتوازن بودن توزیع کلاس‌ها (به‌ویژه فراوانی کمتر برخی هیجانات)، معیار اصلی گزارش‌شده برای F۱ از نوع Macro-F۱ بوده است. به این معنا که ابتدا مقدار دقّت و بازخوانی برای هر کلاس هیجانی به‌صورت مستقل محاسبه شده و سپس میانگین ساده آن‌ها بدون وزن‌دهی بر اساس فراوانی کلاس‌ها گرفته شده است. انتخاب Macro-F۱ کاملاً آگاهانه انجام شده است، زیرا این معیار در مسائل دارای عدم توازن داده عملکرد مدل را بر روی کلاس‌های کم‌تکرار (مانند شادی یا خشم) نیز به همان اندازه کلاس‌های پرتکرار در نظر می‌گیرد. بنابراین، گزارش Macro-F۱ رویکردی سخت‌گیرانه‌تر و علمی‌تر برای ارزیابی مدل در این مسئله محسوب می‌شود.در نتیجه، اعداد گزارش‌شده صرفاً میانگین ساده روی نمونه‌ها نیستند، بلکه بر مبنای Macro-F۱ (به‌عنوان معیار اصلی) محاسبه شده‌اند که در شرایط نامتوازن داده، معیار معتبرتر و منصفانه‌تری برای مقایسه مدل‌ها محسوب می‌شود.

علاوه بر این، نتایج مدل پیشنهادی با چند مدل پایه شامل SVM، Naïve Bayes، LSTM و CNN مقایسه شد تا میزان بهبود عملکرد به‌طور کمی نشان داده شود. انتخاب ترکیبParsBERT و Bi-GRU بر اساس چند دلیل سودمند و برتر بود. مهم‌ترین این مزایا به شرح زیر می‌باشند:

- قدرت درک معناییParsBERT: این مدل با یادگیری دوسویه و آموزش بر داده‌های فارسی، قابلیت درک معانی پیچیده و ظرایف زبانی را دارد.
- تواناییBi-GRU در وابستگی‌های زمانی: بیماران اغلب احساسات خود را در قالب جملات پی‌درپی بیان می‌کنند. Bi-GRU با پردازش دوطرفه می‌تواند اطلاعات پیشین و پسین را به‌طور هم‌زمان در نظر بگیرد.
- انعطاف‌پذیری در برابر داده‌های غیررسمی: ترکیب این دو رویکرد باعث می‌شود مدل در مواجهه با متون غیر رسمی، غلط‌های املایی یا زبان محاوره‌ای عملکرد بهتری داشته باشد.
- کارایی بالا: در آزمایش‌های انجام‌شده، این ترکیب توانست نسبت به مدل‌های پایه دقت بالاتری ارائه دهد، به‌ویژه در هیجانات پیچیده‌ای مانند امید و ناامیدی که در مطالعات قبلی کمتر پوشش داده شده بودند.

روش پیشنهادی این پژوهش با ترکیب دو رویکرد قدرتمند در NLP، یعنیParsBERT و Bi-GRU، مدلی کارآمد و بومی برای تحلیل هیجانات بیماران سرطانی فارسی‌زبان ارائه می‌دهد. مراحل گردآوری و پیش‌پردازش داده‌ها، طراحی معماری مدل و معیارهای ارزیابی همگی به‌گونه‌ای انتخاب شده‌اند که علاوه بر دقت علمی، امکان کاربرد عملی در محیط‌های سلامت و روان‌شناسی بالینی نیز فراهم شود.



شکل۲: نمودار عملکرد دسته‌ها (ParsBERT + Bi-GRU).

٤- پیاده‌سازی و ارزیابی نتایج

در این بخش، نتایج حاصل از ارزیابی مدل پیشنهادی (ParsBERT + Bi-GRU) ارائه و تحلیل می‌شود. ابتدا تنظیمات آزمایش و معیارهای ارزیابی معرفی شده، سپس عملکرد کلی مدل بر روی داده‌های آزمایشی بررسی و با مدل‌های پایه و سایر پژوهش‌های مرتبط مقایسه می‌گردد.

