



دانشکده مهندسی کامپیوتر

## تشخیص موضع متنی در شبکه‌های اجتماعی

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته مهندسی کامپیوتر  
گرایش هوش مصنوعی و رباتیک

غزاله محمودی

استاد راهنما

دکتر سید صالح اعتمادی

بهمن ۱۴۰۲

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

## تأییدیه‌ی هیأت داوران جلسه‌ی دفاع از پایان‌نامه

نام دانشکده: دانشکده مهندسی کامپیوتر

نام دانشجو: غزاله محمودی

عنوان پایان‌نامه: تشخیص موضع متنی در شبکه‌های اجتماعی

تاریخ دفاع: بهمن ۱۴۰۲

رشته: مهندسی کامپیوتر

گرایش: هوش مصنوعی و رباتیک

ردیف	سمت	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضا
۱	استاد راهنما	دکتر سید صالح اعتمادی	استادیار	دانشگاه علم و صنعت ایران	
۲	استاد مدعو داخلی	دکتر حسین رحمانی	استادیار	دانشگاه علم و صنعت ایران	
۳	استاد مدعو خارجی	دکتر سعیده ممتازی	دانشیار	دانشگاه امیرکبیر	

## تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب غزاله محمودی به شماره دانشجویی ۴۰۰۷۲۲۱۵۶ دانشجوی رشته مهندسی کامپیوتر مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه‌ی نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه‌برداری‌شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم ( قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ... ) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض درخصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذیصلاح (اعم از اداری و قضایی) به عهده‌ی اینجانب خواهد بود و دانشگاه هیچ‌گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: غزاله محمودی

تاریخ و امضا:

## مجوز بهره‌برداری از پایان‌نامه

بهره‌برداری از این پایان‌نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما به شرح زیر تعیین می‌شود، بلامانع است:

- ☐ بهره‌برداری از این پایان‌نامه برای همگان بلامانع است.
- ☐ بهره‌برداری از این پایان‌نامه با اخذ مجوز از استاد راهنما، بلامانع است.
- ☐ بهره‌برداری از این پایان‌نامه تا تاریخ ..... ممنوع است.

استاد راهنما: دکتر سید صالح اعتمادی

تاریخ:

امضا:

## قدردانی

سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را زیور عقل آراست و به ما توفیق حرکت در مسیر کسب علم را ارزانی داشت. در آغاز وظیفه خود می‌دانم از زحمات استاد گران‌قدر و فرزانه جناب آقای دکتر سید صالح اعتمادی که راهنمایی اینجانب را در دوره کارشناسی ارشد عهده‌دار بودند، بی‌نهایت سپاس‌گزار هستم. همچنین از آقای بابک بهکام‌کیا که به عنوان همکار در بخشی از پژوهش حضور داشتند کمال تشکر را دارم. در آخر از خانواده عزیزم به پاس حمایت‌های بی دریغ و عاطفه سرشارشان در طول این مدت کمال تشکر و قدردانی را دارم.

غزاله محمودی

بهمن ۱۴۰۲

## چکیده

امروزه شبکه‌های اجتماعی بستری جهت بیان آزادانه عقاید و اشتراک گذاشتن نظرات می‌باشد. این موضوع سبب شده است که با تحلیل داده‌های موجود در شبکه‌های اجتماعی بتوان دید وسیع و جامعی از موضع کاربران متفاوت نسبت به موضوعات مختلف به دست آورد. از جمله این موضوعات می‌توان به مسائل سیاسی، اقتصادی، اجتماعی و فرهنگی اشاره کرد. در پردازش زبان طبیعی، به فرآیند تشخیص خودکار موضع متن نسبت به موضوعی مشخص و معین، تشخیص موضع گفته می‌شود.

در مسائل پردازش زبان طبیعی از جمله تشخیص موضع، نحوه پیش‌پردازش داده‌های متنی در عملکرد مدل آموزش دیده تاثیر به سزایی دارد. در این پژوهش هفت سطح مختلف پیش‌پردازش معرفی شده و مورد بررسی قرار می‌گیرد. علاوه بر این، برای یافتن معماری مدل تشخیص موضع، از ایده جستجو معماری عصبی (NAS) الهام گرفته شد. در این روش با تقسیم معماری مدل به چهار بخش اصلی و تعریف فضای جستجو برای هر بخش و استفاده از الگوریتم جستجو تطبیقی، معماری نهایی طراحی می‌شود. در نهایت بهترین مدل پیشنهاد شده از کدگذار BERTweet و رده‌بند CNN استفاده می‌کند. معماری طراحی شده توانست به 74.47 درصد در معیار F1 دست یابد و نسبت به مدل پایه 19.97 درصد بهبود داشته باشد. همچنین روش ارائه شده رتبه سوم را بین ۱۹ شرکت‌کننده رویداد تشخیص موضع در تغییرات اقلیمی کسب کرد. از سوی دیگر با توجه به کمبود داده‌های آموزشی برای موضوعات متفاوت، تشخیص موضع بدون داده آموزشی نیز مورد بررسی قرار گرفت. در این روش که از مدل‌های زبانی بزرگ و مهندسی پرامپت استفاده می‌کند، چهار رویکرد بر اساس انواع مختلف پرامپت معرفی شد. سپس عملکرد پرامپت‌های پیشنهادی با سایر روش‌های تشخیص موضع بدون داده آموزشی مورد مقایسه قرار گرفت. رویکرد معرفی شده توانست به مقدار 57.33 درصد در معیار F1 دست یابد و نسبت به رویکردهای مشابه 2.03 درصد بهبود داشته باشد.

واژگان کلیدی: تشخیص موضع، پردازش زبان طبیعی، یادگیری عمیق، مدل‌های زبانی بزرگ

# فهرست مطالب

د	فهرست شکل‌ها
ر	فهرست جدول‌ها
۱	فصل ۱: مقدمه
۲	۱-۱ تعریف مسئله تشخیص موضع
۳	۲-۱ ساختار پایان‌نامه
۱	فصل ۲: تعریف مسئله
۱	۱-۲ سطوح مسئله تشخیص موضع
۲	۲-۲ تقسیم‌بندی تشخیص موضع بر اساس موضوع
۲	۱-۲-۲ تشخیص موضع برای یک موضوع خاص
۳	۲-۲-۲ تشخیص موضع برای موضوعات چندگانه مرتبط
۳	۳-۲-۲ تشخیص موضع ادعا محور
۴	۴-۲-۲ تفاوت تشخیص موضع و تحلیل احساسات
۶	۵-۲-۲ کاربردهای تشخیص موضع
۸	فصل ۳: تعاریف و مفاهیم ابتدایی
۸	۱-۳ شبکه‌های اجتماعی
۹	۲-۳ هوش مصنوعی
۹	۳-۳ پردازش زبان طبیعی



۴-۳	تعبیه کلمات	۱۰
۵-۳	یادگیری ماشین	۱۳
۶-۳	یادگیری عمیق	۱۳
۷-۳	شبکه‌های عصبی همگشتی	۱۴
۸-۳	ساز و کار توجه	۱۵
۹-۳	ترنسفورمر	۱۶
۱۰-۳	مهندسی پرامپت	۱۶
۱۱-۳	جستجو پارامترها	۱۷
۱۲-۳	جمع بندی	۱۸

#### فصل ۴: مروری بر کارهای مرتبط

۱-۴	رقابت‌های برگزار شده در مسئله تشخیص موضع	۲۰
۱-۱-۴	رویداد SemEval-2016	۲۰
۲-۱-۴	رویداد NLPCC-ICCPOL-2016	۲۱
۳-۱-۴	رویداد IberEval-2017	۲۲
۴-۱-۴	رویداد SadriStance@EVALITA-2020	۲۲
۵-۱-۴	رویداد VaxxStance@IberLFE-2021	۲۴
۶-۱-۴	رویداد ClimateActivismStance@CASE-2024	۲۵
۲-۴	مجموعه داده‌های مسئله تشخیص موضع	۲۶
۱-۲-۴	مجموعه داده SemEval-2016	۲۸
۲-۲-۴	مجموعه داده ClimaConvo	۲۹
۳-۴	پژوهش‌های انجام شده در مسئله تشخیص موضع	۳۰
۱-۳-۴	رویکردهای یادگیری ماشین مبتنی بر ویژگی	۳۱
۲-۳-۴	رویکردهای یادگیری عمیق	۳۲
۳-۳-۴	مدل‌های از قبل آموزش دیده	۳۴
۴-۴	معیار ارزیابی مسئله تشخیص موضع	۳۷

۳۸	فصل ۵: روش پیشنهادی برای تشخیص موضع با نظارت
۳۹	۱-۵ دادگان آموزشی
۴۰	۲-۵ روش حل مسئله
۴۰	۱-۲-۵ پیش پردازش داده‌ها
۴۲	۲-۲-۵ افزایش داده
۴۳	۳-۲-۵ جستجو معماری
۴۵	۴-۲-۵ انتخاب ابرپارامترها
۴۶	۳-۵ آزمایش‌ها و تحلیل نتایج
۴۶	۱-۳-۵ گام اول: تحلیل نحوه عملکرد الگوریتم جستجو
۴۷	۲-۳-۵ گام دوم: به دست آوردن پارامترهای مدل با عملکرد مطلوب
۴۹	۳-۳-۵ گام سوم: ارزیابی تکمیلی مدل پیشنهادی
۵۰	۱-۳-۳-۵ اثربخشی نوع پیش پردازش داده
۵۱	۲-۳-۳-۵ اثربخشی افزایش داده
۵۳	۳-۳-۳-۵ اثربخشی رده‌بند
۵۳	۴-۳-۳-۵ اثربخشی تابع ضرر
۵۴	۵-۳-۳-۵ بررسی تاثیر دامنه بر نتایج مدل آموزش دیده
۵۵	۴-۳-۵ تحلیل خطا
۵۶	۵-۳-۵ مزایا و معایب روش پیشنهادی
۵۷	۶-۳-۵ نحوه پیاده‌سازی و اجرا آزمایش‌ها
۵۸	۴-۵ جمع‌بندی
۵۹	فصل ۶: روش پیشنهادی برای تشخیص موضع بدون داده آموزشی
۵۹	۱-۶ روش حل مسئله
۶۰	۱-۱-۶ پرامپت Zero Shot
۶۱	۲-۱-۶ پرامپت Zero Shot + Chain of Thought
۶۱	۳-۱-۶ پرامپت Zero Shot + Chain of Thought + Context Description

۶۲ . . . . .	۴-۱-۶ پرامپت Few Shot + Chain of Thought + Context Description
۶۳ . . . . .	۲-۶ آزمایش‌ها و تحلیل نتایج
۶۶ . . . . .	۱-۲-۶ مزایا و معایب روش پیشنهادی
۶۶ . . . . .	۲-۲-۶ نحوه پیاده‌سازی و اجرا آزمایش‌ها
۶۷	فصل ۷: نتیجه‌گیری و کارهای آینده
۶۷ . . . . .	۱-۷ نتیجه‌گیری
۶۷ . . . . .	۲-۷ پیشنهادها و کارهای آینده
۶۹	مراجع
۷۸	واژه‌نامه فارسی به انگلیسی
۸۰	واژه‌نامه انگلیسی به فارسی

## فهرست شکل‌ها

- ۱-۱ تشخیص موضع در شبکه‌های اجتماعی . . . . . ۲
- ۱-۲ توزیع برچسب‌های تحلیل احساسات نسبت به برچسب‌های تشخیص موضع . . . . . ۵
- ۲-۲ موضوعات تحقیقاتی مرتبط با تشخیص موضع . . . . . ۶
- ۱-۳ تولید بازنمایی متن با استفاده از شبکه CNN . . . . . ۱۲
- ۱-۴ نمودار تعداد مقالات چاپ شده در رابطه با تشخیص موضع به تفکیک سال . . . . . ۱۹
- ۲-۴ مروری بر ویژگی‌های مورد استفاده در تشخیص موضع . . . . . ۳۱
- ۳-۴ معماری روش مبتنی بر Classifier Stacked . . . . . ۳۲
- ۴-۴ معماری مدل BiCond . . . . . ۳۴
- ۵-۴ معماری مدل PNEM . . . . . ۳۵
- ۶-۴ معماری مدل MeLT . . . . . ۳۶
- ۷-۴ معماری مدل BiCond . . . . . ۳۶
- ۱-۵ توزیع تعداد تک واژه‌های مجموعه داده ClimaConvo . . . . . ۳۹
- ۲-۵ توزیع کلاس‌ها قبل و بعد از افزایش داده. . . . . ۴۲
- ۳-۵ رده‌بند با شبکه عصبی پیچشی . . . . . ۴۴
- ۴-۵ عملکرد مدل آموزش دیده در هر epoch آموزش . . . . . ۴۹
- ۵-۵ مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده . . . . . ۵۱
- ۶-۵ بازه اطمینان مدل پیشنهادی بر اساس روش‌های پیش‌پردازش داده . . . . . ۵۱
- ۷-۵ بازه اطمینان مدل پیشنهادی آموزش دیده با داده‌های افزایش داده شده . . . . . ۵۲

۵-۸	ماتریس درهم‌ریختگی برای داده ارزیابی مجموعه ClimaConvo	۵۵
۵-۹	خروجی میزان توجه مدل به کلمات ورودی برای پیشبینی‌های درست	۵۶
۵-۱۰	خروجی میزان توجه مدل به کلمات ورودی برای پیشبینی‌های اشتباه	۵۶
۵-۱۱	میزان مصرف GPU در زمان آموزش مدل پیشنهادی	۵۸
۶-۱	الگو پرامپت Zero Shot	۶۰
۶-۲	الگو پرامپت Zero Shot + Chain of Thought	۶۱
۶-۳	الگو پرامپت Zero Shot + Chain of Thought + Context Description	۶۲
۶-۴	الگو پرامپت Few Shot + Chain of Thought + Context Description	۶۳
۶-۵	مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده	۶۵
۶-۶	مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده	۶۵
۶-۷	میزان مصرف GPU در زمان اجرا آزمایش‌های مدل‌های زبانی بزرگ	۶۶

