

# مروری بر روش‌های یادگیری عمیق برای طبقه‌بندی هیجان در متن: پیشرفت‌ها، چالش‌ها و فرصت‌ها

مهدی رسولی<sup>۱</sup>، دانشجوی کارشناسی ارشد، وحید کیانی<sup>۲\*</sup>، استادیار

<sup>۱</sup> دانشکده مهندسی - دانشگاه بجنورد - بجنورد - ایران - mrasouli.edu@gmail.com

<sup>۲</sup> دانشکده مهندسی - دانشگاه بجنورد - بجنورد - ایران - v.kiani@ub.ac.ir

چکیده: امروزه افراد در بستر وب احساسات و هیجان خود را با کمک ابزارهای ارتباطی مختلف به اشتراک می‌گذارند، که یکی از رایج‌ترین آن‌ها بیان احساسات در محتوای متنی مانند پست‌های رسانه‌های اجتماعی، نظرات فروشگاه‌های آنلاین و مرورهای کاربران است. تشخیص هیجان در متن شاخه‌ای از تحلیل احساسات است که هدف آن شناسایی انواع مختلف هیجان نویسنده در متن است. این حوزه علمی به تولیدکنندگان و ارائه‌کنندگان خدمات کمک می‌کند تا از نقاط ضعف و قوت خود آگاه شده و خدمات بهتری را برای مشتریان فراهم آورند. در سال‌های اخیر، تشخیص هیجان در متن به دلیل کاربردهای گسترده‌اش در تجارت، اقتصاد، سیاست، پزشکی، روانشناسی و جامعه‌شناسی به یک زمینه تحقیقاتی جذاب تبدیل شده است. در این مقاله مسئله طبقه‌بندی هیجان در متن و روش‌های حل آن با تأکید بر یادگیری عمیق بررسی خواهد شد. همچنین شرح مختصری بر جدیدترین راهکارهای یادگیری عمیق ارائه خواهد شد که در سال‌های اخیر برای طبقه‌بندی هیجان در متن مورد استفاده قرار گرفته‌اند. علاوه بر این، تعدادی مجموعه داده برچسب‌گذاری شده، مهم‌ترین مسائل باز در تشخیص هیجان و جهت‌گیری‌های تحقیقاتی آینده نیز مطرح خواهند شد که می‌تواند راهنمای خوبی برای محققین جدید این حوزه باشد.

واژه‌های کلیدی: پردازش زبان طبیعی، تحلیل احساسات، شناسایی هیجان، طبقه‌بندی هیجان در متن، یادگیری عمیق.

\* نویسنده مسئول، v.kiani@ub.ac.ir

# A survey on deep learning methods for text-based emotion classification: Advances, challenges, and opportunities

---

Mahdi Rasouli<sup>1</sup>, MSc. Student, Vahid Kiani<sup>2\*</sup>, Assistant Professor

<sup>1</sup> Dept. of Engineering, University of Bojnord, Bojnord, Iran, mrasouli.edu@gmail.com

<sup>2</sup> Dept. of Engineering, University of Bojnord, Bojnord, Iran, v.kiani@ub.ac.ir

**Abstract:** Today, people on the web share their feelings and emotions with the help of various communication tools, one of the most common of which is the expression of feelings in textual content such as social media posts, online store reviews, and user reviews. Emotion detection in text is a branch of sentiment analysis that aims to identify different types of human emotion in the text. This scientific field helps manufacturers and service providers to be aware of their weaknesses and strengths, and to provide better services to customers. In recent years, emotion recognition in text has become an attractive research field due to its wide applications in business, economics, politics, medicine, psychology, and sociology. In this article, the problem of emotion classification in text and its solution methods will be investigated with emphasis on deep learning. Also, a brief description of the latest deep learning solutions that have been used in recent years to classify emotion in text will be discussed. In addition, a number of labelled datasets, the most important open issues in emotion recognition, and future research directions will also be presented, which can be a good guide for new researchers in this field.

**Keywords:** *natural language processing; sentiment analysis, emotion detection; emotion classification in text; deep learning.*

\* Corresponding author, v.kiani@ub.ac.ir

## ۱. مقدمه

عملکرد خود اعمال کنند [۸]. نمونه‌هایی از کاربردهای تحلیل هیجان در متن شامل ارزیابی اثرات تصمیمات یا رویدادهای سیاسی بر جامعه، درک نقاط ضعف و قوت محصولات و اصلاح محصول، اتخاذ استراتژی‌های بازاریابی مناسب، توصیه فیلم و موسیقی، پیش‌بینی تغییرات احتمالی قیمت‌ها در بازارهای مالی، اتخاذ روش‌های آموزشی مناسب در محیط‌های آموزشی، تعیین اثرات روانی اخبار بر جامعه، تحلیل اثرات رویدادهای اجتماعی مانند همه‌گیری بیماری کرونا و افزایش ادراک ربات‌ها در تعامل با انسان است [۶].

این مقاله قصد دارد خواننده را با مسئله طبقه‌بندی هیجان در متن و روش‌های مختلف حل آن به خصوص روش‌های مبتنی بر یادگیری عمیق آشنا کند. ساختار بخش‌های بعدی مقاله به این صورت است: در بخش دوم این مقاله، مدل‌های مختلف هیجان مطرح خواهند شد. در بخش سوم، مسئله تشخیص هیجان در متن معرفی خواهد شد. بخش چهارم به معرفی بعضی مجموعه‌داده‌های برچسب‌گذاری شده برای طبقه‌بندی هیجان در متن اختصاص دارد. در بخش پنجم، مباحث بنیادین در یادگیری عمیق تشریح خواهند شد. در بخش ششم، بعضی راهکارهای رایج در حوزه یادگیری عمیق معرفی شده‌اند، که در سال‌های اخیر برای طبقه‌بندی هیجان در متون غیر فارسی مورد استفاده قرار گرفته‌اند. در بخش هفتم، آخرین تحقیقات انجام شده در زمینه طبقه‌بندی هیجان در متن فارسی بررسی شده‌اند. بخش هشتم به بررسی چالش‌های باقیمانده و فرصت‌های پژوهشی پیش‌رو پرداخته است. در نهایت، بخش نهم به نتیجه‌گیری از مقاله می‌پردازد.

## ۲. مدل‌های هیجان

هیجان را می‌توانیم به‌عنوان تجربه و جریان قوی احساسات در افراد در پاسخ به یک رویداد یا وضعیت تعریف کنیم [۹].

در طول پنجاه سال اخیر، هوش مصنوعی به ارائه راه‌حل‌های موثری برای مسائل عمده انسانی و اجتماعی در حوزه‌های مختلف از جمله تحلیل معنایی متن و ترجمه ماشینی نائل آمده است [۱]. پردازش زبان طبیعی<sup>۱</sup>، شاخه‌ای از هوش مصنوعی است، که از فنون زبانی و محاسباتی برای کمک به رایانه‌ها در درک و تولید زبان‌های انسانی به شکل متن و گفتار استفاده می‌کند [۲]. از جمله مهم‌ترین مسائل در حوزه پردازش زبان طبیعی می‌توان به ترجمه خودکار، سیستم‌های پرسش و پاسخ، بازاریابی و جستجوی اطلاعات، خلاصه‌سازی متن و تحلیل احساسات اشاره کرد [۳]، [۴]. هدف تحلیل احساسات در متن، تحلیل زبان انسانی است به گونه‌ای که عقاید، ایده‌ها و تفکرات انسان به صورت خودکار از متن استخراج شود. در تحلیل احساسات، عموماً قطبیت نظر نویسنده متن در حالت‌های مثبت، منفی و خنثی ارزیابی می‌شود [۵]. تشخیص هیجان در متن شاخه‌ای از تحلیل احساسات است که به جای در نظر گرفتن سه حالت احساسی کلی مثبت، منفی و خنثی، بر تشخیص انواع مختلف هیجان در انسان تمرکز دارد که حالت‌های بیشتری مانند ترس، شادی، غم، عصبانیت و تعجب را پوشش می‌دهد [۶]. تشخیص هیجان می‌تواند در متن، تصویر، صوت، یا ویدئو انجام شود. متن نسبت به سایر رسانه‌ها در دسترس‌تر است، اما برداشت‌های احساسی و هیجانی مختلفی می‌تواند از یک متن انجام شود. به همین دلیل، تشخیص هیجان در متن یک حوزه تحقیقاتی جذاب، پرکاربرد و چالش برانگیز است [۷].

تحلیل هیجان در متن کمک می‌کند تا صاحبان شرکت‌ها و صنایع، سیاستمداران، تولیدکنندگان، خبرگزاری‌ها و شبکه‌های خبری از احساسات جوامع انسانی حوزه خود آگاه شده، تصمیم‌های مناسبی را اتخاذ کرده و اصلاحات لازم را در

<sup>1</sup> Natural Language Processing (NLP)

ترس و تعجب ابروها بالا می‌روند و در هر دو حالت نفرت و عصبانیت بینی چروکیده می‌شود، در تحقیق مذکور انواع هیجان اساسی انسان به جای مدل شش دسته‌ای اکمن، به تنها چهار دسته عصبانیت، ترس، شادی و غم تقسیم شده است. حالت هیجانی اساسی و حالات چهره متناظر با آن‌ها برای مدل اکمن و مدل راشل‌جک در شکل (۱) نمایش داده شده‌اند.

۳. مدل پاروت: مدل پاروت<sup>۴</sup> نیز شش هیجان را به عنوان هیجان‌های اساسی در نظر گرفته است، که شامل ترس، غم، تعجب، عصبانیت، عشق و شادی هستند [۱۴]. علاوه بر این‌ها پاروت حالت‌های هیجانی ثانویه را نیز برای انسان مطرح کرده و تا ۱۰۰ حالت هیجانی را در یک ساختار درختی سازماندهی کرده است. دو سطح ابتدایی از درخت هیجان پاروت در شکل (۲) نمایش داده شده است.

۴. مدل پلاچیک: رابرت پلاچیک<sup>۵</sup> تعداد هیجان اساسی را هشت نوع هیجان در نظر گرفته است [۱۵]، که در مقابل یکدیگر مطرح شده‌اند. در این مدل، شادی در مقابل غم، اعتماد در مقابل نفرت، عصبانیت در مقابل ترس و تعجب در مقابل اضطراب مطرح شده است. مدل پلاچیک نسبت به مدل اکمن شامل دو هیجان بیشتر یعنی اعتماد و اضطراب است. همچنین، پلاچیک بر خلاف اکمن به امکان ترکیب حالت‌های هیجانی نیز معتقد بود. پلاچیک علاوه بر هشت نوع هیجان اساسی، در «دایره هیجان پلاچیک»<sup>۶</sup> بیست و چهار نوع هیجان فرعی را نیز مطرح کرده است، که از ترکیب انواع هیجان با یکدیگر حاصل می‌شوند.

هیجان یک واکنش شدید و کوتاه مدت است که فرد معمولاً نسبت به وقوع آن کاملاً آگاه است. به عنوان مثال شادی، غم، عصبانیت و ترس نمونه‌هایی از انواع هیجان هستند. علت هیجان می‌تواند یک عامل خارجی مانند دیدن یک فیلم یا رخداد یک واقعه خاص در جامعه، یا یک احساس درونی فرد مانند حس افسردگی یا شادمانی باشد. هیجان یک جزء جدانشدنی از زندگی ما انسان‌ها است، که به ما کمک می‌کند تا با دیگران ارتباط برقرار کنیم و با بقیه افراد جامعه همسو و هم‌صدا شویم [۱۰].