٤-۱ تنظیمات آزمایش

برای آموزش و ارزیابی مدل، مجموعه داده ۱۰,۰۰۰ تایی به سه بخش تقسیم شد: ۷۰٪ برای آموزش (۷,۰۰۰ نمونه)، ۱۵٪ برای اعتبارسنجی (۱,۵۰۰ نمونه) و ۱۵٪ برای آزمون (۱,۵۰۰ نمونه). عملکرد مدل با استفاده از معیارهای استاندارد دقت، پرسیژن، بازخوانی و معیار-F۱ ارزیابی گردید. فرایند آموزش مدل با بهینه‌ساز AdamW و نرخ یادگیری ۲×۱۰^{-۵} به مدت ۵ دوره آموزشی انجام شد.

٤-۲ عملکرد کلی مدل پیشنهادی

مدل ترکیبی سیستم پیشنهادی پژوهش حاضر (ParsBERT + Bi-GRU) بر روی مجموعه داده آزمایشی ارزیابی شد و به دقت کلی ۸۶٫۸٪ و میانگینF1-Score معادل ۸۴٪ دست یافت. این نتایج نشان‌دهنده توانایی بالای مدل در تشخیص و طبقه‌بندی شش دسته هیجانی پیچیده در متون غیررسمی فارسی است. جزئیات عملکرد مدل به تفکیک هر دسته هیجانی در جدول ۲ ارائه شده است.

همچنین شکل ۲ نیز نمودار عملکرد دسته‌ها (ParsBERT + Bi-GRU) را نمایش می‌دهد.

تحلیل یافته‌ها نشان می‌دهد که مدل در تشخیص هیجانات پرتکرار و با نشانه‌های واژگانی مشخص، عملکردی عالی داشته است. به‌طور خاص، برای هیجانات«غم» (معیار-F۱: ۹۲٪) و «ناامیدی» (معیار-F۱: ۹۰٪) بالاترین امتیازات به ثبت رسید. در مقابل، عملکرد مدل در تشخیص هیجانات کم‌تکرارتر مانند «شادی» (معیار-F۱: ۷۶٪) و «خشم» (معیار-F۱: ۷۸٪) ضعیف‌تر بود که این امر می‌تواند به دلیل تعداد کمتر نمونه‌های آموزشی برای این دسته‌ها باشد.

٤-۳ تحلیل خطاهای مدل

در خصوص توضیح‌پذیری مدل و عوامل مؤثر بر پیش‌بینی هیجانات، ابتدا برای درک خطاهای مدل، از ماتریس درهم‌ریختگی استفاده شد (جدول ۳).

جدول۳: ماتریس درهم‌ریختگی عملکرد مدل پیشنهادی روی داده‌های آزمایشی.

هیجان واقعی	شادی (پیش‌بینی)	غم (پیش‌بینی)	خشم (پیش‌بینی)	ترس (پیش‌بینی)	امید (پیش‌بینی)	ناامیدی (پیش‌بینی)
شادی	۱۱۲	۵	۰	۲	۲۵	۶
غم	۳	۵۵۸	۴	۱۰	۵	۲۰
خشم	۰	۸	۵۷	۱۰	۰	۰
ترس	۱	۱۲	۸	۱۸۷	۳	۱۴
امید	۱۸	۴	۰	۵	۱۸۰	۱۸
ناامیدی	۲	۲۱	۱	۱۱	۱۵	۳۲۵

تحلیل ماتریس نشان می‌دهد که بیشترین خطاها به دلیل هم‌پوشانی معنایی بین برخی هیجانات رخ داده است:

هم‌پوشانی «شادی» و «امید»: مدل ۲۵ نمونه واقعی «شادی» را به اشتباه «امید» و ۱۸ نمونه «امید» را «شادی» طبقه‌بندی کرده است. این خطا به دلیل وجود واژگان مشترک و نزدیکی معنایی این دو هیجان مثبت است.هم‌پوشانی «غم» و «ناامیدی»: همچنین، مدل ۲۰ نمونه «غم» را «ناامیدی» و ۲۱ نمونه «ناامیدی» را «غم» تشخیص داده است که نشان‌دهنده ارتباط تنگاتنگ این دو هیجان منفی در تجربیات بیماران است.