## فهرست جدول‌ها

۱-۲	نمونه‌ای از مجموعه داده SemEval	۵
۱-۳	نمونه پرامپت برای مسئله تحلیل احساسات	۱۶
۱-۴	نتایج رویداد ClimateActivismStance برای زیر مسئله سوم (تشخیص موضع)	۲۶
۲-۴	توزیع نمونه‌ها در مجموعه داده Multi-Target	۲۷
۳-۴	چند نمونه از داده‌های مجموعه داده SemEval	۲۸
۴-۴	چند نمونه از داده‌های مجموعه داده ClimaConvo	۲۹
۵-۴	چند نمونه از داده‌های مجموعه داده ClimaConvo	۳۰
۶-۴	چند نمونه از پژوهش‌های مسئله تشخیص موضع	۳۳
۱-۵	توزیع کلاس‌ها در بخش‌های آموزش، توسعه و ارزیابی مجموعه داده ClimaConvo	۳۹
۲-۵	توزیع کلاس‌ها در بخش‌های ارزیابی مجموعه داده SemEval	۴۰
۳-۵	فضای جستجو معماری پیشنهادی	۴۳
۴-۵	فضا جستجو ابرپارامترها	۴۶
۵-۵	نتایج به دست آمده بر روی مجموعه داده ClimateCanvo Stance بخش ارزیابی	۴۸
۶-۵	مقایسه بهترین معماری طراحی شده با مدل پایه	۴۸
۷-۵	پارامترهای بهترین مدل پیشنهادی	۴۹
۸-۵	نتایج بررسی اثربخشی نوع پیش‌پردازش داده	۵۰
۹-۵	نتایج بررسی اثربخشی افزایش داده	۵۲
۱۰-۵	نتایج بررسی اثربخشی افزایش داده به ازای هر کلاس	۵۲

۵۳	۱۱-۵ نتایج بررسی اثربخشی رده‌بند
۵۴	۱۲-۵ نتایج بررسی اثربخشی تابع ضرر بر حسب F1-Score
۵۴	۱۳-۵ بررسی تاثیر دامنه بر نتایج مدل آموزش دیده بر اساس معیار F1-Score
۵۵	۱۴-۵ چند نمونه پیشبینی‌های اشتباه مدل
۶۴	۱-۶ مقایسه نتایج پرامپت‌های معرفی شده با سایر مدل‌های Zero-Shot
۶۴	۲-۶ نتایج پرامپت‌های معرفی شده روی مجموعه داده ClimaConvo
۶۵	۳-۶ تاثیر پیش‌پردازش داده‌ها در نتایج به دست آمده

# فصل ۱

## مقدمه

امروزه شبکه‌های اجتماعی یکی از اجزا اصلی تعامل اجتماعی افراد را تشکیل می‌دهد. شبکه‌های اجتماعی ابزار قدرت‌مندی برای بیان نظرات و دیدگاه‌های افراد در رابطه با موضوعات متفاوت می‌باشد. این بستر برای افراد این امکان را فراهم می‌کند که آزادانه نظرات خود را بیان کنند و بازخورد فوری دریافت کنند. همچنین دیدگاه سایر افراد در رابطه با موضوع بیان شده را در این بستر می‌توان مورد بررسی قرار داد.

وابستگی افراد به استفاده از شبکه‌های اجتماعی این امکان را فراهم می‌کند تا داده‌های جمع‌آوری شده از شبکه‌های اجتماعی را به عنوان منبع مهم برای مطالعه جنبه‌های مختلف موضع‌گیری افراد در زمینه‌های سیاسی و اجتماعی مورد استفاده قرار دهیم.

استخراج خودکار اطلاعات از متون زبان طبیعی از مهم‌ترین زمینه‌های تحقیقاتی می‌باشد. در این راستا می‌توان مسئله تشخیص موضع<sup>۱</sup> را تعریف کرد. تشخیص موضع به معنای بیان و تشخیص دیدگاه نویسنده‌ی متن نسبت به یک گزاره‌ی (موضوع) معین است.

تشخیص موضع در مطالعات تحلیلی برای تشخیص جهت‌گیری افکار عمومی در رسانه‌های اجتماعی، مانند مسائل سیاسی و اجتماعی، نقش کلیدی ایفا می‌کند. از این رو می‌تواند در شبکه‌های اجتماعی در بستر اهداف مختلفی از جمله تصمیمات دولت، تبلیغات، اقناع افکار عمومی مورد استفاده قرار بگیرد. به عبارت دیگر نتایج به دست آمده منجر به دید وسیع‌تری به نظرات عمومی و اخذ تصمیمات بهتر می‌شود.

---

<sup>1</sup> Stance Detection

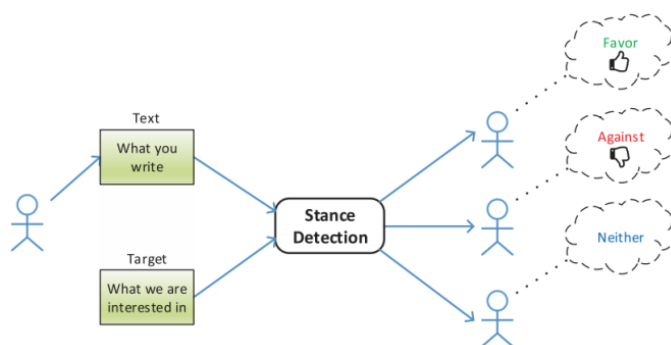


## ۱-۱ تعریف مسئله تشخیص موضع

موضع گیری<sup>۲</sup> به بیان دیدگاه و نظرات افراد راجع به یک گزاره‌ی (موضوع) معین گفته می‌شود [۲۱]. تشخیص موضع<sup>۳</sup> نقشی اساسی در شناخت جهت گیری عمومی در رابطه با مسائل اجتماعی و سیاسی، انتخابات و موضع تصمیمات دولت ایفا می‌کند [۴، ۲۰].

مسئله تشخیص موضع در شبکه‌های اجتماعی به این صورت تعریف می‌شود که یک متن نوشته شده در شبکه اجتماعی (توییت، نظرات و ادعا) و یک موضوع یا هدف مشخص<sup>۴</sup> (شخص، سازمان، جنبش، سیاست) به عنوان ورودی در نظر گرفته می‌شود. هدف تشخیص خودکار موضع متن داده شده نسبت به موضوع مشخص شده است [۴]. این یک مسئله به صورت رده‌بندی است و خروجی یکی از موارد زیر می‌تواند باشد [۲۸].

۱. موافق<sup>۵</sup>: متن مورد نظر، موضوع تعیین شده را بدون هیچ گونه ابهام تایید می‌کند.
۲. مخالف<sup>۶</sup>: متن مورد نظر، موضوع تعیین شده را بدون هیچ گونه ابهام رد می‌کند.
۳. بدون نظر<sup>۷</sup>: متن و موضوع به هم مرتبط نیستند و یا متن به صورت صریح موضوع را رد یا تایید نمی‌کند.



شکل ۱-۱: تشخیص موضع در شبکه‌های اجتماعی [۲۵]

<sup>2</sup>Stance

<sup>3</sup>Stance Detection

<sup>4</sup>Target

<sup>5</sup>Favor/Support

<sup>6</sup>Against/Oppose

<sup>7</sup>None/Neither/Neutral

## ۱-۲ ساختار پایان نامه

این پایان نامه در هفت فصل تنظیم شده است. فصل اول مقدمه‌ای از حوزه تحقیقات را شرح می‌دهد. در فصل دوم تعریف مسئله تشخیص موضع به صورت دقیق بیان می‌شود. همچنین جایگاه تشخیص موضع در مقایسه با سایر مسائل مشابه به طور دقیق مشخص می‌شود. فصل سوم به توضیح تعاریف و مفاهیم ابتدایی مورد نیاز برای فهم تشخیص موضع و روش‌های حل مسئله می‌پردازد. فصل چهارم شامل مروری بر کارهای مرتبط به تشخیص موضع می‌باشد. در فصل پنجم رویکرد حل تشخیص موضع با نظارت شرح داده شده و بررسی می‌شود. فصل ششم نیز شامل معرفی رویکرد پیشنهادی برای تشخیص موضع بدون داده آموزشی می‌باشد. در نهایت در فصل هفتم (فصل آخر) جمع‌بندی از روش‌های حل مسئله انجام شده و پیشنهادهایی برای کارهای آینده ارائه می‌شود.

## فصل ۲

### تعریف مسئله

تشخیص موضع در مطالعات تحلیلی برای تشخیص جهت گیری افکار عمومی در رسانه‌های اجتماعی، مانند مسائل سیاسی و اجتماعی، نقش کلیدی ایفا می‌کند. از این رو می‌تواند در شبکه‌های اجتماعی در بستر اهداف مختلفی از جمله تصمیمات دولت، تبلیغات، اقناع افکار عمومی مورد استفاده قرار بگیرد. به عبارتی دیگر نتایج به دست آمده منجر به دید وسیع تری به نظرات عمومی و اخذ تصمیمات بهتر می‌شود.

#### ۱-۲ سطوح مسئله تشخیص موضع

قبل از بررسی گونه‌های مختلف تشخیص موضع، شناخت دقیق سطوح مختلف تشخیص موضع از اهمیت بالایی برخوردار است. مسئله تشخیص موضع معمولاً در دو سطح بررسی می‌شود. در ادامه توضیحات مختصری از این دو سطح را بیان می‌شود.

##### ۱. سطح بیانی<sup>۱</sup>

در این سطح هدف این است موضع یک متن نوشته شده نسبت به موضوع از قبل مشخص شده، تعیین گردد. این سطح معمول‌ترین روش به کارگرفته شده در تشخیص موضع است. به صورتی که ویژگی‌های متن نوشته شده استخراج می‌شود و تشخیص موضع صرفاً بر اساس متن انجام می‌شود. این مفهوم معمولاً به صورت رده‌بندی با سه کلاس (موافق، مخالف، بدون نظر) تعریف می‌شود. بدون

---

<sup>1</sup>The Statement Level

نظر یعنی متن نوشته شده نسبت به موضوع بی ربط است. حالت دیگر این است که موضوع در آن بررسی شده ولی موضع مشخصی در رابطه با آن (موافق یا مخالف) ندارد.

## ۲. سطح کاربر<sup>۲</sup>

در این سطح هدف پیشبینی موضع کاربر نسبت به یک موضوع می‌باشد. متغیرهای مختلفی که یک کاربر از خود در شبکه‌های اجتماعی به جای می‌گذارد، در بررسی‌ها مورد استفاده قرار می‌گیرد. پست‌های کاربر، پست‌هایی که پسندیده و یا مجدد به اشتراک گذاشته، از جمله اطلاعاتی است که در این فرآیند مورد استفاده قرار می‌گیرد. در این سطح معمولاً رده بندی با دو کلاس (موافق و مخالف) تعریف می‌شود. چرا که این فرض در نظر گرفته شده که حتی اگر بعضی از پست‌های کاربر در رابطه با موضوع معین بدون موضع باشد، اما به هر حال کاربر یک موضع مشخصی نسبت به آن موضوع دارد.

آشنایی با سطوح مختلف تشخیص موضع و تفاوت آن‌ها بسیار با اهمیت است. پست‌های مختلف یک کاربر نسبت به یک موضوع می‌تواند موضع‌های مختلفی داشته باشد. ولی زمانی که در سطح کاربر مسئله تشخیص موضع را تعریف می‌کنیم، موضع کلی کاربر نسبت به موضوع مطرح می‌شود.

## ۲-۲ تقسیم بندی تشخیص موضع بر اساس موضوع

در تشخیص موضع همان‌طور که در گذشته بیان کردیم، نیاز به مشخص بودن یک موضوع (هدف) برای تشخیص موضع در رابطه با آن وجود دارد.

### ۱-۲-۲ تشخیص موضع برای یک موضوع خاص<sup>۳</sup>

رایج ترین نوع تشخیص موضع در شبکه‌های اجتماعی، تشخیص موضع برای یک موضوع خاص است. در این نوع تشخیص موضع متن یا کاربر، ورودی اصلی برای پیش‌بینی موضع نسبت به یک هدف مشخص است. بنابراین برای هر هدف یک مدل رده‌بندی جداگانه آموزش داده می‌شود.

<sup>2</sup>The User Level

<sup>3</sup>Target-specific Stance Detection

۲-۲-۲ تشخیص موضع برای موضوعات چندگانه مرتبط<sup>۴</sup>

در این روش، تشخیص موضع نسبت به چندین هدف مرتبط انجام می‌شود. این هدف‌ها می‌توانند یک موضوع کلی، رویداد خاص و یا اشخاص مختلف باشند. داده‌های یک هدف می‌تواند دانشی برای هدف‌های دیگر داشته باشد. به عنوان مثال فردی که موافقت خود را با یکی از نامزدهای انتخابات ریاست جمهوری اعلام می‌کند، به صورت متوالی مخالف خود را با نامزد حزب (گروه) مقابل نیز اعلام کرده است.

لازم به ذکر است نوع دیگری از تشخیص موضع با عنوان تشخیص موضع برای موضوع متقابل<sup>۵</sup> نیز تعریف شده است. در این روش آموزش مدل بر روی یک موضوع مشخص انجام می‌شود. در مرحله ارزیابی داده‌های موضوع متفاوت اما مرتبط با موضوع زمان آموزش مورد استفاده قرار می‌گیرد. در شرایطی که مجموعه داده کافی برای آموزش مدل با بهترین عملکرد در دسترس نباشد، یکی از ایده‌های قابل استفاده این مورد می‌باشد. به عنوان مثال در پژوهش [۶۳] آموزش مدل بر روی داده‌ها با برچسب هیلاری کلینتون انجام می‌شود. در زمان ارزیابی از داده‌ها با برچسب دونالد ترامپ استفاده می‌شود. در واقع همچون حالت قبل مدل به نوعی آموزش می‌بیند که بتواند برای دو موضوع متفاوت ولی مرتبط عملیات تشخیص موضع را انجام دهد.