مدل‌های هیجان جزء بنیادین و اساسی در سیستم‌های تشخیص هیجان هستند، که نحوه بازنمایی انواع هیجان را در سیستم تشخیص هیجان تعیین می‌کنند. در طبقه‌بندی هیجان، از مدل‌های گسسته هیجان استفاده می‌شود. این مدل‌ها، هیجان انسان را به کمک تعدادی حالت گسسته توصیف می‌کنند. مهم‌ترین مدل‌های گسسته هیجان عبارتند از:

۱. مدل اکمن: رایج‌ترین دسته‌بندی برای انواع هیجان نظریه بزرگ شش دسته‌ای پائول اکمن<sup>۲</sup> است که انواع حالت‌های هیجانی انسان را به شش نوع ترس، عصبانیت، شادی، غم، نفرت و تعجب تقسیم می‌کند [۱۱]، [۱۲]. مدل پائول اکمن تاکنون با موفقیت برای توصیف حالات چهره انسان به کار بسته شده است. با این وجود، در بعضی کاربردها، شش حالت هیجانی پیشنهاد شده توسط اکمن برای توصیف انواع هیجان انسان کافی نیستند.

۲. مدل راشل‌جک<sup>۳</sup>: در یکی از آخرین مطالعه‌ها توسط راشل‌جک [۱۳] در سال ۲۰۱۶، تنها حالت‌هایی در انسان به عنوان هیجان اساسی در نظر گرفته شده‌اند که حالات چهره متفاوتی را در انسان ایجاد کنند. از آنجایی که در هر دو حالت

<sup>4</sup> Parrot

<sup>5</sup> Plutchik

<sup>6</sup> Plutchik's Wheel of Emotions

<sup>2</sup> Paul Ekman

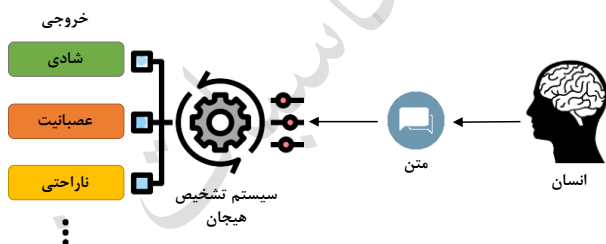
<sup>3</sup> Rachael Jack

تشخیص دهد. کائو و همکاران مسئله تشخیص هیجان در متن را به زبان ریاضی به شکل زیر بیان کرده‌اند [۱۶]:

$$r: A * T \rightarrow E \quad (1)$$

که در آن  $T$  متن ورودی است که باید هیجان آن مشخص شود،  $A$  نویسنده متن و  $E$  هیجان موجود در متن است. سمبل  $r$  رابطه بین نویسنده، متن و هیجان را نشان می‌دهد. این رابطه بیان می‌کند که هیجان خروجی به متن نوشته شده و نویسنده آن بستگی دارد. یک سیستم تشخیص خودکار هیجان معمولاً تنها متن ورودی را دریافت می‌کند و از هویت نویسنده یا تفکرات او ناآگاه است. به همین دلیل، تشخیص هیجان در متن می‌تواند چالش برانگیز باشد.

تشخیص هیجان در متن، در واقع یک مسئله طبقه‌بندی چندکلاسه است. در سیستم طبقه‌بندی هیجان همان‌طور که در شکل (۳) مشاهده می‌کنید، داده ورودی به صورت متن به سیستم تشخیص وارد می‌شود. سپس سیستم با پردازش متن ورودی، خروجی را به صورت انواع هیجان مانند شادی، غم، عصبانیت و غیره اعلام می‌کند.

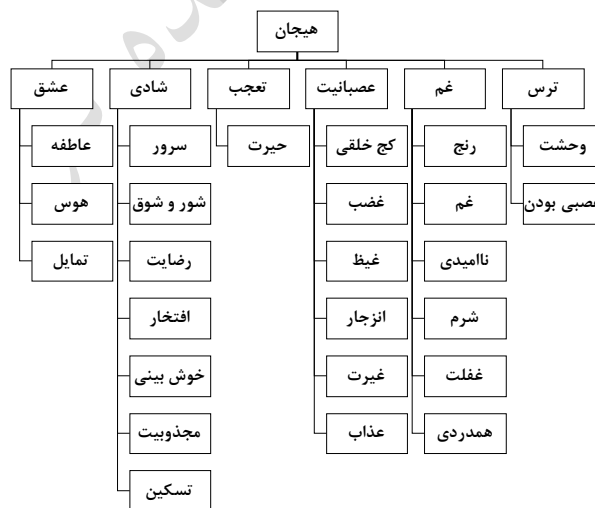


شکل (۳): عملکرد سیستم تشخیص هیجان در متن

طبقه‌بندی هیجان در متن یک فعالیت چالش برانگیز است. زیرا متن ممکن است حاوی خطاهای گرامری باشد، به صورت غیررسمی و محاوره‌ای نوشته شده باشد، بسیار کوتاه باشد و کنایه و ضرب‌المثل نیز در آن وجود داشته باشد. همچنین، انواع هیجان را نمی‌توان به نشانه‌های یکتا و ویژه‌ای در متن مرتبط



شکل (۱): مدل شش مقوله‌ای پائول اکمن و مدل چهار مقوله‌ای راشل جک برای انواع هیجان در انسان



شکل (۲): مدل درختی پاروت برای انواع هیجان در انسان

از مجموعه انواع مدل‌های هیجان که در این بخش معرفی شدند، مدل شش مقوله‌ای اکمن بیشتر از بقیه در تحلیل هیجان متن مورد استفاده قرار گرفته است، هر چند استفاده از مدل‌های پیچیده‌تر می‌تواند به مدل‌سازی دقیق‌تر انواع هیجان منجر شود.

### ۳. تشخیص هیجان در متن

تشخیص خودکار هیجان در متن شامل فرآیندهایی مبتنی بر هوش مصنوعی است که طی آن‌ها سیستم هوشمند باید با مدل‌سازی و بررسی متون، انواع هیجان موجود در آن‌ها را

#### ۴. مجموعه داده‌ها

یکی از چالش‌ها در تولید سیستم‌های تشخیص هیجان در متن، یافتن مجموعه داده‌ای با برچسب‌گذاری مناسب برای ارزیابی دقت نهایی سیستم است. تعداد قابل توجهی از مجموعه داده‌های معیار به راحتی در دسترس بوده و به صورت رایگان قابل دریافت هستند، که در این بخش به معرفی برخی از آن‌ها می‌پردازیم. مشخصات این مجموعه داده‌ها در جدول (۱) خلاصه شده است. توزیع نمونه‌ها بین طبقات هیجانی مختلف در اکثر این مجموعه داده‌ها یکسان نیست و تشخیص هیجان در آن‌ها یک مسئله طبقه‌بندی نامتوازن است.

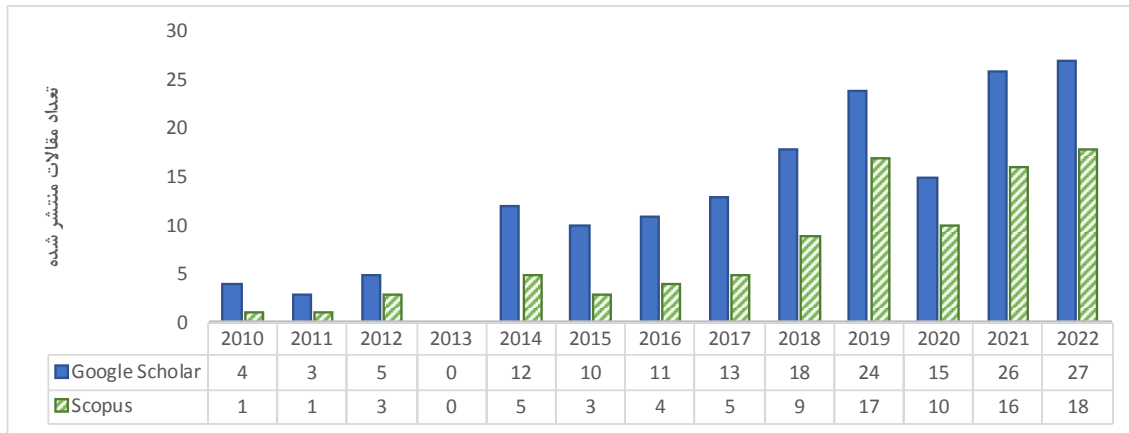
#### ۴.۱. مجموعه داده‌های فارسی

تا کنون، تحقیقات معدودی در زمینه تشخیص هیجان در متون فارسی انجام شده است. به همین دلیل، تعداد مجموعه داده‌های فارسی در دسترس برای تشخیص هیجان در متن فارسی محدود است. از بین مجموعه داده‌های فارسی مناسب برای طبقه‌بندی هیجان، تنها سه مجموعه داده [۱۷] EmoPars، Persian Tweets [۱۸] و ArmanEmo [۱۹] در دسترس عموم قرار دارند. هر سه مجموعه داده حاوی توثیتهای فارسی هستند که طبق مدل اکمن در دسته‌های ترس، عصبانیت، شادی، غم، نفرت و تعجب برچسب‌گذاری شده‌اند. مجموعه داده EmoPars چندبرچسبی است و برای هر نمونه در طبقات هیجانی مختلف از پنج کاربر انسانی رأی‌گیری شده است. در مقابل، مجموعه داده‌های PersianTweets و ArmanEmo به صورت تک‌برچسبی برچسب‌گذاری شده‌اند و هر نمونه در آن‌ها تنها یک برچسب دارد. مجموعه داده ArmanEmo علاوه بر توثیتهای فارسی، حاوی نظرات کاربران دیجی کالا و نظرات کاربران در اینستاگرام نیز هست. مجموعه داده EmoPars [۲۰]، مجموعه داده Persian Tweets [۱۸] و مجموعه داده ArmanEmo [۲۱] در دسترس عموم قرار دارند.

کرد و تفسیر معنای هر کلمه به زمینه متن و کلمات اطراف آن نیز وابسته است [۶]. علیرغم این چالش‌ها، تشخیص هیجان در متن امری ضروری و اجتناب ناپذیر است. تشخیص هیجان در متن به بهبود تعامل انسان با کامپیوتر کمک می‌کند و در بسیاری از کاربردها تنها رسانه در دسترس برای تشخیص هیجان، متن نوشته شده توسط افراد است. همچنین، امروزه استفاده از متن برای بیان عقاید و نظرات بسیار رایج شده است. سیستم تشخیص هیجان در متن می‌تواند در ساخت چت بات‌های هوشمند برای مشاوره، شناسایی جملات توهین آمیز در مکالمات، تحلیل وضعیت بیماران روانی و افراد مستعد خودکشی، تحلیل نظرات کاربران در شبکه‌های اجتماعی، بررسی رضایت کاربران در فروشگاه‌های آنلاین و بسیاری کاربردهای دیگر به کار برده شود [۶]، [۸]، [۱۴].

#### ۳.۱. حجم تحقیقات تشخیص هیجان در متن

برای ارزیابی میزان تحقیقات انجام شده در زمینه تشخیص هیجان در متن در طول سال‌های مختلف، دو موتور جستجوی علمی Google Scholar و Scopus مورد بررسی قرار گرفتند. در هر یک از این موتورهای جستجوی بین‌المللی، مقالاتی که در عنوان آن‌ها سه کلمه Emotion Detection Text آمده بود، مورد جستجو قرار گرفتند. تعداد مقالات یافت شده در سال‌های ۲۰۱۰ تا ۲۰۲۲ به تفکیک موتورهای جستجو در شکل (۴) نمایش داده شده‌اند. نتایج نشان می‌دهند که از سال ۲۰۱۴ به بعد، به تدریج تحقیقات در زمینه شناسایی هیجان در متن سیر صعودی به خود گرفته است و در سال ۲۰۲۲ به حداکثر مقدار خود رسیده است. این موضوع نشان می‌دهد تشخیص هیجان در متن یک حوزه داغ تحقیقاتی است، که می‌تواند مورد پژوهش بیشتر قرار بگیرد.