همچنین برای خروجی‌های طبقه‌بندی چندبرچسبی هیجانات، از روش‌های انتساب اهمیت ویژگی‌ها در سطح توکن[1] استفاده شده است تا مشخص شود کدام کلمات یا عبارات در متن بیماران بیشترین سهم را در فعال شدن هر برچسب هیجانی دارند. خروجی لایه Bi-GRU به‌عنوان تجمیع‌کننده وابستگی‌های ترتیبی عمل کرده و بردار زمینه‌ای نهایی، مبنای تصمیم‌گیری طبقه‌بند چندبرچسبی است که براساس میانگین وزن انتسابی توکن‌ها روی کل مجموعه آزمون محاسبه شده است. به‌عنوان مثال، برای هیجان «غم»، واژگانی با بار معنایی منفی مرتبط با درد، درمان ناموفق و ناامیدی وزن بالاتری داشته‌اند؛ در حالی‌که برای «شادی»، واژگان مرتبط با بهبود وضعیت درمان، حمایت اجتماعی و نتایج مثبت درمان بیشترین سهم را در تصمیم مدل ایفا کرده‌اند. این همخوانی معنایی بین واژگان مؤثر و برچسب هیجانی، اعتبار معنایی[2] تصمیم‌های مدل را تقویت می‌کند.

٤-٤ بحث و مقایسه

به‌منظور سنجش اثربخشی معماری ترکیبی، عملکرد آن با مدل پایه (ParsBERT به‌تنهایی) و همچنین پژوهش‌های مرتبط داخلی و بین‌المللی مقایسه شد (جدول ۴):

نتایج مقایسه نشان‌دهنده چند دستاورد کلیدی است:

- اثربخشی معماری ترکیبی: افزودن لایه Bi-GRU به ParsBERT منجر به بهبود ۲٫۳ درصدی دقت (از ۸۴٫۵٪ به ۸۶٫۸٪) شد. این امر مؤید آن است که مدل‌سازی وابستگی‌های ترتیبی نقش مهمی در افزایش دقت تشخیص و تحلیل هیجانات دارد.برتری در وظیفه پیچیده: اگرچه دقت مدل ما اندکی از پژوهش یزدانی و همکاران (با دقت ~۹۰٪) پایین‌تر است، اما باید توجه داشت که وظیفه پژوهش حاضر (طبقه‌بندی شش رده هیجان) بسیار پیچیده‌تر از تحلیل ۳ دسته احساسی (مثبت، منفی، خنثی) بوده و بر روی داده‌های غیررسمی و پرنویز شبکه‌های اجتماعی انجام شده است.
- کارایی بالاتر نسبت به مدل بین‌المللی: مدل پیشنهادی در مقایسه با رویکرد بین‌المللی Wang و همکاران که از مدل‌های زبان بزرگ برای زبان انگلیسی استفاده کرده‌اند، برتری قابل توجه ۱۸.۹ درصدی در دقت را نشان می‌دهد. این نتیجه، اهمیت و کارایی بالای استفاده از مدل‌های بومی و تخصصی (مانند ParsBERT) برای زبان‌های کم‌منبع را برجسته می‌کند.

٤-٥ دستاوردها و نوآوری‌های پژوهش حاضر

پژوهش حاضر در پاسخ مستقیم به شکاف‌های شناسایی‌شده در ادبیات، خود را به عنوان یک پژوهش بنیادین و کاربردی در حوزه تحلیل هیجان بیماران سرطانی فارسی‌زبان معرفی می‌کند. نوآوری‌های کلیدی این تحقیق که جایگاه آن را در مرزهای دانش مشخص می‌سازد، عبارتند از:

- تمرکز بر زبان فارسی به عنوان یک زبان کم‌منبع: برخلاف جریان غالب پژوهش‌های بین‌المللی، این تحقیق به طور انحصاری بر روی جمعیت بیماران سرطانی فارسی‌زبان متمرکز شده و تلاش می‌کند تا شکاف عمیق منابع و پژوهش را برای این زبان کم‌داده پر کند.
- افزایش تفکیک‌پذیری هیجانات: این پژوهش با فاصله گرفتن از تحلیل قطبیتکلان‌نگر (مثبت/منفی/خنثی)، بر روی شناسایی طیف شش‌گانه‌ای از هیجانات کلیدی و بالینی مهم (شادی، غم، خشم، ترس، امید و ناامیدی) تمرکز دارد. این رویکرد، درک عمیق‌تر و کاربردی‌تری از وضعیت روانی بیماران ارائه می‌دهد.
- ایجاد اولین پیکره تخصصی از داده‌های غیررسمی: این تحقیق با گردآوری و برچسب‌گذاری داده‌ها از پلتفرم‌های تعاملی و غیررسمی مانند ایکس (توییتر سابق)، اینستاگرام و تلگرام، اولین مجموعه داده تخصصی را برای این حوزه ایجاد می‌کند. این رویکرد امکان ثبت صدای اصیل و بدون واسطه بیماران را فراهم می‌آورد که در داده‌های رسمی بالینی یافت نمی‌شود.
- ارائه یک معماری ترکیبی نوآورانه (ParsBERT + Bi-GRU): در بعد فنی، این پژوهش یک مدل ترکیبی نوآورانه را پیشنهاد می‌دهد. این معماری از قدرت ParsBERT برای استخراج بازنمایی‌های متنی غنی از زمینه و از تواناییبیک شبکه Bi-GRU برای مدل‌سازی وابستگی‌های ترتیبی و طولانی‌مدت در متن به‌طور هم‌زمان بهره می‌برد.

در نهایت لازم به ذکر است، فراتر از یک پژوهش نظری، نتایج این تحقیق دارای ارزش کاربردی مستقیم و ملموس برای حوزه پزشکی و سلامت عمومی است. ابزار توسعه‌یافته می‌تواند به عنوان یک سیستم پایش غیرتهاجمی در اختیار متخصصان سلامت روان، انکولوژیست‌ها و

1. Token-level Attribution
2. Semantic Validity

جدول۴: مقایسه کمی عملکرد مدل پیشنهادی با سایر رویکردها.

مدل	دقت کلی	معیار-F۱
سیستم پیشنهادی پژوهش حاضر	٪۸۶٫۸	٪۸۴
ParsBERT منفرد [۳۴]	٪۸۴٫۵	٪۸۲
مدل چندوظیفه ای [۳۱]	+٪۲٫۳	+٪۲٫۰
یزدانی و همکاران [۱۰]	~٪۹۰ (۳دسته)	N/A
ونگ و همکاران [۲۸]	~٪۷۳ (۲–۴ دسته)	~٪۷۲

سیاست‌گذاران قرار گیرد تا با درک بهتر هیجانات بیماران، مداخلات روان‌شناختی پیشگیرانه و به‌موقع را طراحی و اجرا نمایند.

۵- نتیجه‌گیری و راهکارهای آتی پژوهش

این پژوهش با هدف طراحی و ارزیابی یک رویکرد نوین یادگیری عمیق برای تحلیل خودکار هیجانات بیماران سرطانی در شبکه‌های اجتماعی فارسی ارائه گردید. با استفاده از رویکرد چندبرچسبی، پژوهش توانسته است به درک واقع‌گرایانه‌تری از تجربیات هیجانی بیماران دست یابد. این سیستم نه‌تنها هیجانات غالب، بلکه الگوهای هم‌زمانی هیجانات (مانند ترکیب امید و ترس) را نیز شناسایی می‌کند که این اطلاعات برای مداخلات روان‌شناختی هدفمند بسیار ارزشمندتر است.با توجه به چالش‌های منحصربه‌فرد این حوزه، از جمله کمبود منابع داده در زبان فارسی و ماهیت غیررسمی متون، یک معماری ترکیبی مبتنی بر مدل زبان بزرگ بومیParsBERT و شبکه بازگشتی دوطرفه Bi-GRU توسعه داده شد. نتایج ارزیابی نشان داد که مدل پیشنهادی با دستیابی به دقت کلی ٪۸۶٫۸ و معیارF1-Score میانگین ٪۸۴ در طبقه‌بندی شش دسته هیجانی، عملکردی موفق و قابل‌اتکا دارد. این دستاورد، بهبودی معادل ٪۲٫۳ در دقت نسبت به مدل پایه (ParsBERT به‌تنهایی) را نشان می‌دهد و برتری معماری ترکیبی در درک وابستگی‌های ترتیبی متن را اثبات می‌کند. علاوه بر این، مدل ما با ثبت دقتی حدود ٪۱۸٫۹۸ بالاتر از رویکردهای بین‌المللی مبتنی بر LLM، پتانسیل بالای مدل‌های تخصصی و بومی‌سازی شده را در وظایف حوزه سلامت به نمایش گذاشت. مهم‌ترین دستاورد این پژوهش، ارائه یک سیستم کارآمد برایبک حوزه پر اهمیت و کم‌منبع است که می‌تواند به عنوان یک ابزار مکمل در کنار روش‌های بالینی برای پایش سلامت روان بیماران به کار گرفته شود.