۲-۲-۳ تشخیص موضع ادعا محور<sup>۶</sup>

در این حالت هدف تشخیص موضع، یک هدف مشخص (صریح) مثل شخص، رویداد نیست. در واقع هدف در اینجا می‌تواند ادعای عنوان خبر یا شایعه باشد. در ادامه برای آشنایی بیشتر با مسائل مرتبط و فهم جایگاه مسئله اصلی، توضیح دقیق‌تر انواع آن و برچسب‌های موجود در هر مورد را بررسی می‌کنیم.

۱. تشخیص موضع در اخبار جعلی<sup>۷</sup>

در این حالت ورودی تیتیر خبر<sup>۸</sup> به همراه یک متن خبری کامل<sup>۹</sup> می‌باشد (تیتیر خبر و متن خبر ممکن است برای دو خبر کاملاً متفاوت باشند). در نهایت هدف این است موضع متن خبری نسبت به ادعا مطرح شده در تیتیر خبر مورد بررسی قرار بگیرد. این مسئله به صورت رده‌بندی با چهار کلاس

<sup>۴</sup>Multi-Related-Target Stance Detection<sup>۵</sup>Cross-Target Stance Detection<sup>۶</sup>Claim-Based Stance Detection<sup>۷</sup>Fake News Stance Detection<sup>۸</sup>News Headline<sup>۹</sup>News Body

(تایید <sup>۱۰</sup>، رد <sup>۱۱</sup>، بحث شده (موضوع مرتبط بدون نظر قطعی) <sup>۱۲</sup>، نامرتبط <sup>۱۳</sup>) تعریف می شود. این مسئله برای حل تشخیص اخبار جعلی مطرح می شود.

## ۲. تشخیص شایعه <sup>۱۴</sup>

در این حالت ورودی شایعه به همراه یک قطعه متن است. هدف این است موضع نویسنده متن نسبت به صحت شایعه مطرح شده سنجیده شود. این مسئله به صورت رده بندی تعریف می شود. در بعضی تعاریف رده بندی با چهار کلاس شامل (تایید <sup>۱۵</sup>، رد <sup>۱۶</sup>، پرس و جو <sup>۱۷</sup> و اظهار نظر <sup>۱۸</sup>) تعریف می شود. در برخی دیگر از تعاریف رده بندی با دو کلاس (تایید <sup>۱۹</sup> و رد <sup>۲۰</sup>) تعریف می شود.

## ۲-۲-۴ تفاوت تشخیص موضع و تحلیل احساسات

مسئله تحلیل احساسات <sup>۲۱</sup> به عنوان مسئله پردازش احساسی متون در نظر گرفته می شود که می تواند بدون داشتن یک هدف مشخص استنباط شود. در این مسئله مشخص می شود یک متن نوشته شده بر اساس محتوای زبان، احساس مثبت <sup>۲۲</sup>، منفی <sup>۲۳</sup> یا خنثی <sup>۲۴</sup> دارد. در حالی که در تشخیص موضع موافق یا مخالف بودن نویسنده متن با موضوع مشخص شده، مد نظر است. شکل ۱-۲ توزیع برچسب های تحلیل احساسات نسبت به تشخیص موضع در مجموعه داده SEM-Eval2016 [۴۹] را نشان می دهد. برچسب های تحلیل احساسات و تشخیص موضع همیشه یکسان نیست. به عنوان مثال تمام دادگانی که در تحلیل احساسات برچسب مثبت دارند، در تشخیص موضع برچسب موافق نگرفته اند. حدود ۳۰ درصد موافق و حدود ۳۳ درصد مخالف و مابقی نیز برچسب خنثی دارند.

<sup>10</sup>Agrees

<sup>11</sup>Disagrees

<sup>12</sup>Discusses

<sup>13</sup>Unrelated

<sup>14</sup>Rumor Stance Detection

<sup>15</sup>Supporting

<sup>16</sup>Denying

<sup>17</sup>Querying

<sup>18</sup>Commenting

<sup>19</sup>Supporting

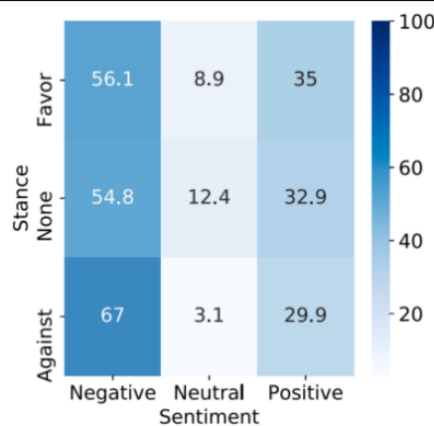
<sup>20</sup>Denying

<sup>21</sup>Sentiment Analysis

<sup>22</sup>Positive

<sup>23</sup>Negative

<sup>24</sup>Neutral Sentiment



شکل ۲-۱: توزیع برجسب‌های تحلیل احساسات نسبت به برجسب‌های تشخیص موضع در مجموعه دادگان

SEM-Eval2016 [۴]

جدول ۲-۱: نمونه‌ای از مجموعه داده SemEval [۴]

متن توییت	موضوع	تشخیص موضع	تحلیل احساسات
I am sad that Hillary lost this presidential race.	Hillary Clinton	موافق	منفی

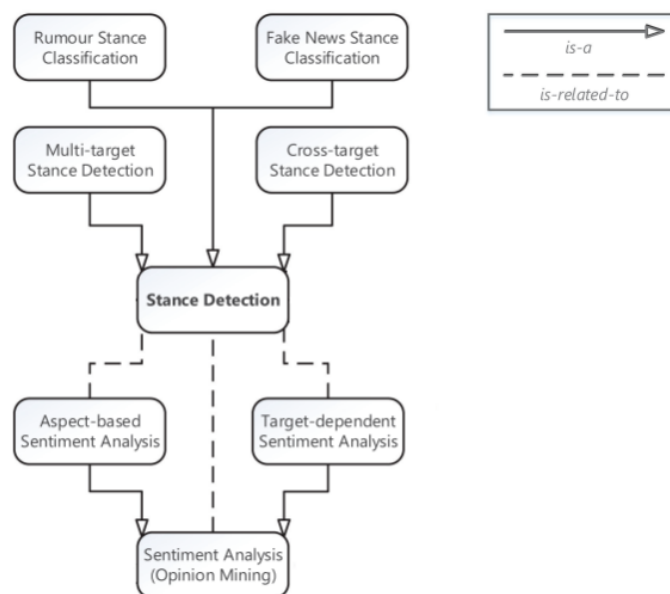
به عنوان مثال و برای درک بهتر جمله و هدف جدول ۲-۱ را در نظر بگیرید [۴۰]. نویسنده متن، با متن نوشته شده نشان می‌دهد با کلینتون موافق است. در حالی که از نظر احساسات، متن احساس منفی منتقل می‌کند. بنابراین می‌توان نتیجه گرفت دو مسئله تحلیل احساسات و تشخیص موضع به هم مرتبط هستند اما یکسان نیستند [۴۹].

پس به طور خلاصه دو تفاوت عمده مسئله تشخیص موضع و تحلیل احساسات شامل موارد زیر می‌شود. (۱) در تشخیص موضع، موضوع که نسبت به آن بررسی‌ها انجام می‌شود به طور کامل مشخص است. (۲) مانند مثال بالا برجسب‌های نهایی این دو مسئله می‌توانند کاملاً با هم متفاوت باشد. بنابراین تفاوت این دو مسئله کاملاً مشهود است.

تحلیل احساسات جنبه گرا<sup>۲۵</sup> شبیه ترین مسئله موجود به تشخیص موضع می‌باشد. در این نوع از تحلیل احساسات موضوع معمولاً شامل محصولات الکترونیکی (لپتاپ)، رستوران و هتل می‌باشد. جنبه‌های مورد بررسی نیز شامل قیمت، کیفیت، طرح می‌باشد. تفاوت تحلیل احساسات جنبه گرا با تشخیص موضع در این

<sup>25</sup> Aspect-Oriented Sentiment Analysis

است که در تحلیل احساسات موضوعی که مورد بررسی قرار می گیرد در متن به صورت دقیق ذکر شده است. اما در تشخیص موضع لزوما این طور نیست. همچنین موضوع مورد بررسی در تشخیص موضع می تواند یک رویداد باشد ولی در تحلیل احساسات جنبه گرا این طور نیست.



شکل ۲-۲: موضوعات تحقیقاتی مرتبط با تشخیص موضع [۲۵]

## ۵-۲-۲ کاربردهای تشخیص موضع

بعد از مطرح شدن تعاریف مختلف مسئله تشخیص موضع، حال به بررسی تعدادی از کاربردهای مطرح شده می پردازیم. در مسئله تشخیص موضع، نظر و موضع افراد نسبت به هدف یا موضوع خاص سنجیده می شود. با توجه به این تعریف کاربردهای زیر مطرح می شود.

۱. نظرسنجی و بررسی افکار عمومی: مطالعات تشخیص موضع معمولاً بر روی محتوا آنلاین انجام می شود که موضوع آن مشخص است. از این رو با استفاده از تشخیص موضع خودکار، موافقت یا مخالف عموم افراد جامعه نسبت به موضوع خاصی را می توان ارزیابی کرد. این روش می تواند جایگزین نظرسنجی های سنتی باشد.

۲. سیستم های توصیه گر: در صورتی که موضع افراد نسبت به هدف و موضوعی معین را بدانیم، توصیه



کردن محصولات به سادگی میسر می شود.

۳. تبلیغات هدفمند: آگاهی از موضع کاربران باعث می شود تبلیغات به صورت موثر و هدفمند انجام شود.

۴. تشخیص اخبار جعلی [۴۳] و تشخیص شایعه نیز از دیگر کاربردهای قابل تعریف به کمک مسئله تشخیص موضع می باشند. در این مسائل معمولا از تشخیص موضع ادعا محور استفاده می شود.

## فصل ۳

# تعاریف و مفاهیم ابتدایی

در این فصل، توضیحات مختصری در ارتباط با مفاهیم و تعاریف اولیه ارائه شده است. توضیحات ارائه شده، پیش زمینه‌ای برای فهم بهتر مسئله تشخیص موضع در شبکه‌های اجتماعی می‌باشد.

### ۳-۱ شبکه‌های اجتماعی

شبکه‌های اجتماعی<sup>۱</sup> یک فناوری مبتنی بر رایانه است که قابلیت به اشتراک گذاری ایده‌ها، افکار و اطلاعات را از طریق شبکه‌ها و جوامع مجازی فراهم می‌کند. شبکه‌های اجتماعی معمولاً توسط مجموعه‌ای از افراد شکل گرفته است که دیدگاه‌های خود را بیان کرده و در رابطه با آن تبادل نظر می‌کنند.

امروزه، بسیاری از تعاملات اجتماعی افراد در بستر شبکه‌های اجتماعی انجام می‌شود. مردم برای برقراری ارتباطات، اطلاع از اخبار و بررسی دیدگاه افراد دیگر از این رسانه‌ها استفاده می‌کنند [۱۰]. در شبکه‌های اجتماعی هر فرد می‌تواند آزادانه دیدگاه خود را مطرح کنند. همچنین از دیدگاه سایر افراد نیز مطلع می‌شود. علاوه بر همه موارد کشف دیدگاه عموم جامعه با رصد شبکه‌های اجتماعی تا حد خوبی امکان پذیر است. استفاده گروه عظیمی از کاربران از شبکه‌های اجتماعی این فرصت را فراهم می‌کند تا جنبه‌های مختلف رفتار انسان از جمله موضع گیری عمومی نسبت به جنبه‌های مختلف اجتماعی و سیاسی، قابل تحلیل باشد [۴].

---

<sup>1</sup>Social Media

## ۲-۳. هوش مصنوعی

هوش مصنوعی<sup>۲</sup> به فرآیند هوشمندسازی کامپیوتر برای یادگیری و تفکر گفته می‌شود. در این فرآیند هدف این است کامپیوتر همچون انسان توانایی حل مسئله و تصمیم‌گیری را یاد بگیرد. این اصطلاح عموماً به پروژه‌هایی اطلاق می‌شود که توانایی استدلال، یادگیری از تجربیات گذشته و کشف معنا و الگو را دارند. همچنین با کمک اطلاعاتی که جمع می‌کند مکرراً می‌تواند توانایی خود را بهبود دهد. سیستم‌های توصیه‌گر<sup>۳</sup>، دستیار هوشمند<sup>۴</sup>، موتورهای جستجو و تشخیص صدا یا دستخط از جمله کاربردهایی هستند که از دانش هوش مصنوعی استفاده می‌کنند.

## ۳-۳. پردازش زبان طبیعی

پردازش زبان طبیعی<sup>۵</sup> یکی از شاخه‌های هوش مصنوعی می‌باشد. هدف این است که کامپیوتر بتواند زبان انسان را درک و تفسیر کند. این فرآیند شامل تبدیل گفتار به متن، آموزش مدل برای تصمیم‌گیری و انجام اقدامات هوشمندانه است. پردازش زبان طبیعی بر روی داده‌های بدون ساختار کار می‌کند و به عوامل مختلفی از جمله دستور زبان، لحن و احساسات وابسته است [۱۳].

برای درک بهتر زبان انسان توسط کامپیوتر مراحل مختلفی از جمله تحلیل واژگانی<sup>۶</sup>، تحلیل نحوی<sup>۷</sup> و تحلیل معنایی<sup>۸</sup> به کار گرفته می‌شود. با توجه به پیچیده بودن فهم زبان انسان، زیر شاخه پردازش زبان طبیعی نیز یکی از پیچیده‌ترین زیرشاخه‌های هوش مصنوعی در نظر گرفته می‌شود. ترجمه ماشینی<sup>۹</sup>، خلاصه‌سازی متون<sup>۱۰</sup> و تحلیل احساسات<sup>۱۱</sup> از جمله کاربردهای پردازش زبان طبیعی می‌باشد. همچنین مترجم گوگل<sup>۱۲</sup> یکی از محبوب‌ترین ابزارهایی می‌باشد که در آن از الگوریتم‌های پردازش زبان طبیعی استفاده شده است.