شکل (۴): روند افزایش تحقیقات در زمینه تشخیص هیجان در متن از سال ۲۰۱۰ تا ۲۰۲۲

جدول (۱): مجموعه داده‌های معیار برای ارزیابی سیستم‌های طبقه‌بندی هیجان در متن

مجموعه داده	زبان	سال	محتوا	تعداد نمونه	مدل هیجان	برچسب‌گذاری
EmoPars	فارسی	۲۰۲۱	توئیت	۳۰۰۰۰	اکمن	چندبرچسبی
PersianTweets	فارسی	۲۰۲۱	توئیت	۱۱۳۰۰۰	اکمن	تک‌برچسبی
ArmanEmo	فارسی	۲۰۲۲	توئیت و نظر	۷۰۰۰	اکمن	تک‌برچسبی
EmoEvent	انگلیسی	۲۰۲۰	توئیت	۷۳۰۳	اکمن	تک‌برچسبی
ISEAR	انگلیسی	۱۹۹۴	جمله	۷۶۶۵	---	تک‌برچسبی
SemEval 2018	انگلیسی	۲۰۱۸	توئیت	۱۰۹۸۳	پلاچیک	چندبرچسبی
EmoInt	انگلیسی	۲۰۱۷	توئیت	۷۰۰۰	راشل‌جک	تک‌برچسبی

## ۴.۲. مجموعه داده‌های انگلیسی

استفاده شده است. مجموعه داده ISEAR<sup>۷</sup> توسط مرکز ملی رقابت‌های تحقیقاتی سوئیس ارائه شده و حاوی ۷۶۶۵ جمله است که به صورت تک‌برچسبی در انواع هیجان ترس، عصبانیت، شادی، غم، نفرت، شرم و گناه برچسب‌گذاری شده‌اند. در سال ۲۰۱۸، مجموعه داده SemEval 2018 Task 5 Subtask 1 برای مسئله طبقه‌بندی هیجان در توئیت‌های انگلیسی، عربی و اسپانیایی ارائه شده است. بخش انگلیسی این مجموعه داده شامل ۱۰۹۸۳ است. هر توئیت به صورت چند برچسبی با برچسب‌های باینری در یازده حالت هیجانی برچسب‌گذاری شده است، که شامل هشت حالت هیجانی مدل

با توجه به اینکه حجم تحقیقات تشخیص هیجان در متون انگلیسی نسبت به سایر زبان‌ها بیشتر است، مجموعه داده‌های انگلیسی زیادی برای تشخیص هیجان به صورت عمومی در دسترس هستند. در این بخش به تعداد کمی از این مجموعه داده‌ها اشاره خواهد شد [۶]. مجموعه داده EmoEvent در سال ۲۰۲۰ ارائه شده است و حاوی ۷۳۰۳ توئیت انگلیسی است. نمونه‌های این مجموعه داده به صورت تک برچسبی و بر اساس مدل اکمن برچسب‌گذاری شده‌اند. برای برچسب‌گذاری هر توئیت از سه کاربر در سامانه Amazon MTurkers

<sup>7</sup> International Survey on Emotion Antecedents and Reactions

متن ابتدا توسط فرایند پیش پردازش اصلاح شده و به یک متن قابل پردازش تبدیل می شود. در مرحله دوم، از آنجایی که مدل های یادگیری ماشینی نمی توانند به صورت مستقیم با کاراکترها و متن کار کنند و معمولاً فقط ورودی های عددی می پذیرند، متن ورودی تمیز شده به کمک روش های جاسازی متن به یک بردار عددی، یک ماتریس، یا دنباله ای از بردارهای عددی تبدیل می شود. سپس، در سومین مرحله یعنی مرحله یادگیری ویژگی ها، بردارهای عددی نماینده متن وارد یک یا چند لایه خاص شبکه های عمیق می شوند، تا ویژگی های معنایی سطح بالاتری از بردارهای ورودی استخراج شوند. در مرحله چهارم، تصمیم گیری نهایی بر اساس بردار ویژگی معنایی انجام می شود. بخش تصمیم گیری معمولاً شامل یک یا چند لایه تماماً متصل<sup>۹</sup> است که در یادگیری ماشینی سنتی نیز استفاده می شد. همانطور که در شکل (۵) مشاهده می کنید، یک مدل یادگیری عمیق از تعداد زیادی لایه تشکیل می شود که لایه های اولیه نقش استخراج ویژگی و لایه های انتهایی نقش تصمیم گیری خودکار را ایفا می کنند.

#### ۵.۱. جاسازی متن

در روش های یادگیری عمیق، جاسازی متن به معنی تبدیل متن ورودی به یک بردار یا دنباله ای از بردارها است. اگر جاسازی متن با توجه به معنای متن انجام شود، سیستم یادگیری ماشینی به دقت بالایی در طبقه بندی هیجان دست خواهد یافت. جاسازی متن می تواند در واحد کلمه، جمله، یا سند انجام شود. در جاسازی کلمات هر کلمه به یک بردار عددی طول ثابت نگاشت داده می شود. روش های جاسازی معنایی کلمات طوری کلمات را به بردارهای عددی نگاشت می دهند که فاصله بین بردارهای کلماتی که مشابهت معنایی دارند کم و فاصله بردارهای کلماتی که تفاوت معنایی دارند زیاد باشد.

پلاچیک همراه با سه حالت عشق، خوش بینی و بدبینی است. در این مجموعه داده، هر توثیت با مشارکت حداقل هفت نفر برچسب گذاری شده است. مجموعه داده EmoInt<sup>۸</sup> در سال ۲۰۱۷ ارائه شده و حاوی ۷۰۰۰ توثیت انگلیسی است که بر اساس مدل هیجان راشل جک در چهار حالت هیجانی به صورت تک برچسبی برچسب گذاری شده اند.

#### ۵. یادگیری عمیق

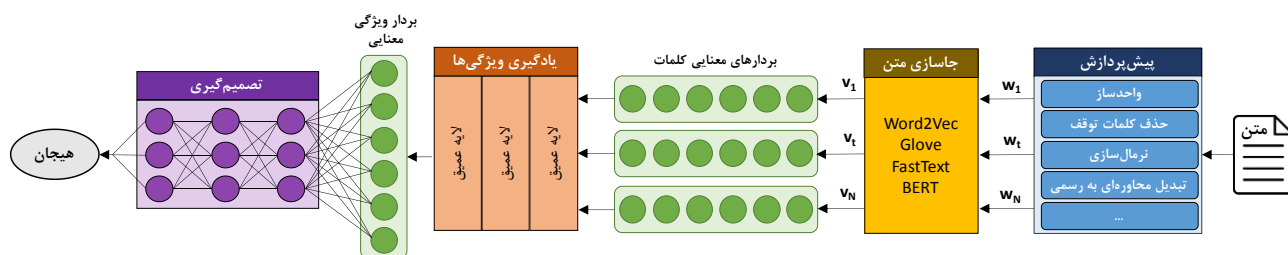
در سال های اخیر، مدل های یادگیری عمیق برای شناسایی هیجان در متن معرفی شده اند، که نسبت به مدل های سنتی یادگیری ماشین از قدرت بیشتری در مدل سازی معنا و توالی کلمات برخوردار هستند [۲]، [۶]، [۸]. مدل های یادگیری عمیق می توانند در یک دنباله از لایه های شبکه عصبی عمیق، جزئیات پنهان و ذاتی معنایی را از متن استخراج کرده و هنگام تصمیم گیری بر اساس معنای متن عمل نمایند [۲۲]، [۲۳]. مدل های عمیق با استفاده از ساختارهای عصبی عمیق و لایه هایی که در خود دارند، به صورت خودکار ویژگی های مطلوب در هر حوزه کاربرد را یاد گرفته و معمولاً باعث دستیابی به نتایج بهتر می شوند.

شکل (۵) فرآیند طبقه بندی هیجان در متن را با رویکرد یادگیری عمیق نشان می دهد. یک سیستم یادگیری عمیق برای طبقه بندی هیجان در متن، از چهار مرحله پیش پردازش، جاسازی متن، یادگیری ویژگی ها با لایه های عمیق و تصمیم گیری تشکیل می شود. متن نوشته شده توسط انسان معمولاً حاوی کلمات محاوره ای، کلمات چند املائی، غلط های املائی، کاراکترهای اضافه و کلمات هجو (کلمات توقف) است. این عوامل نامطلوب در متن ورودی می توانند باعث شوند که عملکرد سیستم در تشخیص صحیح هیجان تضعیف شود. به همین دلیل،

<sup>9</sup> Fully Connected Layer

<sup>8</sup> Emotion Intensity





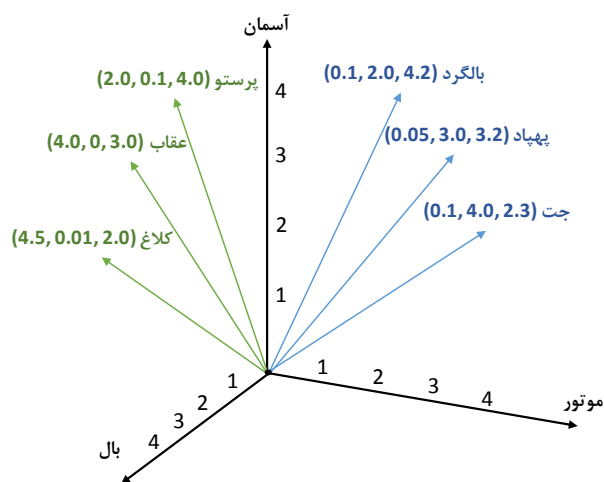
شکل (۵): سیستم تشخیص هیجان در متن مبتنی بر یادگیری عمیق

کلمات قبل یا بعد آن وابسته نبود. به عبارت دیگر، روش‌های جاسازی کلمه، هنگام جاسازی کلمه به زمینه آن توجهی ندارند. در مقابل، مدل‌های زبانی BERT هنگام جاسازی هر کلمه به کلیه کلمات جمله توجه دارند و بردار معنایی تولید شده برای هر کلمه علاوه بر خود آن کلمه به موقعیت آن در جمله و سایر کلمات جمله نیز بستگی دارد [۲۸]. در واقع مدل‌های BERT روش‌های جاسازی توالی هستند و یک توالی از کلمات را به عنوان ورودی دریافت کرده و یک توالی از بردارهای معنایی را به عنوان بازنمایی عددی آن تولید می‌کنند.

## ۵.۲. لایه‌های یادگیری عمیق

در مرحله یادگیری ویژگی‌های عمیق، لایه‌های مختلفی در مدل شبکه عصبی عمیق می‌توانند مورد استفاده قرار بگیرند، که ویژه مدل‌های یادگیری عمیق بوده و در سال‌های اخیر مطرح شده‌اند. هر یک از این لایه‌ها کمک می‌کند تا معانی سطح بالاتری از متن استخراج شده و مدل یادگیری ماشین به فهم معنایی بالاتری از متن برسد. در این راستا، شبکه عصبی بازگشتی<sup>۱۳</sup> برای مدل‌سازی توالی‌های کوتاه مدت، شبکه عصبی حافظه کوتاه مدت-طولانی مدت<sup>۱۴</sup> برای مدل‌سازی توالی‌ها کوتاه مدت و بلند مدت، شبکه عصبی کانولوشنی<sup>۱۵</sup> برای مدل‌سازی زیرعبارت‌ها و کلمات همجوار، لایه توجه<sup>۱۶</sup> برای مدل‌سازی

در شکل (۶) نحوه جاسازی کلمات در یک فضای سه بعدی بصری سازی شده است. در فرایند جاسازی کلمات، بردار متناظر با هر کلمه همواره یک بردار ثابت است و به کلمات قبل یا بعد آن وابسته نیست. برای یادگیری نگاشت بین کلمات با بردارهای معنایی متناظر با آن‌ها و تولید بردار معنایی هر کلمه الگوریتم‌های مختلفی ارائه شده است، که چند نمونه از رایج‌ترین آن‌ها روش‌های جاسازی کلمات Word2Vec<sup>۱۰</sup> [۲۴]، FastText [۲۵] و Glove<sup>۱۱</sup> [۲۶] هستند.