در نهایت به عنوان راهکارهای آتی پژوهش،درصدد توسعه تحلیل مقطعی فعلی به یک تحلیل طولی و گسترش یافته هستیم تا بتوان مسیر هیجانی بیماران را در طول زمان، از تشخیص تا بهبودی ردیابی کرد. این رویکرد می‌تواند به شناسایی نقاط بحرانی روانی در فرآیند درمان و طراحی مداخلات پیشگیرانه منجر شود.

مراجع

[1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1-135, 2008.

[2] J. Winman, "Health psychology," *Journal of Philosophy and Theology*, vol. 9, no. 33-34, pp. 109-129, May 2004.

[3] R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping*, Springer Publishing Company, 1984.

[4] A. Krebber, *et al.*, "Prevalence of depression in cancer patients: A meta-analysis of diagnostic interviews and self-report instruments," *Psycho-Oncology*, vol. 23, no. 2, pp. 121-130, Feb. 2014.

[5] C. Strapparava, "WordNet-Affect: an affective extension of WordNet," in *Proc. of the 4th Int. Conf. on Language Resources and Evaluation*, pp. 1083-1086, Lisbon, Portugal, 26-28 May 2004.

[6] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *Proc. of the 7th Int. Conf. on Language Resources and Evaluation*, pp. 2200-2204, Valletta, Malta, 17-23 May 2010.

[7] D. Seal, U. Roy, and R. Basak, "Sentence-level emotion detection from text Based on semantic rules," in *Information and Communication Technology for Sustainable Development,* Advances in Intelligent Systems and Computing, vol. 933, pp. 423-430, Jul. 2020.

[8] M. Bazargani, N. Fareghzadeh, M. Afzali, S. Karimi, "A hybrid cumulative knowledge framework for friend recommendation in social networks", *International Journal of Nonlinear Analysis and Applications*, vol. 17, no. 1, pp. 43-54, Jan. 2026.

[9] L. Wikarsa and S. N. Thahir, "A text mining application of emotion classifications of Twitter's users using Naïve Bayes method," in *Proc. 1st. Int. Conf. on Wireless and Telematics*, 6 pp., Manado, Indonesia, 17-18 Nov. 2015.

[10] A. Yazdani, *et al.*, "Use of sentiment analysis for capturing hospitalized cancer patients' experience from free-text comments in the Persian language." *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, Article ID: 275, Dec. 2023.

[11] Y. K. Meena, P. M. Kumar, V. V. Kumar., "Classification of online patient reviews based on effectiveness using machine learning algorithms," in *Proc. 3rd Int. Conf. on Trends in Electronics and Informatics*, pp. 838-842, Apr. 2019.

[12] N. Bahrawi., "Sentiment analysis using random forest algorithm-online social media based," *Journal of Information Technology and Its Utilization*, vol. 2, no. 2, pp. 29-33, Dec. 2019.

[13] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, Oct. 2001.

[۱۴] س. ذ. هاشمی‌پور، "تشخیص سرطان سینه با استفاده از الگوریتم جنگل تصادفی،" *هشتمین کنگره ملی تازه‌های مهندسی برق و کامپیوتر ایران*، ۶ صص، تهران، ایران، ۲۸–۲۷ اردیبهشت ۱۴۰۰.