<sup>۲</sup> Artificial Intelligence<sup>۳</sup> Recommendation System<sup>۴</sup> Intelligent Assistants<sup>۵</sup> Artificial Intelligence<sup>۶</sup> Lexical Analysis<sup>۷</sup> Syntactical Analysis<sup>۸</sup> Semantic Analysis<sup>۹</sup> Machine Translation<sup>۱۰</sup> Text Summarization<sup>۱۱</sup> Sentiment Analysis<sup>۱۲</sup> Google Translate

## ۳-۴. تعبیه کلمات

در پردازش زبان طبیعی نمی‌توان متن را به صورت ساده و خام به عنوان ورودی به الگوریتم داد. بنابراین نیاز است کلمات در قالب بردارهای عددی قابل فهم و پردازش توسط کامپیوتر به آن وارد شوند. تعبیه کلمات<sup>۱۳</sup> به فرآیند نگاشت کلمات یا عبارات به بردارهای عددی قابل پردازش توسط کامپیوتر گفته می‌شود. این روش به طور معمول برای مدل‌سازی زبان و آماده‌سازی متن جهت استفاده از آن در الگوریتم‌های پردازش زبان طبیعی استفاده می‌شود. رویکردهای مختلفی برای بازنمایی کلمات وجود دارد که در ادامه معروف‌ترین روش‌ها را مورد بررسی قرار خواهیم داد.

## ۱. کدگذاری one-hot

روش کدگذاری one-hot ساده‌ترین و ابتدایی‌ترین روش استفاده شده برای تعبیه کلمات است. در این روش طول بردار تولید شده برای هر کلمه، برابر با تعداد کلمات یکتا در مجموعه دادگان مورد بررسی می‌باشد. در بردارهای تولید شده توسط این روش، یکی از درایه‌های بردار مقدار یک و بقیه درایه‌ها مقدار صفر می‌گیرند. پیاده‌سازی این روش بسیار آسان است. اما در صورت بزرگ بودن مجموعه داده، طول بردار بازنمایی کلمات بسیار بزرگ خواهد بود. همچنین عیب دیگر این روش این است که مفهومی از کلمات را در خود جای نداده است. به این دلیل که فاصله کلیه کلمات مجموعه دادگان از یکدیگر یکسان است. بنابراین بازنمایی تولید شده، فرقی مابین کلمات مترادف و یا کلمات بی ربط ایجاد نمی‌کند. در صورتی که انتظار ما این است فاصله بردار کلمات مترادف از یکدیگر نسبت به کلمات بی ربط به یکدیگر متفاوت باشد.

## ۲. کدگذاری TF-IDF

روش TF-IDF<sup>۱۴</sup> یکی دیگر از روش‌های مورد استفاده است. هدف از این روش، تنظیم کردن و یک دست کردن کلماتی است که بارها در متن تکرار شده‌اند. متن می‌تواند شامل یک سند یا مجموعه‌ای از اسناد مختلف باشد. این روش ترکیبی از دو معیار است.

● TF : که عبارت است از تقسیم تعداد تکرار کلمه بر کل کلمات محتوا.

<sup>13</sup>Machine Learning

<sup>14</sup>Term Frequency-Inverse Document Frequency

● IDF: فراوانی سند معکوس. عبارت است از لگاریتم تقسیم تعداد کل اسناد موجود بر محتوایی که شامل کلمه مورد نظر می‌شوند.

لازم به ذکر است TF-IDF نمی‌تواند مفهوم کلمه در یک جمله را به خوبی نمایش دهد. در واقع کلمات به بردارهایی تبدیل می‌شوند ولی توجهی به معنی کلمه در جمله نمی‌شود.

### ۳. کدگذاری Word2Vec

روش Word2Vec [۳۹] یکی دیگر از روش‌های مورد استفاده برای تعبیه کلمات است. این روش کل پیکره را به عنوان ورودی می‌گیرد و کلمات را به یک بردار چند بعدی نگاشت می‌کند. برخلاف روش one-hot در این روش بردارهای تولید شده مفاهیمی از کلمه را در خود جای داده‌اند. بنابراین کلمات مشابه، بردارهای مشابه دارند. این الگوریتم دو مدل مبتنی بر شبکه عصبی ارائه می‌دهد که در ادامه هر کدام را بررسی می‌کنیم.

● CBOW: در این روش کلمات اطراف و نزدیک به یک کلمه به عنوان ورودی به شبکه داده می‌شود و هر کلمه میانی به عنوان کلمه هدف در نظر گرفته می‌شود. وظیفه مدل این است با توجه به کلمات اطراف، بردار مناسب برای کلمه هدف (کلمه میانی) را تولید کند.

● Skip-Gram: این روش دقیقاً برعکس روش CBOW می‌باشد. در این روش شبکه عصبی یک کلمه را به عنوان ورودی می‌گیرد و وظیفه دارد چند کلمه قبل و چند کلمه بعد از کلمه ورودی را پیشبینی کند.

### ۴. کدگذاری GloVe

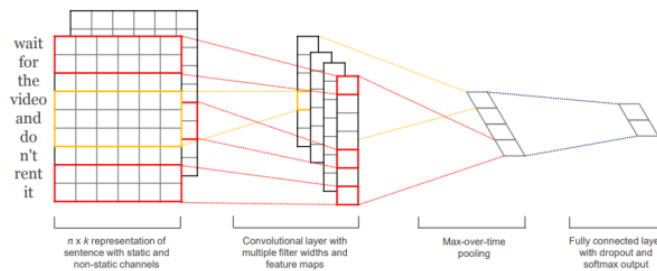
یکی دیگر از روش‌های تولید بردار تعبیه کلمات GloVe<sup>۱۵</sup> [۴۲] می‌باشد. این روش در سال ۲۰۱۴ در تیم پردازش زبان طبیعی استنفورد معرفی شده است. این روش به صورت بدون ناظر با تجمیع ماتریس هم‌زمانی کلمه به کلمه سراسری<sup>۱۶</sup> یک پیکره، یک بردار جانمایی برای هر کلمه تولید می‌کند.

### ۵. تولید بازنمایی متن با استفاده از CNN, GRU, LSTM

<sup>۱۵</sup>Global Vector

<sup>۱۶</sup>Global word-word co-occurrence matrix

با پیشرفت یادگیری عمیق در دهه‌های اخیر، پژوهشگران برای بازنمایی متن و استخراج ویژگی از شبکه‌های CNN [۲۴]، LSTM و GRU نیز استفاده می‌کنند. برای استخراج ویژگی با استفاده از CNN بردارهای کلمات در کنار یکدیگر قرار می‌گیرند. سپس به لایه پیچشی<sup>۱۷</sup> یک بعدی داده می‌شوند و طی مراحل مختلف فیلترهای مختلفی بر روی آن‌ها اعمال می‌شود. در انتها پس از عبور از لایه max-pooling ویژگی‌های مورد نظر از متن به دست می‌آید. بدین ترتیب از شبکه CNN در پردازش زبان طبیعی برای استخراج ویژگی از متن استفاده می‌شود. شکل ۳-۱ فرآیند تولید بازنمایی متن با استفاده از شبکه CNN را نشان می‌دهد. برای تولید بازنمایی متن توسط LSTM و GRU کافیست جمله (عبارت) مورد نظر به عنوان ورودی به شبکه داده شود. خروجی آخرین گام زمانی به عنوان بازنمایی جمله (عبارت) ورودی در نظر گرفته می‌شود.



شکل ۳-۱: تولید بازنمایی متن با استفاده از شبکه CNN [۲۴]

## ۶. کدگذاری BERT

مدل BERT [۱۸] در سال ۲۰۱۹ توسط گوگل معرفی شد. BERT یک مدل از پیش آموزش دیده برای مسائل پردازش زبان طبیعی می‌باشد. از این مدل برای تولید بازنمایی متن با کیفیت بالا می‌توان استفاده کرد. به دلیل اینکه شبکه BERT بر روی حجم بسیار داده آموزش دیده، بازنمایی بهتری از کلمه تولید می‌کند. همچنین با آموزش مجدد لایه آخر، برای وظایف متفاوت قابل استفاده است. مدل BERT مبتنی بر ترنسفورمر و شامل رمزگذار دو طرفه عمیق می‌باشد.

انواع دیگری از مدل‌های مبتنی بر BERT نیز معرفی شدند. مدل RoBERTa [۳۴] با الگوبرداری از مدل BERT آموزش دیده است. تفاوت اصلی RoBERTa با BERT در ابرپارامترهای استفاده شده

<sup>17</sup>Convolution Layer

است. در RoBERTa اندازه دسته و نرخ یادگیری بزرگتر می‌باشد. برخلاف BERT در آموزش اولیه RoBERTa از مسئله next-sentence prediction استفاده نمی‌شود. مدل BERTweet [۴۱] با روش آموزشی RoBERTa بر داده‌های توئیتر انگلیسی آموزش دیده است. DEBERTa [۳۴] مدل BERT و RoBERTa را با کمک توجه از هم گسسته، بهبود می‌بخشد. XLM-RoBERTa [۱۷] یک مدل چندزبانه است که بر داده‌های CommonCrawl آموزش دیده است.

### ۳-۵ یادگیری ماشین

یادگیری ماشین<sup>۱۸</sup> یکی از شاخه‌های مهم هوش مصنوعی می‌باشد. در این روش هدف یادگیری یک مدل ریاضی با استفاده از داده‌های موجود به جای تعیین صریح قوانین در مسئله می‌باشد. بنابراین از الگوریتم‌هایی برای یادگیری الگوها و هم بستگی‌های موجود در داده‌ها استفاده می‌شود. در نهایت با استفاده از الگوهای به دست آمده یک مدل ریاضی تعریف می‌شود که از آن برای تصمیم‌گیری و پیشبینی استفاده می‌شود. هر چه به داده‌های بیشتری دسترسی داشته باشیم، الگوهای بهتر و مطمئن‌تری از داده‌ها استخراج کرده و بنابراین پیشبینی‌های دقیق‌تری خواهیم داشت.

### ۳-۶ یادگیری عمیق

یادگیری عمیق<sup>۱۹</sup> شاخه‌ای از یادگیری ماشین است که مبتنی بر شبکه‌های عصبی مصنوعی<sup>۲۰</sup> می‌باشد. فرآیند یادگیری، عمیق نامیده می‌شود چرا که از چندین نورون ورودی و خروجی و چندین لایه پنهان ساخته شده است.

در حالی که الگوریتم‌های یادگیری ماشین قادر هستند بسیاری از مسائل پیچیده را حل کنند، اما همچنان در برخورد با داده‌های غیرساختاری از جمله متن، عکس و فیلم عملکرد ضعیفی دارند. روش‌های یادگیری عمیق با استفاده از شبکه‌های عصبی عمیق (که یک نوع معماری الهام گرفته از مغز انسان می‌باشد)، این مشکل را حل می‌کنند. هر لایه از شبکه عمیق ویژگی‌هایی از داده‌های ورودی استخراج می‌کند که در نهایت

<sup>18</sup> Machine Learning

<sup>19</sup> Deep Learning

<sup>20</sup> Artificial Neural Networks

می‌توان توسط ویژگی‌های استخراج شده، مسئله مورد نظر را حل کرد. اطلاعات ورودی از طریق لایه‌ها به جلو حرکت می‌کند.

### ۷-۳ شبکه‌های عصبی همگشتی

شبکه‌های عصبی بازگشتی<sup>۲۱</sup> نوعی از شبکه عصبی مصنوعی است که در تشخیص گفتار، پردازش زبان طبیعی، پردازش داده‌های ترتیبی<sup>۲۲</sup> و داده‌های سری زمانی<sup>۲۳</sup> مورد استفاده قرار می‌گیرد. معماری این شبکه به گونه‌ای است که امکان ذخیره اطلاعات قبلی را در خود فراهم می‌کند. در بسیاری از شبکه‌های عمیق مانند شبکه عصبی پیشخور<sup>۲۴</sup> داده‌ها در یک جهت حرکت می‌کنند. یعنی از لایه ورودی به لایه‌های پنهان و سپس به سمت لایه خروجی حرکت می‌کنند و داده‌های قبلی به حافظه سپرده نمی‌شوند. در حالی که شبکه‌های عصبی بازگشتی یک لایه بازخورد دارند که خروجی شبکه به همراه ورودی جدید، به شبکه وارد می‌شوند. این شبکه‌ها اطلاعات مربوط به ورودی قبلی را برای تاثیر بر روی ورودی و خروجی فعلی در حافظه خود ذخیره می‌کنند. در حالی که در شبکه‌های عصبی عمیق خروجی‌های مختلف از یکدیگر مستقل هستند، خروجی شبکه عصبی بازگشتی علاوه بر ورودی فعلی به عناصر قبلی توالی نیز وابسته است.

مشکل اصلی شبکه عصبی بازگشتی حافظه کوتاه مدت آن است. در بلند مدت این شبکه توانایی نگهداشتن اطلاعاتی که در گام‌های زمانی بسیار قبل‌تر به عنوان ورودی به شبکه داده شده است را ندارد. علت این مشکل پدیده محو شدگی گرادیان<sup>۲۵</sup> می‌باشد. این پدیده بیان می‌کند در فرآیند بازگشت خطا به عقب<sup>۲۶</sup> زمانی که به ابتدا شبکه می‌رسیم، گرادیان به تدریج کوچک می‌شود. در نتیجه تغییرات وزن بسیار کند و ناچیز می‌شود و به عبارت دیگر آموزش کند می‌شود. این مسئله معمولاً به خاطر عمق زیاد شبکه اتفاق می‌افتد. معماری‌های دیگری از جمله LSTM [۲۲] و GRU [۱۴] برای رفع این مشکل پیشنهاد شده است.

ویژگی متمایز LSTM<sup>۲۷</sup> قابلیت یادگیری وابستگی بلند مدت است که توسط شبکه‌های عصبی بازگشتی امکان پذیر نبود. شبکه عصبی بازگشتی تنها قادر به یادگیری تعداد محدودی از وابستگی‌های کوتاه مدت بود.