شکل (۶): جاسازی کلمات مبتنی بر شباهت معنایی

یک رویکرد جدید برای جاسازی متن، استفاده از مدل‌های زبانی BERT<sup>۱۲</sup> است [۲۷]. در روش‌های جاسازی کلمه‌ای، بردار معنایی تولید شده در مرحله جاسازی برای هر کلمه، به

<sup>13</sup> Recurrent Neural Network (RNN)

<sup>14</sup> Long Short-term Memory Network (LSTM)

<sup>15</sup> Convolutional Neural Network (CNN)

<sup>16</sup> Self-Attention

<sup>10</sup> Word to Vectors

<sup>11</sup> Global Vectors for Word Representation

<sup>12</sup> Bidirectional Encoder Representations from Transformers (BERT)

## ۶.۱. جاسازی هیجانی

روش‌های جاسازی عمومی که تا اینجا معرفی شدند معمولاً بر روی یک مجموعه از متون عمومی آموزش دیده‌اند و در آن‌ها بردار معنایی هر کلمه معنای عمومی آن کلمه را بازنمایی می‌کند. بعضی محققین در سال‌های اخیر برای افزایش دقت در شناسایی هیجان به جنبه هیجانی هر کلمه نیز علاوه بر جنبه معنایی آن توجه کرده‌اند. در این رویکرد، در مرحله جاسازی متن، از روش‌های جاسازی هیجانی نیز در کنار روش‌های جاسازی معنایی عمومی استفاده می‌شود. به عنوان مثال، باتباتار و همکاران در [۳۰] از روش‌های جاسازی هیجانی<sup>۲۰</sup> همراه با روش‌های جاسازی معنایی **Glove**، **Word2Vec** و **FastText** در یک معماری دو کاناله استفاده کرده‌اند. بردارهای حاصل از روش‌های جاسازی معنایی به یک لایه **BiLSTM** داده شده است، در حالی که بردارهای حاصل از روش‌های جاسازی هیجانی به یک لایه کانولوشن وارد شده است. در نهایت خروجی این دو مسیر با هم الحاق شده و به یک لایه تماماً متصل داده شده است تا طبقه هیجانی متن شناسایی شود. به طور مشابه، چترجی و همکاران در [۳۱] از جاسازی احساسی **SSWE**<sup>۲۱</sup> در کنار جاسازی معنایی **Glove** برای بازنمایی برداری کلمات متن استفاده کرده و بردارهای حاصل از این روش‌های جاسازی را به لایه‌های **LSTM** مجزایی داده‌اند. سپس خروجی نهایی لایه **LSTM** معنایی را به خروجی نهایی لایه **LSTM** احساسی الحاق کرده و به عنوان ورودی به یک لایه تماماً متصل داده‌اند.

زمینه، شبکه عصبی بازگشتی گیتی<sup>۱۷</sup> و بسیاری از مدل‌های عمیق دیگر مطرح شده‌اند. در یک سیستم یادگیری عمیق برای طبقه‌بندی هیجان در متن، این مدل‌های یادگیری عمیق با یکدیگر در لایه‌های متوالی ترکیب می‌شوند تا امکان پردازش قدرتمندتر داده ورودی و دستیابی به دقت بالا فراهم آید.

یکی از رویکردهای جدید در یادگیری عمیق استفاده از لایه توجه است [۲۹]. لایه توجه در متن امکان مدل‌سازی وابسته به زمینه معنای کلمات را فراهم می‌کند، طوری که بردار بازنمایی تولید شده برای هر کلمه علاوه بر خود آن کلمه به کلمات دیگر همان جمله نیز وابسته است. لایه توجه در متن به دو شکل مطرح شده است که شامل لایه توجه چند سر<sup>۱۸</sup> و لایه خودتوجه<sup>۱۹</sup> است. استفاده از خودتوجه نسبت به لایه توجه چندسر رایج‌تر است. لایه خودتوجه شامل وزنهایی است که در فرایند آموزش تنظیم می‌شوند و به مدل کمک می‌کند تا بر وظیفه خاصی تمرکز کند. یک شبکه عمیق می‌تواند شامل چند لایه توجه باشد و هر لایه توجه می‌تواند توجه به بخش خاصی از جمله را یاد بگیرد.

## ۶. مروری بر مطالعات اخیر در متون غیر فارسی

در این بخش تعدادی از سیستم‌های پیشنهادی محققین مختلف را بررسی خواهیم کرد که مبتنی بر یادگیری عمیق بوده و برای شناسایی هیجان در متون غیر فارسی در سال‌های ۲۰۱۹ تا ۲۰۲۲ مطرح شده‌اند. ما تحقیقات پیشین را در این بخش بر اساس رویکردهای نوآورانه آن‌ها برای افزایش دقت در شناسایی هیجان متن بررسی خواهیم کرد. جدول (۲) این تحقیقات را جمع‌بندی کرده است.

<sup>17</sup> Gated Recurrent Unit (GRU)

<sup>18</sup> Multi-head Attention

<sup>19</sup> Self-Attention

<sup>20</sup> Emotional Word Embedding (EWE)

<sup>21</sup> Sentiment-Specific Word Embedding (SSWE)

جدول (۲): آخرین تحقیقات انجام شده در زمینه طبقه‌بندی هیجان در متن برای زبان‌های غیر فارسی

محقق اول	سال	زبان	جاسازی	یادگیری ویژگی‌ها	مجموعه داده	معماری و نوآوری	معیار ارزیابی	دقت
عبدالله [۳۳]	۲۰۱۸	عربی	AraVec	CNN + LSTM	SemEval-2018	ترکیب کانولوشن و LSTM   ترکیب مدل عربی و انگلیسی   استفاده از جاسازی AraVec	Spearman Correlation	۵۶٪
چترجی [۳۱]	۲۰۱۹	انگلیسی	SSWE + Glove	LSTM	۲۲۲۶ مکالمه سه نوبتی در توییتر	ترکیب جاسازی احساسی SSWE با جاسازی GloVe	F1	۷۱٪
باتیاتار [۳۰]	۲۰۱۹	انگلیسی	EWE + Word2Vec/ Glove/ FastText	BiLSTM   CNN	۸۳۲۵۹ متن در ۱۰ مجموعه داده مجزا شامل توییتر و مکالمات	ترکیب جاسازی هیجانی EWE با جاسازی معنایی   ترکیب کانولوشن و BiLSTM	F1	۵۱٪ تا ۹۰٪
سان [۳۴]	۲۰۱۹	انگلیسی	Word2Vec	CNN	یک میلیون پست در Weibo و ۵۰ هزار دیالوگ سریال‌ها	مدل سازی زیرعبارت‌ها با لایه کانولوشن	F1	۹۳٪
بالی [۳۵]	۲۰۱۹	عربی	Word2Vec	CNN	SemEval-2018	مدل سازی زیرعبارت‌ها با لایه کانولوشن	F1	۹۹٪
راغب [۳۶]	۲۰۱۹	انگلیسی	AWD-LSTM	BiLSTM + Self-Attention	۳۸۰۰۰ مکالمه از مجموعه داده SemEval-2019	استفاده از سه لایه خودتوجه همراه با BiLSTM	F1	۷۵٪
شرویاستاوا [۳۷]	۲۰۱۹	انگلیسی	Word2Vec	CNN + MaxPooling + FullyConnected + Self-Attention	۱۳۳۵۴ گفته از نمایش تلویزیونی Charmed	استفاده از لایه خودتوجه و کانولوشن	F1	۷۲٪
ساسیدهار [۳۲]	۲۰۲۰	هندی-انگلیسی	Word2Vec	CNN + BiLSTM	۱۲۰۰۰ متن هندی-انگلیسی مخلوط با کدهای برنامه‌نویسی	یادگیری انتقالی در جاسازی Word2Vec	F1	۸۳٪
آدوما [۳۸]	۲۰۲۰	انگلیسی	BERT	BiLSTM	۷۶۶۶ جمله در مجموعه داده ISEAR	جاسازی با BERT   مدل سازی توالی با BiLSTM	F1	۷۳٪
العمری [۳۹]	۲۰۲۰	انگلیسی	Glove + BERT	BiLSTM	۳۸۲۲۴ دیالوگ در مجموعه داده SemEval-2019	جاسازی با BERT   مدل سازی توالی با BiLSTM	F1	۶۷٪
سونیدانگ کو [۴۰]	۲۰۲۱	اسپانیایی	XML-RoBERTa	CNN + MaxPooling	بیش از ۸ هزار توییت اسپانیایی از مجموعه داده EmoEval-2021	جاسازی با XLM-RoBERTa   مدل سازی زیرعبارت‌ها با کانولوشن	F1	۵۵٪
داس [۴۱]	۲۰۲۱	بنگالی	BERT/Bangla-BERT/XML-R	----	۶ هزار متن بنگالی از منابع مختلف آنلاین و آفلاین	جاسازی با BERT, Bangla-BERT و XLM-R	F1	۶۹٪
دیراج [۴۲]	۲۰۲۱	انگلیسی	Glove	BiLSTM + Multi-head Attention + CNN + MaxPooling	متون مربوط به سوالات کاربران در مورد سلامت روان شامل ۲۰۸۶ سوال از سایت Webmd و ۵۳۲۸ سوال از سایت Healthtap	مدل سازی وابسته به زمینه با Multi-head Attention   ترکیب BiLSTM و کانولوشن   مدل سازی زیرعبارت‌ها با لایه کانولوشن	F1	۷۸٪ تا ۸۹٪
فو [۴۳]	۲۰۲۱	انگلیسی	SSOA + Deep Averaging Network (DAN)	CNN	۵۵۳۱ گفته از مجموعه داده IEMOCAP	استفاده از واژه‌نامه احساسی ANEW برای محاسبه هیجان	F1	۶۹٪
لین [۴۴]	۲۰۲۱	انگلیسی	Word2Vec	BiLSTM + Self-Attention	۱۰ هزار نظر کاربران در اپلیکیشن سفر Low Carbon	ترکیب خودتوجه و BiLSTM	F1	۹۷٪
لو [۴۵]	۲۰۲۱	اسپانیایی	BETO	----	۸۲۲۳ توییت از مجموعه داده TASS-2020	افزایش داده‌های متنی اسپانیایی با ترجمه برگشتی	Accuracy	۷۳٪
آبونیزو [۴۶]	۲۰۲۱	انگلیسی	BERT/ERNIE	CNN/GRU/LSTM	شش مجموعه داده شامل Financial.JEMOCAP.Phrase Bank.Amazon.Ethos.SEMAINE.Yelp	افزایش داده‌های متنی با روش‌های داده‌افزایی ساده، ترجمه برگشتی، BART و PREDATOR	F1	۹۳٪
لو [۴۷]	۲۰۲۱	انگلیسی	Glove	CNN/LSTM	۹ هزار نظر در مورد فیلم‌های سینمایی در مجموعه Stanford Sentiment Treebank. ۲۱ هزار توییت از مجموعه Hate Speech	افزایش داده‌های متنی با مدل SeqGAN به عنوان شبکه مولد تخصصی	Accuracy	۸۲٪
کومار [۴۸]	۲۰۲۲	انگلیسی	BERT	CNN   BiLSTM	چهار مجموعه داده ISEAR.Aman.EmotionLines.AffectiveText	جاسازی با BERT   معماری دو کاناله با کانولوشن و BiLSTM	F1	۷۳٪ تا ۸۳٪
مجید [۴۹]	۲۰۲۲	اردو	Trained Word2Vec	CNN   BiLSTM + Self-Attention	۱۸ هزار جمله به زبان اردو از منابع مختلف	استفاده از توجه بعد از BiLSTM   دو کانال با کانولوشن و BiLSTM	F1	۸۲٪
ژو [۵۰]	۲۰۲۲	انگلیسی، چینی، اسپانیایی، هندی	FastText/BERT	CNN/LSTM	۳۵۳۰ جمله چینی-انگلیسی، ۲۸۸۳ توییت انگلیسی-اسپانیایی، ۳۸۷۹ جمله هندی-انگلیسی که هر سه مجموعه با کدهای برنامه‌نویسی نیز مخلوط شده‌اند	استفاده از دو رمزگذار تک زبانه در دو کانال موازی و یک رمزگذار دو زبانه   آموزش رمزگذارها با یادگیری تخصصی   رمزگذارهای مبتنی بر لایه توجه چندس	F1	۶۵٪