[15] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. on Information Theory*, vol. 13, no. 1, pp. 21-27, Jan. 1967.

[16] A. D. W. Sumari, A. F. Huda, and M. A. Fauzi, "Sentiment analysis of patient's satisfaction in health services using k-nearest neighbor and Chi-Square," *Journal of Computer Science*, vol. 17, no. 1, pp. 57-65, Jan. 2021.

[۱۷] ن. فارغزاده، م. بازرگانی ن. جعفری، "تشخیص سرطان سینه با استفاده از الگوریتم جنگل تصادفی،" *اسلام شناسی و قرآن پژوهی در جهان معاصر*، شناسه مقاله: e734628، بهمن ۱۴۰۴.

[18] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Computational Intelligence Magazine,* vol. 13, no. 3, pp. 55-75, Aug. 2018.

[19] Y. Kim, "Convolutional neural networks for sentence classification," in *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing,* pp. 1746-1751, Doha, Qatar, 25-29 Oct. 2014.

[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

[21] K. Cho, *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation in *Proc. of the Conf. on Empirical Methods in Natural Language Processing,* pp. 1724-1734, Doha, Qatar, 25-29 Oct. 2014.

[22] C. Wang, J. Yang, M. Yu, L. Lv, and X. Wang, "A hybrid CNN-LSTM model for sentiment analysis of online reviews," in *Proc. of the 12th Int. Conf. on Intelligent Computation Technology and Automation*, pp. 248-252, Xiangtan, China, 26-27 Oct. 2019.

[23] B. K. Crannell, E. Clark, C. C. Jones, and D. G. T. III, "Characterizing cancer-related patient-reported comments on Twitter," in *Proc. of the IEEE Int. Conf. on Healthcare Informatics*, pp. 526-532, Chicago, IL, USA, 4-7 Oct. 2016.

[24] A. C. Sosea and C. Caragea, "CancerEmo: A dataset for fine-grained emotion detection," in *Proc. of the 28th Int. Conf. on Computational Linguistics*, pp. 5987-6000, Barcelona, Spain, 8-13 Dec. 2020.

[25] A. Vaswani, *et al.*, "Attention is all you need," in *Proc. of the 31st Annual Conf. on Neural Information Processing Systems*, pp. 5998-6008, Long Beach, CA, USA, 4-9 Dec. 2017.

[26] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. of the* Conf. *of the North American Chapter of the Association for Computational Linguistics: Human Language*