<sup>21</sup> Recurrent Neural Network (RNN)

<sup>22</sup> Sequential Data

<sup>23</sup> Time Series

<sup>24</sup> Feedforward Neural Network

<sup>25</sup> Vanishing Gradient

<sup>26</sup> Backpropagation

<sup>27</sup> Long Short-Term Memory



برخلاف شبکه عصبی بازگشتی که در آن محتوا در هر گام از ابتدا بازنویسی می‌شود، در معماری LSTM شبکه قادر است در یک گام زمانی محتوا حافظه را بدون تغییر به گام بعدی منتقل کند. به عنوان مثال اگر در گام‌های ابتدایی شبکه اطلاعات مهمی را تشخیص دهد، می‌تواند آن را برای گام‌های طولانی بدون تغییر در حافظه خود ذخیره کند.

معماری GRU<sup>۲۸</sup> در سال ۲۰۱۴ معرفی شد. معماری GRU نسخه ساده شده LSTM می‌باشد. GRU علاوه بر رفع مشکل محو شدگی گرادیان در شبکه عصبی بازگشتی، سربار محاسباتی LSTM را نیز کاهش می‌دهد. همچنین با کم کردن تعداد دروازه‌ها، سرعت محاسبات را افزایش داده است.

### ۳-۸ ساز و کار توجه

هدف استفاده از ساز و کار توجه<sup>۲۹</sup> بازیابی اطلاعات از بردارهای زمینه<sup>۳۰</sup>  $y_j$  در رابطه با بردار پرس‌وجو<sup>۳۱</sup>  $x$  می‌باشد. ساز و کار توجه ابتدا امتیاز  $\alpha_j$  را بین بردار پرس‌وجو  $x$  و بردار زمینه  $y_j$  محاسبه می‌کند.

$$a_j = \text{Score}(x, y_j) \quad (۱-۳)$$

$$\alpha_j = \frac{\exp(a_j)}{\sum_k \exp(a_k)} \quad (۲-۳)$$

خروجی لایه توجه میانگین وزن‌دار امتیاز  $\alpha_j$  به ازای بردارهای زمینه می‌باشد. محاسبات انجام شده مشابه لایه softmax است. اگر بردار پرس‌وجو  $x$  مجموعه‌ای از بردار زمینه  $\{y_j\}$  باشد، امتیاز به دست آمده از معادله ۳-۳ توجه به خود<sup>۳۲</sup> نامیده می‌شود.

$$\text{Att}_{x \rightarrow \{y_j\}} = \sum_j \alpha_j y_j \quad (۳-۳)$$

<sup>۲۸</sup>Gated Recurrent Unit (GRU)

<sup>۲۹</sup>Attention

<sup>۳۰</sup>Context Vector

<sup>۳۱</sup>Query Vector

<sup>۳۲</sup>self-attention

## ۳-۹ ترنسفورمر

ترنسفورمر [۵۷]، جدیدترین معماری معرفی شده برای مسائل پردازش زبان طبیعی می‌باشد. این معماری در سال ۲۰۱۷ معرفی شد. ترنسفورمرها با استفاده از مکانیزم توجه، امکان پردازش موازی داده‌ها را فراهم می‌کنند. با استفاده از ترنسفورمرها برای هر کلمه می‌توان بردار بازنمایی تولید کرد. همچنین به دلیل آموزش موازی، سرعت آموزش نیز بهبود قابل توجهی پیدا می‌کند. ترنسفورمرها این قابلیت را برای ما فراهم می‌کند که به جای پردازش دنباله ورودی از ابتدا تا انتها، تنها به قسمت‌های مهم توجه کنیم. این مدل‌ها معمولاً بر روی حجم بالایی از داده‌ها از پیش آموزش می‌بینند.<sup>۳۳</sup>

مدل‌های زبانی بزرگ<sup>۳۴</sup> دسته‌ای از مدل‌های زبانی هستند که توانایی درک و تولید متنی شبیه انسان را دارند. این مدل‌ها معمولاً شامل میلیون‌ها تا میلیارد‌ها پارامتر قابل آموزش می‌باشند. برخی از آن‌ها با معماری مبتنی بر ترنسفورمرها ساخته شده‌اند. LLMs ها بر روی میلیون‌ها داده آموزش دیده‌اند تا بتوانند تا بتوانند زبان انسان و انواع دیگر داده‌های پیچیده را تفسیر کنند. برخی از معروف‌ترین مدل‌های زبانی بزرگ شامل GPT-3، LLaMA [۵۳] و Orca [۳۸] می‌باشند. Orca یک نمونه از مدل زبانی کوچک است که در دو نسخه ۷ میلیارد و ۱۳ میلیارد پارامتری عرضه شده است. این مدل با fine-tuning مدل LLaMA بر داده‌های مصنوعی با کیفیت بالا ایجاد شده است.

## ۳-۱۰ مهندسی پرامپت

پرامپت یک دستور متنی است که وظیفه مدل را توصیف می‌کند. جدول ۳-۱ نمونه‌ای از دستور متنی برای رده‌بندی در مسئله تحلیل احساسات را نشان می‌دهد (بخش‌های مختلف پرامپت نیز مشخص شده است).

جدول ۳-۱: نمونه پرامپت برای مسئله تحلیل احساسات

خروجی	ورودی	دستورالعمل
Sentiment:	Text: I think the food was okay.	Classify the text into neutral, negative, or positive.

مهندسی پرامپت<sup>۳۵</sup> توسعه و بهینه‌سازی پرامپت برای استفاده کارآمد از مدل‌های زبانی بزرگ است. این مهارت شامل طراحی یک ورودی (دستورالعمل) مناسب برای LLMs می‌باشد به گونه‌ای که مدل بتواند بهترین

<sup>۳۳</sup>Pre-train

<sup>۳۴</sup>Large Language Models (LLMs)

<sup>۳۵</sup>Prompt Engineering

خروجی را تولید کند. مهندسی پرامپت با یادگیری درون متنی<sup>۳۶</sup> معنا پیدا می‌کند. منظور از یادگیری درون متنی این است که مدل توانایی یادگیری موقت از پرامپت را داشته باشد. برای طراحی یک پرامپت کارآمد رویکردهای متفاوتی وجود دارد که در ادامه معروف‌ترین رویکردها را معرفی می‌کنیم<sup>۳۷</sup>.

۱. رویکرد Zero-Shot: در این رویکرد فقط یک دستورالعمل به مدل داده می‌شود (مشابه جدول ۱-۳).

۲. رویکرد Chain of thought: این رویکرد توانایی استدلال پیچیده را از طریق استدلال میانی برای مدل فراهم می‌کند. اضافه کردن "Let's think step by step" به انتهای متن یکی از معمول‌ترین دستورالعمل‌های این تکنیک به حساب می‌آید.

۳. رویکرد Few-Shot: در این رویکرد علاوه بر دستورالعمل برای انجام مسئله تعریف شده، تعدادی مثال از ورودی و خروجی مطلوب نیز در پرامپت گنجانده می‌شود.

## ۱۱-۳ جستجو پارامترها

مدل‌های یادگیری ماشین شامل پارامترها و فراپارامترهایی<sup>۳۸</sup> هستند که می‌توانند مقادیر مختلفی داشته باشند. انتخاب مقادیر این پارامترها همواره یکی از چالش‌های یادگیری به حساب می‌آید. روش‌های مختلفی برای انتخاب پارامترها وجود دارد. ساده‌ترین رویکرد برای انتخاب پارامترها جستجو تصادفی<sup>۳۹</sup> می‌باشد. در این رویکرد برای هر یک از پارامترها یک فضا جستجو تعریف می‌شود. در هر بار اجرا به صورت کاملاً تصادفی برای هر پارامتر مقداری انتخاب می‌شود. این روش لزوماً بهترین مقادیر را هر پارامتر را به دست نمی‌آورد. روش جستجو شبکه‌ای<sup>۴۰</sup> تمام حالت‌های ممکن فضا تعریف شده را جستجو می‌کند. مشکل اصلی این روش نفرین ابعاد<sup>۴۱</sup> می‌باشد. به این معنی که هر چه ابعاد فضا جستجو بزرگ‌تر شود، پیچیدگی زمانی به صورت نمایی افزایش پیدا خواهد کرد.

برخی از روش‌های بهینه‌سازی فراپارامترها در هر بار انتخاب فراپارامترها برای آزمایش بعدی، از اطلاعات

<sup>36</sup>In-context learning

<sup>37</sup><https://www.promptingguide.ai/>

<sup>38</sup>Hyperparameter

<sup>39</sup>Random Search

<sup>40</sup>Grid Search

<sup>41</sup>Curse of Dimensionality

ترکیب قبلی نیز استفاده می‌کنند. این روش‌های انتخاب تطبیقی<sup>۴۲</sup>، با انتخاب فراپارامترهایی که احتمال موفقیت‌شان بیشتر است، سرعت جستجو را به طور قابل توجهی افزایش می‌دهند. یکی از ابزارهای معرفی شده در این روش، Optuna [۳] می‌باشد. این رویکرد فضا جستجو پارامترها را به صورت پویا ساختاردهی می‌کند. برای انتخاب هر نمونه از فضا جستجو از الگوریتم TPE<sup>۴۳</sup> که بر مبنای بهینه‌سازی Bayesian است استفاده می‌کند. همچنین Optuna با قابلیت توقف زود هنگام<sup>۴۴</sup> با استفاده از روش‌های Pruning احتمال رسیدن به پارامترهای بهینه در کوتاه‌ترین زمان ممکن را افزایش می‌دهد.

## ۳-۱۲ جمع بندی

در این فصل، ابتدا شبکه اجتماعی و پردازش زبان طبیعی معرفی شد. سپس مکانیزم‌های تعبیه کلمات که یکی از مهم‌ترین روش‌های کار با متن است مورد بررسی قرار گرفت. همچنین مفاهیم یادگیری عمیق، مهندسی پرامپت و روش‌های جستجو به عنوان مهم‌ترین رویکرد حل مسئله بررسی شد. آشنایی با این مفاهیم، درک مسئله تشخیص موضع را آسان تر می‌کند.

---

<sup>42</sup> Adaptive Selection

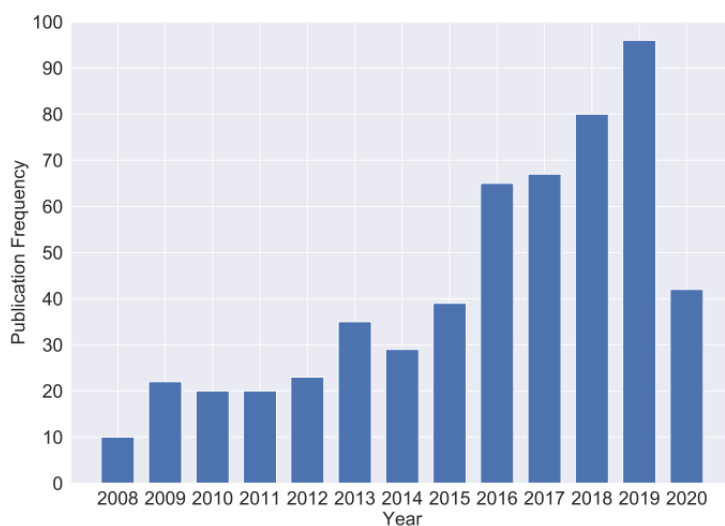
<sup>43</sup> Tree-Structured Parzen Estimator

<sup>44</sup> Early Stopping

## فصل ۴

### مروری بر کارهای مرتبط

در سال‌های اخیر تحقیقات در تشخیص موضع افزایش چشم‌گیری داشته است (شکل ۴-۱). در این بخش ابتدا رقابت‌های برگزار شده در زمینه تشخیص موضع را مورد بررسی قرار می‌دهیم. سپس تعدادی از مجموعه داده‌های معرفی شده را بررسی می‌کنیم. در ادامه رویکردهای معرفی شده برای تشخیص موضع را معرفی می‌کنیم.



شکل ۴-۱: نمودار تعداد مقالات چاپ شده در رابطه با تشخیص موضع به تفکیک سال. جستجو براساس کلمات کلیدی «تشخیص موضع»، «پیش بینی موضع» و «رده‌بندی موضع» انجام شده است [۴].

## ۴-۱ رقابت‌های برگزار شده در مسئله تشخیص موضع

با بررسی منابع موجود، مشخص شد تا کنون شش رقابت در رابطه با مسئله تشخیص موضع برگزار شده است. در این رقابت‌ها مجموعه داده‌های جدید و مدل‌های مرتبط با آن‌ها معرفی می‌شوند. بنابراین برگزاری چنین رویدادهایی نقش مهمی در رشد و تقویت تحقیقات در تشخیص موضع دارند [۲۵]. مهم‌ترین رویدادهای برگزار شده به شرح زیر است.

### ۴-۱-۱ رویداد SemEval-2016

اولین رقابت برگزار شده برای مسئله تشخیص موضع، رقابت **SEMEval2016:Task6**<sup>۱</sup> می‌باشد. در این رقابت چندین مسئله تعریف شده که مسئله ششم مرتبط با تشخیص موضع می‌باشد. این رقابت شامل دو زیر مسئله است [۴۰].

#### ۱. تشخیص موضع با نظارت<sup>۲</sup>

مجموعه داده دارای برچسب برای ۵ موضوع<sup>۳</sup> که دارای ۲۸۱۴ توییت در قسمت داده‌های آموزشی و ۱۲۴۹ توییت در داده‌های ارزیابی می‌باشد.

#### ۲. تشخیص موضع با نظارت ضعیف<sup>۴</sup>

شامل حدوداً ۷۸۰۰۰ توییت بدون برچسب به عنوان داده‌های آموزشی می‌باشد. داده‌های ارزیابی شامل ۷۰۷ توییت در رابطه با هدف دیگری می‌باشد.