## ۶,۲. یادگیری انتقالی و تنظیم دقیق

در یادگیری انتقالی، استفاده از مدل‌های عمیق از پیش آموزش دیده می‌تواند به ما کمک کند تا با مجموعه محدودی از داده‌های آموزشی نیز به دقت مناسبی در مسائل طبقه‌بندی برسیم [۸]. بعضی محققین از یادگیری انتقالی و تنظیم دقیق برای دستیابی به دقت بالا در شناسایی هیجان متن استفاده کرده‌اند. به عنوان مثال، ساسیدهار و همکاران در [۳۲] از یک روش جاسازی از پیش آموزش دیده دو زبانه هندی-انگلیسی برای بازنمایی عددی و از لایه‌های کانولوشن و BiLSTM برای طبقه‌بندی هیجان در متون هندی-انگلیسی حاوی کدهای برنامه‌نویسی استفاده کرده‌اند. در روش پیشنهادی ایشان، روش جاسازی از پیش آموزش دیده قبلاً روی مجموعه‌ای از ۲۵۰ هزار توئیت هندی-انگلیسی آموزش دیده است، اما با متون حاوی کدهای برنامه‌نویسی آشنا نیست. به همین دلیل، ساسیدهار و همکارانش این روش جاسازی را با مجموعه متون جدید به‌روزرسانی کرده‌اند تا عملکرد کل سیستم در شناسایی هیجان در متون هندی-انگلیسی حاوی کدهای برنامه‌نویسی بهبود یابد. کراتر والد و همکاران در [۸] رویکرد `sent2affect` را مطرح کرده‌اند، که در آن مدل یادگیری عمیقی که قبلاً برای تشخیص قطبیت نظرات در متن آموزش دیده است، با یادگیری انتقالی و تنظیم دقیق در کاربرد جدید یعنی شناسایی هیجان در متن به کار بسته شده است. در روش پیشنهادی ایشان ویژگی‌های استخراج شده با جاسازی کلمات با ویژگی‌های حاصل از یادگیری انتقالی `sent2affect` ترکیب شده و سپس به یک لایه شبکه عصبی بازگشتی، لایه چشم‌پوشی تصادفی<sup>۲۲</sup> و یک لایه تماماً متصل داده شده است. در یادگیری انتقالی `sent2affect` مدل عمیقی که قبلاً برای طبقه‌بندی قطبیت نظرات آموزش داده شده است به کاربرد جدید کپی شده و وزن‌های آخرین لایه آن بر روی داده‌های تشخیص هیجان در متن به روزرسانی شده است. این رویکرد

باعث کاهش زمان لازم برای آموزش شده و دقت را نیز افزایش داده است.

## ۶,۳. مدل‌های دو زبانه

بعضی محققین برای دستیابی به دقت بالا در شناسایی هیجان متن به خصوص در زبان‌هایی با منابع آموزشی محدود، از ترکیب مدل‌های انگلیسی پیش آموزش دیده با مدل زبان هدف استفاده کرده‌اند. در این رویکرد، مدل‌های انگلیسی از قبل روی مجموعه متون انگلیسی با حجم زیاد آموزش دیده‌اند و انتظار می‌رود با ترکیب با مدل زبان هدف بتوانند به طبقه‌بندی دقیق‌تر کمک کنند. به عنوان مثال، عبدالله و همکاران در [۳۳] یک مدل ترکیبی را با رویکرد دوزبانه برای شناسایی هیجان در توئیت‌های زبان عربی ارائه کرده‌اند. مدل عربی از لایه‌های LSTM، کانولوشن<sup>۲۳</sup> و تماماً متصل تشکیل شده است؛ در حالی که مدل انگلیسی از تعداد زیادی طبقه‌بند پیش‌آموزش دیده انگلیسی تشکیل شده است. خروجی این دو مدل توسط یک لایه تماماً متصل با یکدیگر ترکیب شده است.

## ۶,۴. استفاده از واژه‌نامه

یکی از رویکردهای سنتی برای تشخیص هیجان در متن، استفاده از واژه‌نامه هیجانی است، که در آن به ازای هر کلمه انواع هیجان موجود در آن کلمه برچسب‌گذاری شده‌اند. از جمله این واژه‌نامه‌های هیجانی می‌توان به واژه‌نامه NRC<sup>۲۴</sup>، `DepecheMood`، `WordNet-Affect`، `EmoSenticNet` و `SentiWordNet` اشاره کرد. در رویکردهای عمیق نیز بعضی محققین از واژه‌نامه هیجانی برای کمک به سیستم تشخیص هیجان استفاده کرده‌اند. به عنوان مثال، فو و همکاران در [۴۳] از واژه‌نامه احساسی ANEW<sup>۲۵</sup> [۵۱] برای ارزیابی شدت هیجان‌های شادی، عصبانیت، غم و خنثی در متن استفاده

<sup>23</sup> CNN Layer

<sup>24</sup> National Research Council Canada (NRC)

<sup>25</sup> Affective Norms for English Words (ANEW)

<sup>22</sup> Drop-out

وابسته به زمینه معنای کلمات در تشخیص هیجان روی متون سلامت روان استفاده کرده‌اند. در روش پیشنهادی ایشان، متن ورودی کاربر بعد از جاسازی وارد یک لایه BiLSTM شده تا توالی کلمات مدل‌سازی شود. سپس از یک لایه توجه چند سر برای مدل‌سازی زمینه کلمات، یک لایه کانولوشن برای شناسایی الگوهای محلی، یک لایه ادغام حداکثری و در نهایت یک لایه تماماً متصل استفاده شده است. مجید و همکاران در [۴۹] از لایه توجه بعد از یک لایه BiLSTM در یک معماری دو کاناله برای شناسایی هیجان در متون زبان اردو استفاده کرده‌اند. در روش پیشنهادی ایشان در کانال اول یک لایه کانولوشن و در کانال دوم یک لایه BiLSTM قرار گرفته است.

#### ۶.۶. جاسازی متن با کمک BERT

محققین زیادی از نسخه‌های مختلف BERT برای جاسازی متن در تشخیص هیجان استفاده کرده‌اند. به عنوان مثال، داس و همکاران در [۴۱] از ترنسفورمرهای BERT، Bangla-BERT و XLM-R<sup>۲۷</sup> برای شناسایی هیجان در متون بنگالی استفاده کرده‌اند، که مدل XLM-R عملکرد بهتری داشته است. العمری و همکاران در [۳۹] از BERT به منظور جاسازی متن و از لایه‌های BiLSTM و تماماً متصل به منظور یادگیری ویژگی برای شناسایی هیجان در دیالوگ‌های زبان انگلیسی استفاده کرده‌اند. سوئیدانگ کو و همکاران در [۴۰] از مدل XLM-RoBERTa به عنوان روش جاسازی و از لایه کانولوشن به عنوان یادگیری ویژگی استفاده کرده‌اند تا هیجان را در متون زبان اسپانیایی تشخیص دهند. آدوما و همکاران در [۳۸] از مدل BERT به عنوان روش جاسازی و از BiLSTM به عنوان یادگیری ویژگی برای شناسایی هیجان در متون مجموعه داده ISEAR<sup>۲۸</sup> استفاده کرده‌اند. کومار و رامان در [۴۸] از مدل

کرده‌اند. سپس، در بخش جاسازی متن، ضرایب وزنی انواع هیجان با نمایش برداری متن ورودی ترکیب شده و به یک لایه کانولوشن و سپس سه لایه تماماً متصل وارد شده‌اند تا هیجان متن شناسایی شود.

#### ۶.۵. مکانیزم توجه

بعضی مطالعات اخیر از مکانیزم توجه در شناسایی هیجان متن استفاده کرده‌اند. به عنوان مثال، راغب و همکاران در [۳۶] از یادگیری انتقالی در بخش جاسازی و از لایه توجه در بخش یادگیری ویژگی‌ها استفاده کرده‌اند. روش جاسازی که قبلاً بر روی مجموعه بزرگی از متون به نام Wikitext-103 آموزش دیده به کاربرد جدید منتقل شده است. نتایج جاسازی بعد از عبور از سه لایه BiLSTM وارد سه لایه خودتوجه شده و بعد از عبور از لایه‌های تماماً متصل طبقه نهایی متن شناسایی شده است. شریواستاوا و همکاران در [۳۷] از Word2Vec به عنوان روش جاسازی و از لایه خودتوجه در بخش یادگیری ویژگی‌ها، برای تشخیص هیجان در متون دیالوگ‌های فیلم‌های سینمایی استفاده کرده‌اند. در روش پیشنهادی ایشان، ابتدا متن به کمک روش جاسازی کلمات Word2Vec به دنباله‌ای از بردارها تبدیل شده است. سپس این دنباله به عنوان یک ماتریس در نظر گرفته شده و توسط لایه‌های کانولوشن دو بعدی پردازش شده است. خروجی کانولوشن‌ها پس از ادغام حداکثری<sup>۲۶</sup> وارد یک لایه تماماً متصل و سپس یک لایه خودتوجه شده و در نهایت طبقه نهایی با کمک تابع فعالیت Softmax مشخص شده است. لین و همکاران در [۴۴] نیز از خودتوجه در بخش یادگیری ویژگی‌ها و از جاسازی Word2Vec در بخش جاسازی متن، برای شناسایی هیجان در متون نظرات کاربران اپلیکیشن مسافرتی Low-Carbon استفاده کرده‌اند. دیراج و راماکریشنادو [۴۲] از مکانیزم توجه چند سر برای مدل‌سازی

<sup>27</sup> Cross-lingual language model (XLM)

<sup>28</sup> International Survey on Emotion Antecedents and Reactions (ISEAR)

<sup>26</sup> Max-pooling

BERT به عنوان روش جاسازی و در لایه‌های بعدی از یک معماری دو کاناله بهره برده‌اند، که کانال اول حاوی لایه کانولوشن و کانال دوم حاوی BiLSTM است. در روش پیشنهادی ایشان از لایه کانولوشن برای مدل‌سازی ویژگی‌های محلی متن و از BiLSTM برای مدل‌سازی توالی کلمات استفاده شده است. ژو و همکاران در [۵۰] از دو رمزگذار تک زبانه در دو کانال موازی و یک رمزگذار دو زبانه برای تحلیل هیجان در متون دو زبانه در زبان‌های انگلیسی، چینی، اسپانیایی و هندی استفاده کرده‌اند. این رمزگذارها به صورت تخصصی آموزش دیده‌اند و عملکرد سیستم روی متون دو زبانه‌ای ارزیابی شده که با کدهای برنامه‌نویسی مخلوط شده‌اند.