Technologies, pp. 4171-4186, Minneapolis, MN, USA, 2-7 Jun. 2019.
[27] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI Blog*, vol. 1, no. 8, 2019.
[28] J. Lee, *et al.*, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234-1240, Feb. 2020.
[29] Z. Wang, K. E. Daniel, L. E. Barnes, and P. I. Chow, CALLM: Understanding Cancer Survivors' Emotions and Intervention Opportunities via Mobile Diaries and Context-Aware Language Models, arXiv preprint arXiv:2503.10707, 2025.
[30] H. Guo, Y. Zhang, and J. Luo, "ClinicalBERT for sentiment analysis and entity extraction in electronic health records," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 4, pp. 1821-1832, Apr. 2023.
[31] A. Allam, G. T. Jones, K. E. Stewart, and D. J. L. Jones, "Sentiment analysis of clinical narratives to identify patients with depression and anxiety: A systematic review," *JMIR Medical Informatics*, vol. 9, no. 7, Article ID: e25860, Jul. 2021.
[32] J. Li, R. Wang, and S. Zhao, "A transformer-based multi-task framework for sentiment analysis in digital health applications," *Expert Systems with Applications*, vol. 201, Article ID: 117091, Sept. 2022.
[33] H. Sharifi, H. Faili, and M. Asad, "Challenges and opportunities in Persian natural language processing," *Journal of Signal and Data Processing*, vol. 16, no. 1, pp. 3-28, Mar. 2019.
[34] M. Rasooli, M. Kouhpaee, and O. Fehresti, "Linguistic challenges of Persian-English machine translation," in *Proc. of the 13th Conf. of the European Chapter of the Association for Computational Linguistics*, pp. 24-33, Avignon, France, 23-27 Apr. 2012.
[35] A. Roshanfekr, M. A. M. Abadi, and S. A. Mirroshandel, "Deep learning for sentiment analysis of Persian texts," in *Proc. 3rd Int. Conf. on Web Research*, pp. 101-106, Tehran, Iran, 19-20 Apr. 2017.
[36] H. Asgarian, H. S. Aghdasi, and M. H. Sadreddini, "ArmanEmo: A New Dataset for Emotion Detection in Persian Texts," *in Proc. of the 5th Int. Conf. on Web Research*, pp. 136-141, Tehran, Iran, 24-25 Apr. 2019.
[37] M. Yazdani, A. R. Hamidi, and S. Kohan, "Sentiment analysis of hospitalized patients' comments in Persian using machine learning," *Journal of Health and Biomedical Informatics*, vol. 10, no. 2, pp. 123-134, Sept. 2023.
[38] S. M. H. Ghavanini, M. R. H. Ghavanini, and A. M. E. Moghadam, "Sentiment analysis of Persian social media texts: A survey," in *Proc. of the 26th Int. Computer Conf. Computer Society of Iran*, 6 pp., Tehran, Iran, 3-4 Mar. 2021.
[39] E. Bagheri, M. H. Saraee, and F. de la Prieta, "Challenges of sentiment analysis in informal Persian texts: a survey," *Artificial Intelligence Review*, vol. 55, no. 1, pp. 1-36, Jan. 2022.
[40] H. Sharifi, H. Faili, and M. Asad, "A survey on the state of resources and tools for Persian natural language processing," *Journal of Advances in Computer Research*, vol. 11, no. 1, pp. 15-32, Feb. 2020.
[41] A. Jafer, "A survey on sentiment analysis for low-resource languages," *ACM Trans. on Asian and Low-Resource Language Information Processing*, vol. 20, no. 4, Article ID: 65, Jul. 2021.
[42] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," *American Scientist*, vol. 89, no. 4, pp. 344-350, Jul./Aug. 2001.

نفیسه فارغزاده مدرک دکترای تخصصی خود را در رشته علوم کامپیوتر و گرایش سیستم‌های نرم افزاری در سال ۱۳۹۷ از دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران دریافت کرد. ایشان عضو هیأت علمی دانشگاه آزاد اسلامی واحد خدابنده، استان زنجان می‌باشد و تاکنون کتاب‌ها و مقالات متعدد در حوزه سیستم‌های هوشمند و سامانه‌های توصیه‌گر منتشر کرده است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: سیستم های خبره و هوش مصنوعی، داده کاوی و سیستم‌های توصیه‌گر.

مهدی قبادی مدرک کارشناسی ارشد خود را در رشته مهندسی کامپیوتر از دانشگاه آزاد اسلامی واحد الکترونیک تهران دریافت کرد. ایشان چندین مقاله علمی در حوزه علوم کامپیوتر منتشر کرده است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: هوش مصنوعی و داده کاوی.

پریسا رحمانی مدرک کارشناسی خود را در سال ۱۳۸۲ و مقطع کارشناسی ارشد در رشته مهندسی نرم‌افزار را در سال ۱۳۸۴ از دانشگاه آزاد اسلامی واحد قزوین اخذ نمود. ایشان در سال ۱۳۹۶ موفق به دریافت درجه دکتری تخصصی در رشته مهندسی نرم‌افزار از دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران گردید. زمینه‌های پژوهشی و تخصصی وی مشتمل بر یادگیری ماشین، الگوریتم‌های فراابتکاری و تخصیص منابع در شبکه‌های بی‌سیم، از جمله تخصیص کانال و توان در شبکه‌های ادهاک می‌باشد.

مهدی بازرگانی مدرک دکترای تخصصی خود را در رشته علوم کامپیوتر و گرایش سیستم های نرم افزاری از دانشگاه آزاد اسلامی واحد قزوین دریافت کرد. ایشان عضو هیأت علمی دانشگاه آزاد اسلامی واحد زنجان می‌باشد و تاکنون مقالات متنوعی در حوزه سیستم های هوشمند منتشر کرده است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: هوش مصنوعی، بلاکچین و اینترنت اشیا.