در مجموع ۱۹ گروه در این رقابت شرکت کردند. شرکت کنندگان در این مسابقه از روش‌های مختلفی استفاده کردند. روش‌های استفاده شده شامل یادگیری ماشین سنتی مبتنی بر ویژگی<sup>۵</sup>، یادگیری عمیق و یادگیری مجمع<sup>۶</sup> می‌باشد. بهترین مدل معرفی شده برای زیر مسئله اول (تشخیص موضع با نظارت)، از یادگیری مجمع شبکه عصبی بازگشتی (RNN) استفاده کرده است. در ارزیابی نهایی، این روش به مقدار

<sup>۱</sup><https://alt.qcri.org/semeval2016/task6/>

<sup>۲</sup>Supervised Stance Detection

<sup>۳</sup>Target

<sup>۴</sup>Weakly Supervised Stance Detection

<sup>۵</sup>Feature-Based Machine Learning

<sup>۶</sup>Ensemble Learning

67.82% برای معیار F1-Score دست یافت [۶۵]. همچنین بهترین مدل (در قیاس با ۹ روش معرفی شده) برای زیر مسئله دوم (تشخیص موضع با نظارت ضعیف) از شبکه عصبی پیچشی (CNN) استفاده کرد که در ارزیابی پایانی مقدار 56.28% برای معیار F1-Score دست یافت [۶۱]. علاوه بر این، مدل مبتنی بر شبکه عصبی پیچشی (CNN) با مقدار 67.33% برای معیار F1-Score جایگاه دوم را برای زیر مسئله اول کسب کرد. مدل پایه‌ای که توسط طراحان مسابقه توسعه یافته از روش SVM استفاده کرده است. در ارزیابی انتهایی به مقدار 68.98% برای معیار F1-Score دست یافت. این بالاترین مقدار در بین همه شرکت کنندگان در رویداد می‌باشد.

## ۴-۱-۲ رویداد NLPCC-ICCPOL-2016

رقابت NLPCC-ICCPOL-2016 تشخیص موضع بر روی وبلاگ‌های کوتاه به زبان چینی است [۶۴]. در این رقابت چندین مسئله تعریف شده که مسئله چهارم مرتبط با تشخیص موضع می‌باشد. در این رقابت نیز دو زیر مسئله مشابه با SemEval-2016 تعریف شد.

### ۱. تشخیص موضع با نظارت

حدود ۴۰۰۰ وبلاگ کوتاه به صورت دستی<sup>۷</sup> برای ۵ موضوع حاشیه نویسی شده است. موضوعات شامل SE iPhone، ترقی در جشنواره بهار، عملیات ضد تروریستی روسیه در سوریه، سیاست دو فرزند و ممنوعیت موتورسیکلت و محدودیت وسایل نقلیه الکترونیکی در شترن می‌باشد. از این میان 75% داده‌ها برای آموزش و مابقی داده‌ها برای ارزیابی استفاده می‌شوند.

### ۲. تشخیص موضع بدون ناظر<sup>۸</sup>

در این روش از مجموعه داده بدون برچسب برای آموزش بدون ناظر استفاده می‌شود. داده‌های این قسمت شامل ۲۴۰۰ وبلاگ کوتاه بدون برچسب است. این داده‌ها شامل دو موضوع موارد غذایی اصلاح شده ژنتیکی و آزمایش‌های هسته‌ای کره شمالی می‌باشند.

در این رقابت شانزده تیم در زیر مسئله اول (تشخیص موضع با نظارت) و پنج تیم در زیر مسئله دوم (تشخیص موضع بدون ناظر) شرکت کردند. برنده رقابت در زیر مسئله اول، یک مدل رده بندی جداگانه برای

<sup>7</sup>Manual

<sup>8</sup>Unsupervised Stance Detection

هر موضوع آموزش داده است. پایه مدل‌های استفاده شده SVM و Forest Random هستند. بالاترین مقدار F1-Score به دست آمده برابر با 71.06% می‌باشد.

ویژگی‌های ورودی بهترین سیستم شامل Unigram، TF-IDF<sup>۹</sup>، کلمات مشابه و بردارهای بازنمایی در سطح کلمه و کاراکتر می‌باشد. سایر ویژگی‌های استفاده شده توسط سایر شرکت کنندگان شامل Bigram و واژگان شامل احساسات<sup>۱۰</sup> می‌باشد. همچنین با بررسی نتایج شرکت کنندگان این نتیجه حاصل شد که نتیجه بهتر در مسئله تحلیل احساسات لزوماً به این معنا نیست که مدل می‌تواند در مسئله تشخیص موضع نیز به نتایج خوبی دست پیدا کند. بالاترین نتیجه کسب شده برای زیر مسئله دوم (تشخیص موضع بدون ناظر) برابر با 46.87% است.

#### ۴-۱-۳ رویداد IberEval-2017

رویداد IberEval-2017 مشابه با دو رویداد قبلی می‌باشد. یکی از مسائل تعریف شده در این رقابت مربوط به تشخیص موضع و تشخیص جنسیت در زبان اسپانیایی و کاتالان می‌باشد [۵۱].

روش‌های متداول استفاده شده توسط شرکت کنندگان شامل SVM، شبکه عصبی، روش‌های یادگیری عمیق از جمله LSTM می‌باشد. همچنین ویژگی‌های ngram و بازنمایی کلمات<sup>۱۱</sup> از جمله متداول‌ترین ویژگی‌های استفاده شده برای این مسئله بود.

ده تیم در این رقابت شرکت کردند. بهترین مدل معرفی شده برای تشخیص موضع در زبان اسپانیایی روش SVM با ترکیبی از چندین ویژگی مطرح شده در قسمت بالا به عنوان ورودی مدل است. این در حالی است که مدل Logistic Regression برای زبان کاتالان به نتیجه بهتری دست یافته است. روش‌های یادگیری عمیق نتایج خوبی در این رقابت کسب نکرده اند.

#### ۴-۱-۴ رویداد SadriStance@EVALITA-2020

رویداد SadriStance@EVALITA-2020 اولین رقابت برگزار شده در تشخیص موضع به زبان ایتالیایی می‌باشد. مجموعه داده جمع‌آوری شده درباره موضوع جنبش **Sadrines** در ایتالیا (سال ۲۰۱۹) می‌باشد. در مجموع

<sup>۹</sup>Term Frequency-Inverse Document Frequency

<sup>۱۰</sup>Sentiment Lexicons

<sup>۱۱</sup>word embedding



۳۲۴۲ توییت از توییت‌ر جمع‌آوری شده است. از این میان ۱۷۷۰ توییت برجسب مخالف، ۷۸۴ توییت برجسب موافق و ۶۸۷ توییت باقی مانده برجسب بدون نظر دارند. این رقابت دو زیر مسئله را برای شرکت کنندگان تعریف می‌کند [۱۵].

#### ۱. تشخیص موضع بر اساس متن توییت<sup>۱۲</sup>

در این زیر مسئله شرکت کنندگان برای تشخیص موضع فقط متن توییت را در اختیار دارند.

#### ۲. تشخیص موضع بر اساس ویژگی‌های محتوایی<sup>۱۳</sup>

در این حالت علاوه بر متن توییت برخی دیگر از ویژگی‌ها از جمله تعداد بازاشتراک گذاری<sup>۱۴</sup>، تعداد افرادی که در روز اول انتشار توییت آن را پسندیده‌اند نیز مورد استفاده قرار می‌گیرد. همچنین تعداد دنبال کنندگان<sup>۱۵</sup>، موقعیت مکانی فرد منتشر کننده توییت نیز در دسترس است. سایر اطلاعاتی که از شبکه اجتماعی توییت‌ر از یک کاربر در دسترس می‌باشد از جمله ویژگی‌هایی است که در آموزش مدل می‌توان از آن استفاده کرد. این ویژگی‌ها شامل quotes و نظرات در ارتباط با توییت نیز می‌باشد.

در این رقابت سیزده تیم در زیر مسئله اول (تشخیص موضع بر اساس متن توییت) و دوازده تیم در زیر مسئله دوم (تشخیص موضع بر اساس ویژگی‌های محتوایی) شرکت کردند. شرکت کنندگان روش‌های متنوعی را برای حل تشخیص موضع پیشنهاد دادند. SVM و Logistic Regression از جمله مدل‌های مبتنی بر یادگیری ماشین استفاده شده هستند. همچنین BiLSTM و CNN و مکنیزم توجه از جمله روش‌های یادگیری ماشین استفاده شده است. در نهایت روشی که از BERT [۲۳] استفاده کرده بهترین نتیجه را در مقایسه با سایر شرکت کنندگان کسب کرده است. نایج نشان می‌دهد استفاده از ویژگی‌های کاربران در شبکه‌های اجتماعی باعث بهبود F1-Score می‌شود.

<sup>12</sup>Textual Stance Detection

<sup>13</sup>Contextual Stance Detection

<sup>14</sup>Retweets

<sup>15</sup>Follower

## ۴-۱-۵ رویداد VaxxStance@IberLFE-2021

رویداد VaxxStance@IberLFE-2021<sup>۱۶</sup> برای تشخیص موضع بر روی دو زبان اسپانیایی و باسکی<sup>۱۷</sup> برگزار شد [۲]. داده‌های جمع‌آوری شده برای این مسابقه، درباره موضوع واکسن کرونا در دو زبان جمع‌آوری شده است. در جمع‌آوری مجموعه داده دقت شده تعداد کاربرها و توییت‌ها متناسب باشد. حالتی که تعداد زیادی توییت متعلق به تعداد کمی کاربر باشد مطلوب نیست. همچون رویداد قبلی، علاوه بر متن توییت سایر ویژگی‌های مربوط به کاربران در شبکه اجتماعی نیز جمع‌آوری شده است. برای جمع‌آوری داده از هشتک‌های مرتبط با موضوع واکسن کرونا استفاده کردند. همچنین چالش بزرگ در جمع‌آوری مجموعه داده، پیدا کردن داده با برچسب مخالف بود. در نهایت برای زبان باسکی ۱۳۸۴ توییت که در مجموع مربوط به ۲۱۰ کاربر بود جمع‌آوری شد. داده جمع‌آوری شده به زبان اسپانیایی نیز شامل ۲۶۵۲ توییت مختلف به ۱۶۷۶ کاربر می‌باشد. این رقابت سه زیر مسئله را برای شرکت کنندگان تعریف می‌کند.

## ۱. زیر بخش اول (Close Track)

در این زیر مسئله تنها مجاز به استفاده از داده‌هایی که در رقابت در اختیار شرکت کننده قرار داده شده هستند. به عنوان مثال نمی‌توان برای آموزش بردار بازنمایی بهتر برای کلمات، از مجموعه داده‌های دیگر استفاده کرد. مدل‌های این زیر بخش به دو صورت آموزش می‌بینند. نوع اول مدل‌هایی هستند که برای آموزش آن‌ها تنها از داده‌های متنی استفاده شده است. دسته دوم مدل‌هایی هستند که علاوه بر داده‌های متنی، ویژگی‌های مربوط به شبکه اجتماعی توییت را نیز در اختیار دارند.

## ۲. زیر بخش دوم (Open Track)

در این بخش بر خلاف قسمت قبل، شرکت کنندگان مجاز به استفاده از همه نوع داده هستند. هدف این است به کمک تکنیک‌های افزایش داده Data Augmentation و انتقال دانش بین دو زبان مختلف کیفیت نتایج خروجی بهبود یابد.

## ۳. زیر بخش سوم (Zero-shot Track)

در این بخش نمی‌توان از توییت‌های مربوط به زبان مبدا در آموزش مدل استفاده کرد. هدف حل تشخیص موضع با استفاده از داده‌های شبکه‌های اجتماعی و بدون استفاده از متن می‌باشد.

<sup>۱۶</sup> <https://vaxxstance.github.io/>

<sup>۱۷</sup> باسکی زبان استفاده شده توسط مردم باسک ساکن در شمال اسپانیا و جنوب فرانسه می‌باشد.

در این رقابت بیست تیم ثبت‌نام کردند اما تنها سه تیم نتایج خود را ارسال کردند. همچنین در زیربخش دوم و سوم تنها یک تیم شرکت کرد. نتیجه به دست آمده نشان می‌دهد مدل آموزش داده شده به زبان اسپانیایی از کیفیت بالاتری برخوردار هستند. همچنین اگر در آموزش مدل علاوه بر متن از ویژگی‌های شبکه اجتماعی استفاده شود، بهبود قابل توجه در نتایج به دست خواهد آمد. برای زیربخش اول و دوم نتایج به زبان اسپانیایی بهتر از نتایج زبان باسکی است. اما در زیر بخش سوم نتایج برای دو زبان به یکدیگر نزدیک است.

#### ۴-۱-۶ رویداد ClimateActivismStance@CASE-2024

رویداد ClimateActivismStance<sup>۱۸</sup> برای تشخیص موضع و تشخیص سخنان نفرت انگیز<sup>۱۹</sup> در موضوع تغییرات اقلیمی به زبان انگلیسی برگزار شد [۵۲]. داده‌های استفاده شده از تویتر جمع‌آوری شدند. این رویداد سه زیرمسئله را برای شرکت کنندگان تعریف می‌کند.

##### ۱. زیر مسئله اول (تشخیص سخنان نفرت انگیز)

در این زیر بخش، هدف اصلی تشخیص وجود سخنان نفرت انگیز به ازای متن موجود می‌باشد. مسئله تعریف شده یک رده‌بند دو کلاسه می‌باشد.

##### ۲. زیر مسئله دوم (تشخیص موضوع سخنان نفرت انگیز)

در این زیر بخش هدف تشخیص موضوع سخنان نفرت آمیز به ازای متن موجود می‌باشد. موضوعات شامل فردی<sup>۲۰</sup>، سازمانی<sup>۲۱</sup> و اجتماع<sup>۲۲</sup> می‌باشد.

##### ۳. زیر مسئله سوم (تشخیص موضع)

در این بخش هدف تشخیص موضع برای متن ورودی است. موضع یکی از سه کلاس تایید، رد و بدون نظر می‌باشد.