۶.۷. مدل‌سازی زیرعبارت‌ها

برای دستیابی به یک فهم معنایی بالاتر نسبت به جاسازی کلمات، بعضی محققین به مدل‌سازی زیرعبارت‌ها و خصوصیات محلی متن توجه کرده‌اند، به این معنی که معنای توالی‌های کوتاه از دو یا چند کلمه نیز در مدل‌سازی لحاظ شود. برای این منظور یک رویکرد رایج استفاده از شبکه عصبی کانولوشن یک بعدی یا دو بعدی روی نمایش عددی کلمات است. به عنوان مثال، سوئیدانگ کو و همکاران در [۴۰] از شبکه عصبی کانولوشنی یک بعدی برای استخراج ویژگی‌های محلی از متن استفاده کرده‌اند. ابتدا متن ورودی توسط مدل XLM-ROBERTa جاسازی شده است، سپس نتیجه توسط لایه کانولوشن یک بعدی پردازش شده و بعد از عبور از لایه ادغام حداکثری، به یک لایه تماماً متصل داده شده است تا طبقه‌بندی هیجان شناسایی شود. عبدالله و همکاران [۳۳] در زیر مدل عربی سیستم پیشنهادی خود، از AraVec برای جاسازی متن عربی، سپس از لایه کانولوشن یک بعدی برای مدل‌سازی زیرعبارت‌ها و بعد از آن از یک لایه LSTM برای مدل‌سازی توالی زیرعبارت‌ها استفاده کرده‌اند. دیراج و راماکریشنادو در [۴۲] از لایه کانولوشن

دو طرفه بعد از یک لایه BiLSTM و توجه چند سر برای مدل‌سازی زیرعبارت‌ها استفاده کرده‌اند. در روش پیشنهادی ایشان، متن توسط روش جاسازی GloVe به نمایش عددی تبدیل شده است، که هم ویژگی‌های محلی و هم ویژگی‌های سراسری را هنگام جاسازی در نظر می‌گیرد. سپس، دنباله بردارهای عددی توسط یک لایه Bi-LSTM پردازش شده است، تا توالی‌های رایج شناسایی شوند. در لایه بعدی یک لایه توجه چند سر در نظر گرفته شده که ارتباط‌های معنایی سطح بالا بین کلمات را مدل‌سازی می‌کند. سپس زیرعبارت‌ها توسط یک لایه کانولوشن شناسایی شده‌اند و نتیجه بعد از ادغام حداکثری به یک لایه تماماً متصل جهت تصمیم‌گیری نهایی داده شده است. سان و همکاران در [۳۴] از لایه کانولوشن برای مدل‌سازی زیرعبارت‌ها بعد از روش جاسازی Word2Vec استفاده کرده‌اند. روش پیشنهادی ایشان در مراحل بعدی از روش زنجیره مارکوف مونت کارلو<sup>۲۹</sup> برای مدل‌سازی توالی زمانی انواع هیجان کاربر استفاده کرده تا تغییرات هیجان کاربر را در طول زمان شناسایی کند. بالی و غنیم نیز در [۳۵] از لایه‌های کانولوشن یک بعدی برای مدل‌سازی زیرعبارت‌ها در متون عربی استفاده کرده‌اند.

#### ۶.۸. افزایش داده‌های متنی

بعضی محققین برای دستیابی به دقت بالا و اجتناب از بیش‌برازش در طبقه‌بندی هیجان متن، از افزایش داده‌های متنی استفاده کرده‌اند. به عنوان مثال، لو در [۴۵] از روش ترجمه برگشتی برای جلوگیری از بیش‌برازش در شناسایی هیجان در متون زبان اسپانیایی استفاده کرده است. در این تکنیک، متون آموزشی اسپانیایی به کمک یک مترجم خودکار به زبان میانی انگلیسی ترجمه شده و مجدداً به زبان اسپانیایی برگشت داده می‌شوند و متن جدیدی با همان برچسب قبلی به دست می‌آید.

<sup>29</sup> Markov Chain Monte Carlo (MCMC)

علامت تعجب، فراوانی واژگان هیجانی، نقش واژگان در جمله و وجود تشدید کننده از متن ورودی استخراج شده‌اند. سپس، این ویژگی‌ها به یک طبقه‌بند ماشین بردار پشتیبان<sup>۳۳</sup> داده شده‌اند. روش پیشنهادی خسروی روی یک مجموعه داده از عناوین خبری روزنامه‌های کثیرالانتشار فارسی با ۱۰۰ نمونه به نرخ دقت طبقه‌بندی برابر با ۸۴٪ و نرخ میانگین F1 برابر با ۳۴٪ در هشت کلاس هیجانی دست یافته است. از جمله محدودیت‌های پژوهش مذکور می‌توان به مجموعه داده کوچک و عدم استفاده از روش‌های یادگیری عمیق اشاره کرد.

صادقی و همکاران در سال ۲۰۲۱ در [۵۷] یک روش ترکیبی برای شناسایی هیجان در متن فارسی ارائه کرده‌اند که از ترکیب ویژگی‌های حاصل از یادگیری عمیق با ویژگی‌های سنتی برای شناسایی دقیق‌تر هیجان در متون فارسی استفاده می‌کند. در بخش ویژگی‌های سنتی، از شمارش تعداد کلیدواژه‌های هیجانی، تعداد گروه‌های نحوی هیجانی که در واقع کلیدواژه‌هایی هیجانی با نقش خاص در جمله هستند و تعداد ساختارهای هیجانی که در حقیقت عبارت‌ها یا کنایه‌های هیجانی هستند، برای استخراج ویژگی استفاده شده است. در بخش ویژگی‌های عمیق از یک مدل یادگیری عمیق شامل جاسازی کلمات Word2Vec، یک لایه شبکه عصبی بازگشتی گیتی، یک لایه چشم‌پوشی تصادفی و سپس یک لایه تماماً متصل با پنج نورون استفاده شده است. در نهایت، در بخش ترکیب ویژگی‌های سنتی با بردار پنج عنصری حاصل از یادگیری عمیق الحاق شده و به عنوان بردار ویژگی نهایی به طبقه‌بند سنتی ماشین بردار پشتیبان داده شده است. برای ارزیابی روش پیشنهادی، ۲۳ هزار جمله از مجموعه داده بیژن خان که حاوی متون خبری است انتخاب شده و به صورت دستی در پنج کلاس ترس، عصبانیت، شادی، غم و تعجب برچسب‌گذاری هیجانی شده‌اند.

آبونیزو و همکاران در [۴۶] از روش‌های افزایش داده ساده<sup>۳۰</sup>، ترجمه برگشتی، BART<sup>۳۱</sup> و PREDATOR برای افزایش داده‌های متنی در طبقه‌بندی احساسات استفاده کرده‌اند. روش افزایش داده ساده شامل چهار تغییر ساده جایگزینی با کلمه هم معنی، درج تصادفی، حذف تصادفی و جابجایی تصادفی کلمات روی متن است. روش BART یک مدل توالی به توالی مبتنی بر خودرمزگذار و نوع خاصی از BERT برای افزایش داده‌های متنی است [۵۲]. یک مدل قدرتمند دیگر برای افزایش داده‌های متنی، مدل BERT شرطی یعنی C-BERT است، که از برچسب جملات آموزشی نیز حین افزایش آن‌ها استفاده می‌کند [۵۳]. لو و همکاران در [۴۷] از شبکه مولد تخصصی SeqGAN<sup>۳۲</sup> برای افزایش داده‌های متنی در تحلیل احساسات استفاده کرده‌اند. مرور جذابی بر روش‌های افزایش داده‌های متنی در [۵۴]، [۵۵] آمده است.

## ۷. مروری بر مطالعات اخیر در زبان فارسی

در این بخش به مرور تحقیقاتی پرداخته خواهد شد که در سال‌های اخیر برای شناسایی هیجان متن در زبان فارسی انجام شده‌اند. خلاصه این تحقیقات در جدول (۳) آمده است.

یکی از اولین پژوهش‌ها در زمینه تشخیص هیجان متن فارسی، پژوهش خسروی و همکاران در سال ۲۰۱۹ در [۵۶] است که در آن از یک روش یادگیری ماشین سنتی برای تشخیص هیجان در اخبار استفاده شده است. در روش پیشنهادی ایشان، از واژه‌نامه NRC برای تشخیص و شمارش واژگان هیجانی استفاده شده است. ویژگی‌های زمینه‌ای متن با شمارش کلمات هیجانی، تعداد و نسبت واژگان مثبت و منفی، تعداد

<sup>30</sup> Easy Data Augmentation (EDA)

<sup>31</sup> Bidirectional Auto-encoder Representations from Transformers (BART)

<sup>32</sup> Sequence Generative Adversarial Nets (SeqGAN)

<sup>33</sup> Support Vector Machines (SVM)

جدول (۳): آخرین تحقیقات انجام شده در زمینه طبقه‌بندی هیجان در متن برای متون زبان فارسی

محقق اول	سال	جاسازی	یادگیری و ویژگی‌ها	تصمیم‌گیری	مجموعه داده	معماری و نوآوری	معیار ارزیابی	دقت
خسروی [۵۶]	۲۰۱۹	NRC Lexicon	-----	SVM	۱۰۰ عنوان خبری فارسی از روزنامه‌های کثیرالانتشار	یادگیری ماشین سنتی   استفاده از واژه‌نامه NRC   استفاده از ویژگی‌های زمینه‌ای متن	F1	۳۴٪
صادقی [۵۷]	۲۰۲۱	Word2Vec	GRU	SVM	۲۳۰۰۰ خبر فارسی از پیکره بیژن‌خان	ترکیب یادگیری عمیق با یادگیری سنتی	F1	۹۳٪
میرزایی [۱۹]	۲۰۲۲	ParsBERT/ XLM-R	----	Fully Connected	مجموعه داده ArmanEmo حاوی ۷۰۰۰ تونیت فارسی	استفاده از مدل‌های جاسازی پیش‌آموزش دیده	F1	۷۵٪
عباس‌کوهی [۵۸]	۲۰۲۲	ParsBERT/ XLM-R	----	Fully Connected	بخشی از مجموعه داده EmoPars حاوی ۱۵۰۰۰ تونیت فارسی، مجموعه داده ArmanEmo حاوی ۷۰۰۰ تونیت فارسی	متوازن‌سازی مجموعه داده فارسی   افزایش داده‌ها با کمک تغییرات تصادفی متن   استفاده از مدل‌های جاسازی پیش‌آموزش دیده	F1	۸۱٪
خدائی [۵۹]	۲۰۲۲	XLM-R	----	CatBoost DT	مجموعه داده JAMFA شامل ۲۲۴۱ جمله ادبی فارسی	ترکیب ویژگی‌های عمیق با طبقه‌بندهای سنتی	F1	۷۰٪

متون فارسی استفاده کرده‌اند. در روش پیشنهادی ایشان، برای کمک به مدل یادگیری عمیق، بعضی ویژگی‌های سنتی از جمله انواع ایموجی در متن، نقش کلمات در جمله<sup>۳۴</sup>، کلمات با املای غلط و هشتگ‌ها نیز استفاده شده است. این ویژگی‌های سنتی به صورت یک توصیف متنی به انتهای متن ورودی الحاق شده‌اند و به عنوان ورودی نهایی به مدل یادگیری عمیق داده شده‌اند. در بخش یادگیری عمیق از مدل‌های پیش‌آموزش دیده ParsBERT، XLM-Roberta-base و XLM-Roberta-large برای جاسازی متن و سپس یک لایه تماماً متصل برای تصمیم‌گیری استفاده شده است. عباس‌کوهی و همکاران به منظور افزایش تعداد نمونه‌ها از روش‌های افزایش داده‌های متنی استفاده کرده‌اند، که در آن نمونه‌های جدید با تغییرات تصادفی شامل حذف، درج، جایگزینی و جابجایی تصادفی بعضی کلمات متن تولید شده‌اند. استفاده از افزایش داده‌های متنی، تابع هزینه F1-Cross-Entropy، کم‌نمونه‌برداری، وزن‌دهی به کلاس‌ها و استخراج ویژگی‌های سنتی باعث شده است نرخ میانگین F1 برای مدل ParsBERT از ۳۲٪ به ۷۶٪ بر روی مجموعه داده

روش پیشنهادی صادق به نرخ دقت طبقه‌بندی برابر با ۹۷٪ و نرخ F1 برابر با ۹۳٪ رسیده است. ترکیب یادگیری عمیق با یادگیری سنتی در مطالعه ایشان، نرخ دقت طبقه‌بندی را به ترتیب از ۹۴٪ و ۹۲٪ به نرخ دقت برابر با ۹۷٪ برای مدل ترکیبی افزایش داده است.

میرزایی و همکاران در سال ۲۰۲۲ در [۱۹] مجموعه داده ArmanEmo را معرفی کرده و تعدادی از روش‌های جاسازی پیش‌آموزش دیده را با یادگیری انتقالی بر روی آن آزموده‌اند. در آزمایش‌های انجام شده نرخ میانگین F1 برای کلاس‌های مختلف به ترتیب برای جاسازی FastText مقدار ۴۷٪، برای جاسازی ParsBERT مقدار ۶۵٪، برای جاسازی XLM-Roberta-base مقدار ۶۹٪ و برای جاسازی XLM-Roberta-large مقدار ۷۵٪ شده است. پژوهش میرزایی و همکاران نشان داد که یادگیری انتقالی و استفاده از مدل‌های جاسازی پیش‌آموزش دیده می‌تواند به دقت قابل قبولی در شناسایی هیجان در متن فارسی منجر شود.

عباس‌کوهی و همکاران در سال ۲۰۲۲ در [۵۸] از افزایش داده‌های متنی برای دستیابی به دقت بالا در شناسایی هیجان

<sup>34</sup> Part of Speech (POS) Tagging



برای کلاس‌های مختلف به راهکارهای خاص نیاز دارد. بعضی از این راهکارها که در سال‌های اخیر ارائه شده‌اند شامل جاسازی هیجانی، جاسازی توالی، مدل‌های ترکیبی، یادگیری انتقالی، مکانیزم توجه، جاسازی با BERT، مدل‌سازی زیرعبارت‌ها، مدل‌سازی زمینه و افزایش داده‌های متنی هستند.

با وجود این که تحقیقات قابل توجهی برای تشخیص هیجان در متون انگلیسی و زبان‌های غیر فارسی انجام شده است، اما مسئله تشخیص هیجان در متن هنوز یک مسئله چالش برانگیز است. از نظر ما بخشی از این چالش به ماهیت روانشناختی هیجان در متن مربوط است. هیجان از نظر روانشناختی مفهومی چند بعدی و پیچیده است و ما انسان‌ها در هر لحظه ممکن است ترکیبی از هیجان‌ها را با هم تجربه کنیم. از طرف دیگر، افراد مختلف ممکن است برداشت‌های متفاوتی از یک متن داشته باشند و معنای یک متن یکتا نیست. این عوامل باعث می‌شوند برچسب‌گذاری هر نمونه را نتوان با قطعیت کامل انجام داد. به عنوان مثال، یک انسان هنگام خواندن جمله "چرا به من پیام نمی‌دهی!" ممکن است آن را به عنوان عصبانیت یا غم تعبیر کند و همین ابهام برای ماشین نیز وجود دارد. عدم قطعیت در برچسب‌گذاری به خصوص برای مجموعه‌داده‌های فارسی که عموماً توسط تعداد کمی از افراد برچسب‌گذاری شده‌اند، باعث می‌شود الگوریتم‌های یادگیری ماشین نتوانند به دقت بسیار بالا در تشخیص هیجان دست یابند.

تشخیص هیجان در متن فارسی به مراتب از متن انگلیسی دشوارتر است. مجموعه‌داده‌های تشخیص هیجان در متن فارسی از نظر صحت برچسب‌گذاری، تعداد نمونه‌ها و رسمی یا محاوره‌ای بودن متن مشکلاتی دارند. این موضوع باعث می‌شود آموزش و ارزیابی دقیق مدل‌ها در زبان فارسی با مجموعه‌داده‌های فعلی دشوار باشد. بنابراین، به مجموعه‌داده‌های استاندارد با تعداد نمونه زیاد در زبان فارسی نیاز است که

EmoPars افزایش یابد. همچنین، مدل جاسازی XLM-Roberta به نرخ میانگین F1 برابر با ۰.۸۱ دست یافته است. روش پیشنهادی عباس‌کوهی بر روی مجموعه‌داده ArmanEmo نیز به نرخ میانگین F1 برابر با ۰.۸۱ دست یافته است. البته نتایج پژوهش عباس‌کوهی روی مجموعه‌داده ArmanEmo با نتایج پژوهش میرزایی روی همین مجموعه‌داده مستقیماً قابل مقایسه نیست، زیرا نسبت اندازه مجموعه آزمون به آموزش در آزمایش‌های این دو محقق یکسان نبوده است.

خدائی و همکاران در سال ۲۰۲۲ در [۵۹] از ترکیب ویژگی‌های حاصل از یادگیری عمیق با روش یادگیری سنتی CatBoost DT برای دستیابی به دقت بالا در طبقه‌بندی هیجان متون ادبی فارسی استفاده کرده‌اند. در روش پیشنهادی ایشان متن ورودی به کمک جاسازی XLM-Roberta به یک دنباله ۱۰۰ تایی از بردارهایی به طول ۱۰۲۴ تبدیل شده است. سپس این ماتریس 100x1024 عنصری به عنوان ورودی به طبقه‌بند درختی CatBoost DT داده شده است. خدائی و همکاران روش پیشنهادی خود را بر روی مجموعه‌داده JAMFA ارزیابی کرده‌اند، که حاوی ۲۲۴۱ جمله ادبی است. روش پیشنهادی خدائی و همکاران در طبقه‌بندی چهار کلاسه هیجان متون ادبی فارسی به نرخ میانگین F1 برابر با ۰.۷۰ دست یافته است. استفاده از جاسازی XLM-Roberta و طبقه‌بند درختی CatBoost DT در روش ایشان، نسبت به مدل پایه که از جاسازی FastText و لایه BiLSTM استفاده می‌کند، نرخ دقت طبقه‌بندی را به میزان ۰.۲۰ بهبود داده است.

## ۸. چالش‌ها و فرصت‌های پژوهشی

مطالعه تحقیقات مختلف نشان داد که مسئله تشخیص هیجان در متن یک مسئله نسبتاً دشوار است که دستیابی به دقت بالا در آن

مجرمین به کمک متون گزارش‌های روزانه ایشان، بررسی افسردگی در بیماران بر اساس یادداشت‌های روزانه ایشان، شناسایی پیام‌های تهدیدآمیز در شبکه‌های اجتماعی برای پیشگیری از جرم، بررسی هیجان در نظرات کاربران در خصوص فیلم‌ها و سریال‌ها، شناسایی پیام‌های توهین‌آمیز در مکالمات و شناسایی هیجان در زیرنویس فیلم‌ها و آهنگ‌ها نمونه‌هایی از کاربردهایی هستند که هنوز در زبان فارسی به آن‌ها پرداخته نشده است.

## ۹. نتیجه

در این مطالعه آخرین تحقیقات انجام شده برای طبقه‌بندی هیجان در متن با کمک رویکردهای یادگیری عمیق بررسی شدند. مطالعه کارهای مختلف نشان داد که مسئله طبقه‌بندی هیجان در متن یک مسئله نسبتاً دشوار است که دستیابی به دقت بالا در آن برای کلاس‌های مختلف به راهکارهای ویژه‌ای نیاز دارد. بعضی از این راهکارها که در سال‌های اخیر ارائه شده‌اند شامل جاسازی هیجانی، جاسازی توالی، مدل‌های ترکیبی، یادگیری انتقالی، مکانیزم توجه، جاسازی با BERT، مدل‌سازی زیرعبارت‌ها، مدل‌سازی زمینه و افزایش داده‌های متنی هستند. با وجود تحقیقات انجام شده، هنوز دستیابی به دقت بالا در طبقه‌بندی متون فارسی، ارائه مجموعه‌داده‌های استاندارد و مناسب برای زبان فارسی، مقابله با کمبود داده در زبان فارسی و به کار بستن تشخیص هیجان در کاربردهای جدید از جمله حوزه‌های پیشنهادی برای تحقیقات آینده هستند.

## مراجع

- [1] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, Sep. 2021, doi: 10.1016/j.jii.2021.100224.

متون آن از نظر هیجانی توسط تعداد قابل توجهی کاربر برچسب گذاری شده باشند. همچنین، ارزیابی همزمان مدل‌های پیشنهادی بر روی مجموعه‌داده‌های انگلیسی و ترجمه ماشینی مجموعه‌داده‌های انگلیسی به فارسی می‌تواند دو راهکار موقت برای حل این مشکل باشد.

در یک سیستم یادگیری عمیق برای تشخیص هیجان در متن، بازنمایی عددی و طبقه‌بندی دو مرحله مهم را تشکیل می‌دهند. در مرحله بازنمایی عددی، متن باید طوری به یک بردار عددی یا دنباله‌ای از بردارهای عددی تبدیل شود، که معنای متن در نمایش عددی تولید شده بازنمایی شود. برای این منظور باید به اطلاعات زمینه‌ای متن و عبارت اطراف هر کلمه در بازنمایی آن توجه شود. مدل‌های جاسازی توالی، جمله، کلمه و سند در سال‌های اخیر برای حل این مشکل مطرح شده‌اند که بخش قابل توجهی از آن‌ها بر معماری ترنسفورمر مبتنی هستند. با این وجود معماری ترنسفورمر نیز مشکلاتی دارد، که از جمله آن‌ها می‌توان به عدم پاسخگویی روی کلمات جدید، پیچیدگی زیاد و بیش‌برازش روی مجموعه‌داده‌های کوچک اشاره کرد. استفاده از مدل‌های جاسازی هوشمندانه‌تر و مدل‌های ترنسفورمر جدید می‌تواند به افزایش دقت تشخیص هیجان در متون فارسی منجر شود. همچنین، ترکیب این مدل‌ها با مدل‌های سنتی، لایه توجه، مدل‌های شبکه عصبی و سایر مدل‌های طبقه‌بندی به صورت طبقه‌بندهای گروهی می‌تواند محققین را به سیستم‌های با کیفیت‌تری برساند.

بررسی ما در این تحقیق نشان داد که تاکنون پژوهش‌های بسیار محدودی در زمینه تشخیص هیجان در متن فارسی انجام شده است. این در حالی است که تشخیص هیجان در متن کاربردهای زیادی دارد. طراحی و ارزیابی سیستم‌های یادگیری عمیق برای تشخیص هیجان در کاربردهای جدید می‌تواند موضوع کارهای آینده محققین این حوزه باشد. نظارت بر افکار

- [11] P. Ekman and D. Cordaro, "What is meant by calling emotions basic," *Emotion Review*, vol. 3, no. 4, pp. 364–370, Oct. 2011, doi: 10.1177/1754073911410740.
- [12] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3–4, Art. no. 3–4, May 1992, doi: 10.1080/02699939208411068.
- [13] R. E. Jack, W. Sun, I. Delis, O. G. B. Garrod, and P. G. Schyns, "Four not six: Revealing culturally common facial expressions of emotion," *J Exp Psychol Gen*, vol. 145, no. 6, pp. 708–730, Jun. 2016, doi: 10.1037/xge0000162.
- [14] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *2021 IEEE World AI IoT Congress (AIoT)*, May 2021, pp. 0255–0261. doi: 10.1109/AIIoT52608.2021.9454192.
- [15] R. Plutchik, "In search of the basic emotions," *PSYCCRITIQUE*, vol. 29, no. 6, pp. 511–513, 1984.
- [16] E. C.-C. Kao, C.-C. Liu, T.-H. Yang, C.-T. Hsieh, and V.-W. Soo, "Towards text-based emotion detection a survey and possible improvements," in *2009 International Conference on Information Management and Engineering*, Apr. 2009, pp. 70–74. doi: 10.1109/ICIME.2009.113.
- [17] N. Sabri, R. Akhavan, and B. Bahrak, "EmoPars: A collection of 30k emotion-annotated Persian social media texts," in *Proceedings of the Student Research Workshop Associated with RANLP 2021*, 2021, pp. 167–173. [Online]. Available: <https://aclanthology.org/2021.ranlp-srw.23/>
- [18] B. Karimi. (2022, Apr. 03). Persian tweets emotional dataset [Online]. Available: <https://www.kaggle.com/behdadkarimi/persian-tweets-emotional-dataset>
- [19] H. Mirzaee, J. Peymanfard, H. H. Moshtaghin, and H. Zeinali, "ArmanEmo: A persian dataset for text-based emotion detection." arXiv, Jul. 24, 2022. doi: 10.48550/arXiv.2207.11808.
- [20] N. Sabri. (2022, Jan. 31). Persian-Emotion-Detection [Online]. Available: <https://github.com/nazaninsbr/Persian-Emotion-Detection>
- [2] D. W. Otter, J. R. Medina, and J. K. Kalita, "A survey of the usages of deep learning for natural language processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 604–624, Feb. 2021, doi: 10.1109/TNNLS.2020.2979670.
- [۳] خسروی آ.، عبدالحسینی ح.، «شخصیت شناسی در شبکه های اجتماعی با استفاده از مدل سازی موضوعی نظرات کاربران»، مجله محاسبات نرم، ۱۴۰۱، doi: 10.22052/scj.2023.243197.1006
- [۴] پورغلامعلی ف.، کاهانی م.، عسگریان ا.، «استفاده از تکنولوژی داده های عظیم در نظرکاوی»، مجله محاسبات نرم، جلد ۹، شماره ۱-پ ۱۷، ص ۲۶-۳۹، ۱۳۹۹، doi: 10.22052/scj.2021.111450
- [5] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, p. 107134, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.
- [6] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," *Engineering Reports*, vol. 2, no. 7, p. e12189, 2020, doi: 10.1002/eng2.12189.
- [7] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Varadarajan, "AI based emotion detection for textual big data: Techniques and contribution," *BDCC*, vol. 5, no. 3, Art. no. 3, Sep. 2021, doi: 10.3390/bdcc5030043.
- [8] B. Kratzwald, S. Ilić, M. Kraus, S. Feuerriegel, and H. Prendinger, "Deep learning for affective computing: Text-based emotion recognition in decision support," *Decision Support Systems*, vol. 115, pp. 24–35, Nov. 2018, doi: 10.1016/j.dss.2018.09.002.
- [9] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human emotion recognition: Review of sensors and methods," *Sensors*, vol. 20, no. 3, Art. no. 3, Jan. 2020, doi: 10.3390/s20030592.
- [10] M. Schreiner, T. Fischer, and R. Riedl, "Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda," *Electron Commer Res*, vol. 21, no. 2, pp. 329–345, Jun. 2021, doi: 10.1007/s10660-019-09353-8.