در مجموع ۲۳ تیم در زیر مسئله اول، ۱۸ تیم در زیر مسئله دوم و ۱۹ تیم در زیر مسئله سوم شرکت کردند. مدل‌های مبتنی بر Transformer نتایج بهتری کسب کردند. بالاترین مقدار F1-Score به دست آمده

<sup>18</sup><https://github.com/therealthapa/case2024-climate>

<sup>19</sup>Hate Speech

<sup>20</sup>Individual

<sup>21</sup>Organization

<sup>22</sup>Community

برای زیر مسئله اول (تشخیص سخنان نفرت انگیز) 91.44% برای زیر مسئله دوم (تشخیص موضوع سخنان نفرت انگیز) 78.58% و برای زیر مسئله سوم (تشخیص موضع) 74.83% می‌باشد. نتایج به دست آمده از این رویداد برای زیر مسئله سوم در جدول ۴-۱ قابل مشاهده است (تیم IUST نیز توانست مقام سوم را در بخش تشخیص موضع کسب کند). تیم اول مسابقات مدل bERTweet با طول ورودی ۹۶ توکن را معرفی کرده است. همچنین تیم دوم مسابقات با استفاده از مهندسی پرامپت و مدل Llama توانست نتایج قابل قبولی کسب کند.

Rank	Team Name	Codalab Username	Accuracy	Precision	Recall	F1-score
1	ARC-NLP (Kaya et al., 2024)	kagankaya1	<b>74.90</b>	<b>78.48</b>	<b>72.26</b>	<b>74.83</b>
2	HAMiSoN-Generative (Fraile-Hernandez and Peñas, 2024)	JesusFraile	74.78	78.27	72.23	74.79
3	IUST (Mahmoudi and Eetemadi, 2024)	gh_mhdi	73.11	78.63	71.45	74.47
4	HAMiSoN-MTL (Rodriguez-Garcia and Centeno, 2024)	Raquel	74.33	77.02	72.42	74.02
5	AAS-T-NLP (El-Sayed and Nasr, 2024b)	AhmedElSayed	74.39	79.31	70.78	73.98
6	MasonPerplexity (Gangul et al., 2024)	Sadiya_Puspo	73.69	77.80	70.90	73.73
7	-	kojiro000	73.43	77.44	70.89	73.58
8	-	refaat1731	72.22	77.49	70.06	73.15
9	HAMiSoN-baselines (Montesinos and Rodrigo, 2024)	juliorremo	74.01	78.17	70.36	73.13
10	-	Nikhil_7280	71.90	76.62	68.13	70.81
11	-	swatirajwal	67.86	70.83	70.05	70.26
12	Bryndza (Suppa et al., 2024)	mareksuppa	71.19	68.72	71.23	69.33
13	NLPDame (Christodoulou, 2024)	christiechris	66.52	71.16	67.94	69.30
14	byteSizedLLM	mdp0999	65.24	72.55	66.85	69.10
15	CUET_Binary_Hackers (Farsi et al., 2024)	SalmanFarsi	66.13	69.08	66.91	67.94
16	Z-AGI Labs (Narayan and Biswal, 2024)	mrutyunjay_research	69.08	79.26	62.94	63.72
17	Team +1	pakapro	32.71	32.66	31.51	28.98
18	-	ankitha11	0.38	1.32	0.16	0.29
19	pokemons	md_kashif_20	0.00	0.00	0.00	0.00

جدول ۴-۱: نتایج رویداد ClimateActivismStance برای زیر مسئله سوم (تشخیص موضع) [۵۲]

## ۴-۲ مجموعه داده‌های مسئله تشخیص موضع

با وجود اینکه تشخیص موضع، یک مسئله نوظهور و جدید است، اخیراً تلاش‌های قابل توجهی برای جمع‌آوری مجموعه داده در این حوزه انجام شده است. اغلب مجموعه داده‌ها به صورت عمومی در دسترس تمام افراد قرار دارد. مجموعه داده‌های جمع‌آوری شده معمولاً از متن‌های مختلف از جمله توییت‌ها، پست‌ها در سایت‌های آنلاین، مقالات خبری و یا نظرات درباره خبرها جمع‌آوری شده‌اند.

مجموعه داده Multi-Target [۴۸] شامل توییت‌هایی در مورد چهار تن از نامزدهای انتخابات ریاست جمهوری آمریکا است. این افراد شامل دونالد ترامپ، هیلاری کلینتون، تد کروزر و برنی سندرز می‌باشند.

هر موضوع شامل دو تن از کاندیدا ریاست جمهوری می‌باشد (جدول ۲-۴). بنابراین برای هر موضوع دو برجسب موجود است (به ازای هر کاندید موضع توییت مشخص می‌شود). مجموعه داده شامل ۴۴۵۵ توییت که به صورت دستی برجسب زده شده است. این مجموعه داده، اولین مجموعه داده معرفی شده به صورت چند موضوعی (Target-Multi) می‌باشد. مجموعه داده WT-WT<sup>۲۳</sup> [۱۶] جمع‌آوری شده به زبان انگلیسی شامل

جدول ۲-۴: توزیع نمونه‌ها در مجموعه داده Multi-Target [۴۸]

Target Pair	# total	# train	# dev	# test
Clinton-Sanders	1366	957	137	272
Clinton-Trump	1722	1240	177	355
Cruz-Trump	1317	922	132	263
Total	4455	3119	446	890

۵۱۲۸۴ توییت می‌باشد. مجموعه داده جهت تشخیص موضع شایعه در بازارهای مالی جمع‌آوری شده است. توییت‌های جمع‌آوری شده درباره بحث‌های صورت گرفته در خصوص واگذاری یک شرکت به شرکت دیگر و یا تجمیع دو شرکت است.

مجموعه داده X-Stance [۵۶] شامل ۶۷۰۰۰ داده در ارتباط با کاندیدا ریاست جمهوری سوییس به چهار زبان انگلیسی، آلمانی، فرانسوی و ایتالیایی می‌باشد. مجموعه داده P-Stance [۲۹] شامل ۲۱۵۷۴ توییت به زبان انگلیسی درباره سه کاندید ریاست جمهوری آمریکا شامل دونالد ترامپ، جو بایدن و برنی سندرز می‌باشد. مجموعه داده‌های قبلی جمع‌آوری شده از تویتر، اغلب شامل داده‌های بسیار کمی بودند. همین موضوع انگیزه‌ای برای جمع‌آوری مجموعه داده

P-Stance شد. مجموعه داده VAST [۵] برای تشخیص موضع بدون داده آموزشی<sup>۲۴</sup> یا با داده آموزشی کم<sup>۲۵</sup> از نظرات موجود در تیویورک تایم جمع‌آوری شده است. این مجموعه داده شامل موضوعات متنوعی است. از هر موضوع تعداد داده کمی در مجموعه داده موجود است.

تا به اینجا معرفی مختصر بر تعدادی از مجموعه داده‌های تشخیص موضع داشتیم. در ادامه به معرفی دقیق‌تر دو مجموعه داده SemEval-2016 و ClimaConvo در این پژوهش می‌پردازیم.

<sup>۲۳</sup> Will-They-Won't-They

<sup>۲۴</sup> Zero-Shot

<sup>۲۵</sup> Few-Shot

## ۴-۲-۱ مجموعه داده SemEval-2016

از معروف ترین مجموعه دادگان به زبان انگلیسی برای مسئله تشخیص موضع که در پژوهش های زیادی مورد استفاده قرار گرفته، مجموعه داده SemEval است. این مجموعه داده برای اولین بار در رویداد SemEval2016 برای دو زیر مسئله تشخیص موضع با نظارت و تشخیص موضع با نظارت ضعیف معرفی شد. این مجموعه داده از شش موضوع شامل انکار وجود خدا<sup>۲۶</sup>، نگرانی از تغییرات آب و هوایی<sup>۲۷</sup>، جنبش فمینیسم<sup>۲۸</sup>، هیلاری کلینتون<sup>۲۹</sup>، قانونی شدن سقط جنین<sup>۳۰</sup> و دونالد ترامپ<sup>۳۱</sup> تشکیل شده است.

مجموعه دادگان SemEval اولین مجموعه داده جمع‌آوری شده از پست‌های تویتر برای موضوع خاص در مسئله تشخیص موضع است. این مجموعه داده علاوه بر برچسب تشخیص موضع، شامل برچسب تحلیل احساسات نیز می‌باشد. مجموعه داده توسط افراد و به صورت دستی، طبق پروتکل خاص برچسب گذاری شده است. چند نمونه از داده‌های موجود در مجموعه داده در جدول ۴-۳ نشان داده شده است.

جدول ۴-۳: چند نمونه از داده‌های مجموعه داده SemEval [۴۰]

متن توییت	موضوع	تشخیص موضع	تحلیل احساسات
I am sad that Hillary lost this presidential race.	Hillary Clinton	موافق	منفی
Life is sacred on all levels. Abortion does not compute with my philosophy. (Red on # OITNB)	Legalization of abortion	مخالف	خنثی
The biggest terror threat in the World is climate change # drought # floods	Climate change is a real concern	موافق	منفی

برای جمع‌آوری مجموعه داده هشتک‌های خاصی در تویتر سرچ شدند. سپس تنها توییت‌هایی که در انتها آن‌ها هشتک وجود داشت نگه داشته و مابقی توییت‌ها حذف شدند. همچنین گاهی از روی هشتک موجود، می‌توان موضع را تشخیص داد. بنابراین هشتک‌های توییت‌ها نیز حذف شدند. برای جمع‌آوری داده‌ها قوانین خاصی در نظر گرفته شد. به عنوان مثال اینکه (۱) توییت‌ها توسط مردم آمریکا قابل فهم باشد. (۲) برای

<sup>26</sup> Atheism<sup>27</sup> Climate Change is Concern<sup>28</sup> Feminist Movement<sup>29</sup> Hillary Clinton<sup>30</sup> Legalization of Abortion<sup>31</sup> Donald Trump

هر موضوع خاص از هر سه کلاس موافق، مخالف و بدون نظر داده به اندازه کافی وجود داشته باشد. (۳) مجموعه داده بهتر است شامل توییت‌هایی باشد که بدون اشاره مستقیم به هدف، موضع خود را نسبت به هدف بیان کند. (۴) مجموعه داده باید شامل داده‌هایی باشد که هدف مورد نظر توییت با موضوع انتخابی توسط ما متفاوت باشد. در نهایت مجموعه داده ای شامل ۴۸۷۰ داده جمع آوری شد. توزیع نمونه ها بر اساس موضوعات در جدول ۴-۴ قابل مشاهده است.

جدول ۴-۴: توزیع نمونه‌ها در مجموعه داده SemEval [۴۰]

Target	# total	# train	% of instances in Train			# test	% of instances in Test		
			favor	against	neither		favor	against	neither
Data for Task A									
Atheism	733	513	17.9	59.3	22.8	220	14.5	72.7	12.7
Climate Change is Concern	564	395	53.7	3.8	42.5	169	72.8	6.5	20.7
Feminist Movement	949	664	31.6	49.4	19.0	285	20.4	64.2	15.4
Hillary Clinton	984	689	17.1	57.0	25.8	295	15.3	58.3	26.4
Legalization of Abortion	933	653	18.5	54.4	27.1	280	16.4	67.5	16.1
All	4163	2914	25.8	47.9	26.3	1249	24.3	57.3	18.4
Data for Task B									
Donald Trump	707	0	-	-	-	707	20.93	42.29	36.78

## ۲-۲-۴ مجموعه داده ClimaConvo

یکی از جدیدترین مجموعه داده‌های معرفی شده برای تشخیص موضع ClimaConvo [۴۵] می‌باشد. این مجموعه داده در رویداد ClimateActivismStance [۵۲] معرفی و استفاده شده است. این مجموعه داده شامل حدود ۱۵ هزار توییت در موضوع تغییرات اقلیمی می‌باشد. شش مسئله متفاوت پردازش زبان طبیعی بر روی این مجموعه داده تعریف شده است. داده‌های این مجموعه داده، مربوط به توییت‌های سال ۲۰۲۲ می‌باشد. برای جمع‌آوری مجموعه داده، توییت‌های شامل هشتک‌های مربوط به تغییرات اقلیمی جمع‌آوری شدند. به عنوان مثال #climatecrisis, #climatechange, #ClimateEmergency, #ClimateTalk, #globalwarming نمونه‌ای از هشتک‌های استفاده شده برای جمع‌آوری داده می‌باشد.