- Available:  
<https://proceedings.neurips.cc/paper/2017/hash/3f5e243547dee91fbd053c1c4a845aa-Abstract.html>
- [30] E. Batbaatar, M. Li, and K. H. Ryu, "Semantic-emotion neural network for emotion recognition from text," *IEEE Access*, vol. 7, pp. 111866–111878, 2019, doi: 10.1109/ACCESS.2019.2934529.
- [31] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, and P. Agrawal, "Understanding emotions in text using deep learning and big data," *Computers in Human Behavior*, vol. 93, pp. 309–317, Apr. 2019, doi: 10.1016/j.chb.2018.12.029.
- [32] T. T. Sasidhar, P. B., and S. K. P., "Emotion detection in Hinglish (Hindi+English) code-mixed social media text," *Procedia Computer Science*, vol. 171, pp. 1346–1352, 2020, doi: 10.1016/j.procs.2020.04.144.
- [33] M. Abdullah, M. Hadzikadicy, and S. Shaikhz, "SEDAT: Sentiment and emotion detection in Arabic text using CNN-LSTM deep learning," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL: IEEE, Dec. 2018, pp. 835–840. doi: 10.1109/ICMLA.2018.00134.
- [34] X. Sun, C. Zhang, and L. Li, "Dynamic emotion modelling and anomaly detection in conversation based on emotional transition tensor," *Information Fusion*, vol. 46, pp. 11–22, Mar. 2019, doi: 10.1016/j.inffus.2018.04.001.
- [35] M. Baali and N. Ghneim, "Emotion analysis of Arabic tweets using deep learning approach," *J Big Data*, vol. 6, no. 1, Art. no. 1, Dec. 2019, doi: 10.1186/s40537-019-0252-x.
- [36] W. Ragheb, J. Azé, S. Bringay, and M. Servajean, "Attention-based modeling for emotion detection and classification in textual conversations," *arXiv:1906.07020 [cs]*, Jun. 2019, [Online]. Available: <http://arxiv.org/abs/1906.07020>
- [37] K. Shrivastava, S. Kumar, and D. K. Jain, "An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network," *Multimed Tools Appl*, vol. 78, no. 20, pp. 29607–29639, Oct. 2019, doi: 10.1007/s11042-019-07813-9.
- [21] Arman Rayan Sharif. (2023, Jan. 06) ArmanEmo [Online]. Available: <https://github.com/Arman-Rayan-Sharif/arman-text-emotion>
- [۲۲] زارع مهرجردی ف.، یزدیان دهکردی م.، لطیف ع.م.، «ارزیابی روش‌های یادگیری کلاسیک و یادگیری عمیق در تجزیه و تحلیل احساسات داده‌های تلگرام فارسی»، مجله محاسبات نرم، جلد ۱۱، شماره ۱-پ، ۲۱، ۱۴۰۱، doi: 10.22052/scj.2023.246553.1077.
- [۲۳] فیضی‌درخشانی م.، متقی نیا ز.، عسگری چناقلو م.، «طبقه‌بندی متون فارسی مبتنی بر شبکه‌های عصبی عمیق»، مجله محاسبات نرم، جلد ۱۱، شماره ۱-پ، ۲۱، ۱۴۰۱، doi: 10.22052/scj.2023.243182.1010.
- [24] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv:1301.3781 [cs]*, Sep. 2013, [Online]. Available: <http://arxiv.org/abs/1301.3781>
- [25] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 135–146, 2017, doi: 10.1162/tacl\_a\_00051.
- [26] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar: Association for Computational Linguistics, 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.
- [27] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv:1810.04805 [cs]*, May 2019, [Online]. Available: <http://arxiv.org/abs/1810.04805>
- [28] K. Clark, U. Khandelwal, O. Levy, and C. D. Manning, "What does BERT look at? An analysis of BERT's attention," in *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, Florence, Italy: Association for Computational Linguistics, Aug. 2019, pp. 276–286. doi: 10.18653/v1/W19-4828.
- [29] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. [Online].

- [45] H. Luo, "Emotion detection for spanish with data augmentation and transformer-based models," in *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2021)*, Malaya, Spain: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2021), 2021, pp. 35–42.
- [46] H. Q. Abonizio, E. C. Paraiso, and S. Barbon Junior, "Toward text data augmentation for sentiment analysis," *IEEE Transactions on Artificial Intelligence*, pp. 1–1, 2021, doi: 10.1109/TAI.2021.3114390.
- [47] J. Luo, M. Bouazizi, and T. Ohtsuki, "Data augmentation for sentiment analysis using sentence compression-based SeqGAN with data screening," *IEEE Access*, vol. 9, pp. 99922–99931, 2021, doi: 10.1109/ACCESS.2021.3094023.
- [48] P. Kumar and B. Raman, "A BERT based dual-channel explainable text emotion recognition system," *Neural Networks*, vol. 150, pp. 392–407, Jun. 2022, doi: 10.1016/j.neunet.2022.03.017.
- [49] A. Majeed, M. O. Beg, U. Arshad, and H. Mujtaba, "Deep-EmoRU: mining emotions from roman urdu text using deep learning ensemble," *Multimed Tools Appl*, May 2022, doi: 10.1007/s11042-022-13147-w.
- [50] X. Zhu, Y. Lou, H. Deng, and D. Ji, "Leveraging bilingual-view parallel translation for code-switched emotion detection with adversarial dual-channel encoder," *Knowledge-Based Systems*, vol. 235, p. 107436, Jan. 2022, doi: 10.1016/j.knosys.2021.107436.
- [51] A. B. Warriner, V. Kuperman, and M. Brysbaert, "Norms of valence, arousal, and dominance for 13,915 English lemmas," *Behav Res*, vol. 45, no. 4, pp. 1191–1207, Dec. 2013, doi: 10.3758/s13428-012-0314-x.
- [52] M. Lewis *et al.*, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension." arXiv, Oct. 29, 2019. doi: 10.48550/arXiv.1910.13461.
- [53] X. Wu, S. Lv, L. Zang, J. Han, and S. Hu, "Conditional BERT contextual augmentation," in *Computational Science – ICCS 2019*, J. M. F. Rodrigues, P. J. S. Cardoso, J. Monteiro, R. Lam, V. V. Krzhizhanovskaya, M. H. Lees, J. J. Dongarra, and P. M. A. Sloot, Eds., in Lecture Notes in Computer Science. Cham: Springer International
- [38] A. F. Adoma, N.-M. Henry, W. Chen, and N. Rubungo Andre, "Recognizing emotions from texts using a bert-based approach," in *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China: IEEE, Dec. 2020, pp. 62–66. doi: 10.1109/ICCWAMTIP51612.2020.9317523.
- [39] H. Al-Omari, M. A. Abdullah, and S. Shaikh, "EmoDet2: Emotion detection in english textual dialogue using BERT and BiLSTM models," in *2020 11th International Conference on Information and Communication Systems (ICICS)*, Irbid, Jordan: IEEE, Apr. 2020, pp. 226–232. doi: 10.1109/ICICS49469.2020.239539.
- [40] S. Q. Suidong Qu, Y. Y. Yanhua Yang, and Q. Q. Qinyu Que, "Emotion classification for spanish with XLM-RoBERTa and TextCNN,," 2021, [Online]. Available: <https://www.semanticscholar.org/paper/17801995b7352f212d2841ad7b59482379de50a5>
- [41] A. Das, O. Sharif, M. M. Hoque, and I. H. Sarker, "Emotion classification in a resource constrained language using transformer-based approach," *arXiv:2104.08613 [cs]*, Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2104.08613>
- [42] K. Dheeraj and T. Ramakrishnudu, "Negative emotions detection on online mental-health related patients texts using the deep learning with MHA-BCNN model," *Expert Systems with Applications*, vol. 182, p. 115265, Nov. 2021, doi: 10.1016/j.eswa.2021.115265.
- [43] Y. Fu, L. Guo, L. Wang, Z. Liu, J. Liu, and J. Dang, "A sentiment similarity-oriented attention model with multi-task learning for text-based emotion recognition," in *MultiMedia Modeling*, J. Lokoč, T. Skopal, K. Schoeffmann, V. Mezaris, X. Li, S. Vrochidis, and I. Patras, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 278–289. doi: 10.1007/978-3-030-67832-6\_23.
- [44] X.-M. Lin, C.-H. Ho, L.-T. Xia, and R.-Y. Zhao, "Sentiment analysis of low-carbon travel APP user comments based on deep learning," *Sustainable Energy Technologies and Assessments*, vol. 44, p. 101014, Apr. 2021, doi: 10.1016/j.seta.2021.101014.

Publishing, 2019, pp. 84–95. doi: 10.1007/978-3-030-22747-0\_7.

- [54] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Text Data Augmentation for Deep Learning," *J Big Data*, vol. 8, no. 1, Art. no. 1, Dec. 2021, doi: 10.1186/s40537-021-00492-0.
- [55] S. Y. Feng *et al.*, "A survey of data augmentation approaches for NLP," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Online: Association for Computational Linguistics, Aug. 2021, pp. 968–988. doi: 10.18653/v1/2021.findings-acl.84.
- [56] A. Khosravi, M. Kelarestaghi, and M. Purmohammad, "Emotion detection in persian text: A machine learning model," *BJCP*, vol. 14, no. 1, Art. no. 1, Aug. 2019, doi: 10.29252/bjcp.14.1.42.
- [57] S. S. Sadeghi, H. Khotanlou, and M. Rasekh Mahand, "Automatic persian text emotion detection using cognitive linguistic and deep learning," *JAIDM*, vol. 9, no. 2, pp. 169–179, Jan. 2021, doi: 10.22044/jadm.2020.9992.2136.
- [58] A. Abaskohi, N. Sabri, and B. Bahrak, "Persian emotion detection using ParsBERT and imbalanced data handling approaches," Nov. 2022, doi: 10.48550/arXiv.2211.08029.
- [59] A. Khodaei, A. Bastanfard, H. Saboohi, and H. Aligholizadeh, "Deep emotion detection sentiment analysis of persian literary texts," Research Square preprint, 2022, [Online]. Available: <https://doi.org/10.21203/rs.3.rs-1796157/v1>.