جدول ۴-۵: چند نمونه از داده‌های مجموعه داده ClimaConvo [۴۵]

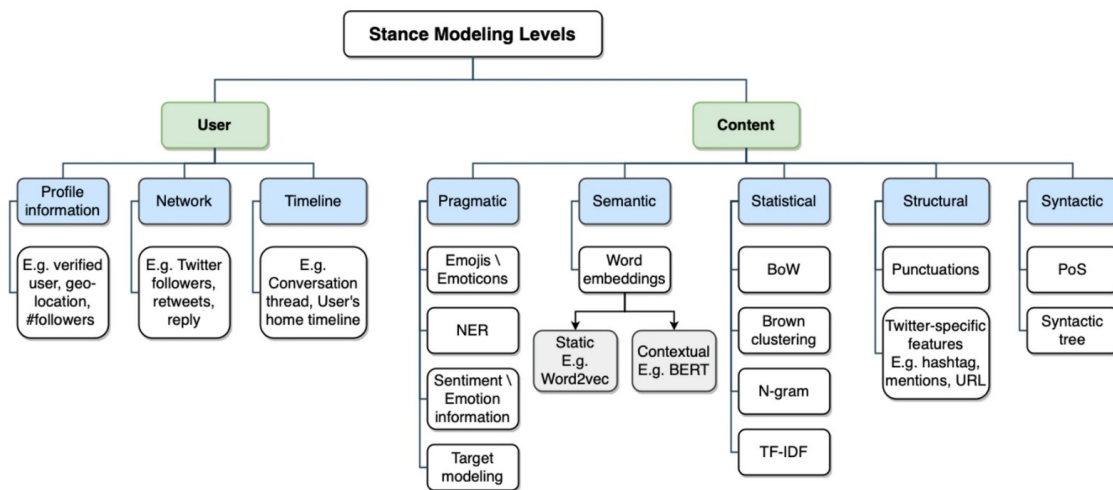
Task	Class	Examples
Task A: Relevance	Relevant	most important youth movement worldwide is the movement against global warming. we want the states to commit. #climatechange #globalwarming
	Non-relevant	Great question! Maybe fear doesn't grow while we sleep. For dreams to #evolve fear needs take a back seat #leonardodavinci #Artist #truth #quote #Jobs #fridaysforfuture
Task B: Stance Detection	Support	I am joining the Global Climate Strike as we demand policymakers and world leaders to prioritize #PeopleNotProfit! #FridaysForFuture
	Neutral	Heat wave in India and Pakistan is so frightening because the period above average is significantly long and seems like the next months will be a real challenge. #climatecrisis
	Oppose	why the wokeness as well as numerous astroturfing like #Unteilbar #FridaysForFuture #LetzteGeneration, which (are supposed to) divert society's attention to sideshows, are financed by the families of the huge fortunes.
Task C: Hate Speech Detection	Hate Speech	[username] Liars and haters spread these #fake-pictures. I would want everyone to know that you are a liar and should be shamed publicly. #climatechange #scammers
	No Hate Speech	We can't leave it all to [username] We all have a responsibility to move away from our fossil fuel addiction and make the right political choices. #climatestrike #canpoli
Task D: Direction of Hate Speech	Directed	Two days ago, [username] was found with [username]. No wonder they both are scamming us together in the name of climate change. #ClimateChange #FridaysForFuture
	Undirected	[username] It is all of us who are to be blamed for this! God will make us burn in hell! I hate everyone! #LossAndDamage #globalwarming
Task E: Targets of Hate Speech	Individual	Greta is brainwashing people on a problem that is non-existent. Brainwashing innocent teenagers isn't cool. #FridaysForFuture #Greta
	Organization	JICA and Sumitomo are trying to build a coal-fired power plant in Bangladesh. Let's kick them out of the nation! #jicanotwelcome #co2emission #ClimateJustice
	Community	climate change is the only problem of white people with certain hairstyles! #not-myproblem #ClimateChange #myth
Task F: Humor Detection	Humor	Trying to explain climate change to my cat: 'You see, Fluffy, the planet is getting warmer because humans are driving around in big metal boxes emitting invisible stuff called greenhouse gases.' #ClimateChange #greenhouse #ClimateConfusion
	No Humor	As global temperatures continue to rise due to human activity, urgent action is needed to mitigate the impacts of climate change on our planet. #ClimateChange #ClimateAction

### ۴-۳ پژوهش‌های انجام شده در مسئله تشخیص موضع

همان‌طور که در بخش ۲-۱ اشاره شد، تشخیص موضع در سطوح متفاوتی تعریف می‌شود. شکل ۴-۲ شامل مروری جامع بر روش‌های ارائه شده در هر سطح می‌باشد. این شکل به صورت طبقه‌بندی شده ویژگی‌های مورد استفاده برای تشخیص موضع را نشان می‌دهد. در ادامه مروری بر روش‌های رده‌بندی در سطح محتوا<sup>۲۲</sup> خواهیم داشت. رویکردهای ارائه شده را به دو دسته‌بندی کلی می‌توان تقسیم کرد. در ادامه پژوهش‌های هر بخش به طور مختصر معرفی خواهد شد.

<sup>32</sup>Content





شکل ۴-۲: مروری بر ویژگی‌های مورد استفاده در تشخیص موضع [۶]

#### ۴-۳-۱ رویکردهای یادگیری ماشین مبتنی بر ویژگی<sup>۳۳</sup>

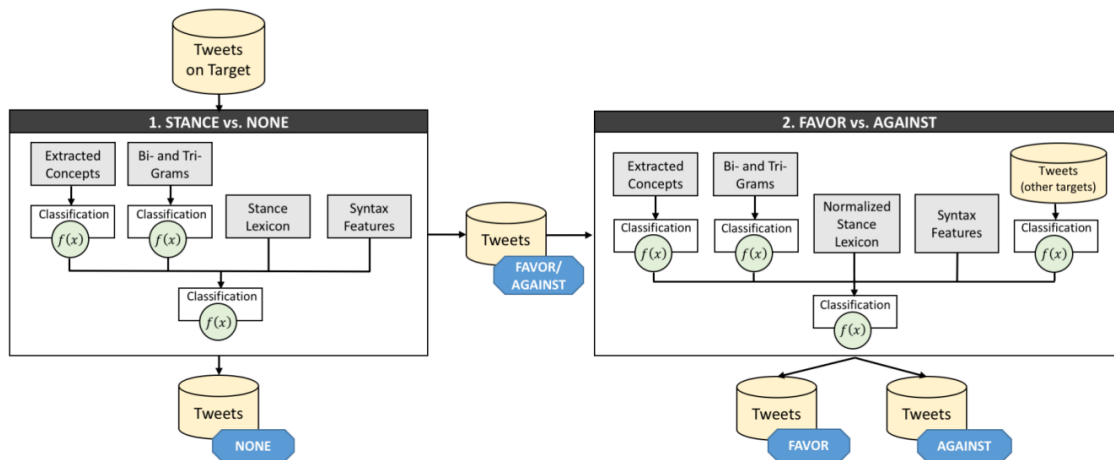
روش SVM متداول‌ترین رویکرد یادگیری ماشین مبتنی بر ویژگی برای تشخیص موضع می‌باشد [۱، ۲۷، ۴۰، ۴۹، ۵۴]. این مدل در بیش از ۴۰ پژوهش مورد استفاده قرار گرفته است. همچنین Logistic Regression [۲۶، ۳۶]، naïve Bayes و درخت تصمیم جز رده‌بندهای متداول استفاده شده هستند. از دیگر الگوریتم‌های استفاده شده می‌توان به KNN و k-means clustering [۴۷] اشاره کرد.

مارتین و همکاران [۵۵] ترکیبی از چندین روش یادگیری را برای حل مسئله تشخیص موضع پیشنهاد می‌کنند. در این پژوهش از چهار مدل شامل SVM، Random Forest، Gradient Boosting و Logistic Regression استفاده می‌شود. این الگوریتم برای ترکیب نتایج به دست آمده از چهار مدل نام برده شده، از الگوریتم ژنتیک استفاده می‌کند. ویژگی‌های واژگان به عنوان ورودی مدل برای رده بندی انتخاب می‌شوند. این ویژگی‌ها شامل بازنمایی لغات، n-gram در سطح کلمه و کاراکتر می‌باشد. همچنین تعدادی ویژگی مخصوص حل مسئله تشخیص موضع نیز تعریف کردند. این ویژگی‌ها شامل تعداد هشتگ‌ها، تعداد برخی کلمات خاص و یا غلط‌های املائی می‌باشد. این ویژگی‌ها به صورت دستی تعریف می‌شوند.

مشیل و همکاران [۶۲]، روش رده بندی پشت سر هم<sup>۳۴</sup> را برای حل تشخیص موضع معرفی کردند. در واقع رده بندی سه کلاسه در دو مرحله، با حل دو رده بندی دو کلاسه انجام می‌شود. مرحله اول مشخص می‌کند

<sup>۳۳</sup>Feature-Based Machine Learning Approaches

<sup>۳۴</sup>Stacked Classifications



شکل ۴-۳: معماری روش مبتنی بر Classifier Stacked [۶۲]

آیا موضعی وجود دارد یا خیر. خروجی این مرحله دو برچسب بدون موضع (None) یا با موضع مشخص (Against/Favor) است. سپس در صورت نیاز، در مرحله دوم برچسب نهایی از بین موافق یا مخالف انتخاب می‌شود. رده بند استفاده شده در این روش، مدل SVM می‌باشد. واژگان موضع<sup>۳۵</sup> از مهم‌ترین ویژگی‌هایی است که برای رده بندی به عنوان ورودی به مدل داده می‌شود.

#### ۴-۳-۲ رویکردهای یادگیری عمیق<sup>۳۸</sup>

شبکه‌های عمیق در تعداد قابل توجهی از مطالعات در تشخیص موضع استفاده شده‌اند. RNN ها و انواع توسعه یافته آن (GRU, LSTM) و شبکه‌های عصبی پیچشی CNN ها از جمله شبکه‌های عمیق استفاده شده هستند. از میان روش‌های مطرح شده، مدل LSTM متداول‌ترین رویکرد استفاده شده برای تشخیص موضع می‌باشد. در ادامه به بررسی و معرفی برخی رویکردهای معروف پرداخته می‌شود. برخی از ویژگی‌های متداول مورد استفاده در روش‌های یادگیری عمیق مرتبط، بازنمایی کلمات Word2vec و GloVe می‌باشد.

در بسیاری از مطالعات اخیر از یک مکانیزم توجه استفاده شده تا عملکرد نهایی را بهبود بخشد [۶۹، ۶۰].

<sup>۳۵</sup> Stance-lexicon Features

<sup>۳۶</sup> embeddings contextualized  
<sup>۳۷</sup> embeddings static

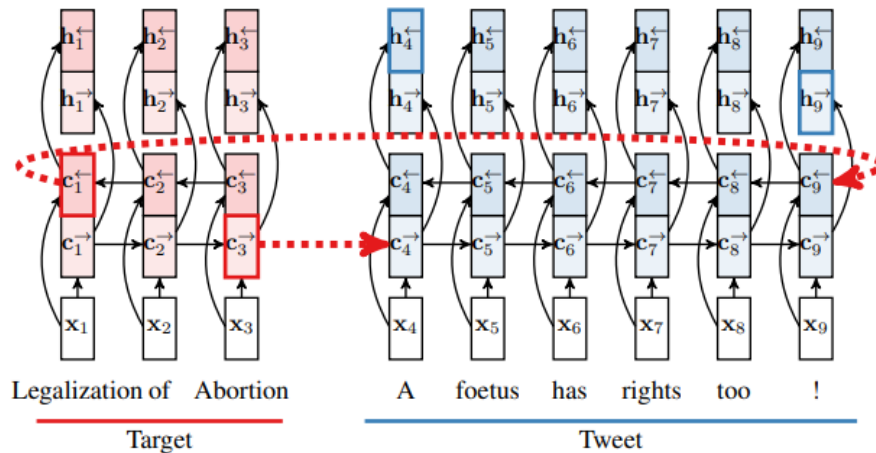
<sup>۳۸</sup> Deep Learning Approaches

جدول ۴-۶: چند نمونه از پژوهش‌های مسئله تشخیص موضع

پژوهش	ویژگی	مدل	سال	مجموعه داده
ARC-NLP [۵۲]	CE <sup>36</sup>	BERTweet	۲۰۲۳	ClimaConvo
HAMiSoN-G [۵۲]	prompt	Llama	۲۰۲۳	ClimaConvo
Baseline [۴۵]	CE	ClimateBert	۲۰۲۳	ClimaConvo
PMINFM [۱۲]	CE, N-gram	Ensemble model (RoBERTa+BiLSTM+attention)	۲۰۲۱	SE16-T6
PE-HCN [۶۸]	CE, Topic modeling	RoBERTa, Hierarchical capsule network	۲۰۲۰	SE16-T6
HAN [۵۰]	N-gram, POS, Structural Sentiment lexicons	LSTM+attention	۲۰۱۸	SE16-T6
[۱۱]	SE <sup>37</sup> , Topic modeling, Sentiment labeling	Multi-Task, Attention	۲۰۱۹	SE16-T6
MITRE [۶۵]	SE, Hashtag prediction	LSTM	۲۰۱۶	SE16-T6
[۴۹]	SE, Hashtag prediction N-gram, Sentiment lexicons	SVM	۲۰۱۶	SE16-T6
[۱۹]	N-gram, Sentiment lexicons Topic modeling	Maximum entropy	۲۰۱۶	SE16-T6
pkudblab [۶۱]	SE	CNN, voting scheme	۲۰۱۶	SE16-T6
GPT3.5 [۶۶]	prompt	GPT3.5	۲۰۲۲	SE16-T6
TPDG [۳۱]	CE, Syntactical dependency, Pragmatic dependency graph	BiLSTM, GCN, attention	۲۰۲۱	SE16-T6
SEKT [۶۷]	SE, Semantic lexicons, Knowledge graph	GCN, BiLSTM+knowledge -aware memory unit	۲۰۲۰	SE16-T6
CrossNet [۶۳]	SE, CE, Target modeling	Attention+MLP	۲۰۱۸	SE16-T6

مدل BiCond<sup>۳۹</sup> توسط ایزابل و همکاران [۷] برای تشخیص موضع با نظارت ضعیف معرفی شده است. این روش از دو شبکه BiLSTM برای حل تشخیص موضع استفاده می‌کند. این دو شبکه به نحوی به یکدیگر وابسته هستند. اولین BiLSTM وظیفه تولید بازنمایی از موضوع را به عهده دارد. دومین BiLSTM بازنمایی از متن توییت تولید می‌کند. نکته مهم در معماری BiCond این است که آخرین مقادیر وزن BiLSTM که موضوع را به عنوان ورودی گرفته، مقادیر اولیه BiLSTM ای است که در ادامه متن را به عنوان ورودی می‌گیرد. این وابستگی در جهت برعکس متن اصلی نیز برقرار است. به همین دلیل است مدل شرطی نام گرفته است. این وابستگی بین موضوع و متن توییت در معماری به وضوح در شکل ۴-۴ (با خط چین قرمز) قابل مشاهده است. این معماری با توجه به وابستگی که بین موضوع و متن توییت ایجاد می‌کند، عملکرد بهتری

<sup>39</sup>Bidirectional Conditional LSTM



شکل ۴-۴: معماری مدل BiCond [۷]

در تشخیص موضع با نظارت ضعیف از خود نشان داده است. همچنین نتایج خوبی در تشخیص موضع برای موضوعاتی که در زمان آموزش ندیده نیز به دست آورده است.

مدل PNEM [۴۶] یک روش ترکیبی برای تشخیص موضع است. در این مدل از LSTM، CNN و مکانیزم توجه استفاده شده است. ابتدا با استفاده از چندین کرنل در شبکه CNN ویژگی‌های سطح بالا از بردار متن و موضوع استخراج می‌شود. در ادامه بردار بازنمایی به دست آمده به عنوان ورودی به DC-BiLSTM<sup>۴۰</sup> و NLSTMs<sup>۴۱</sup> داده می‌شود. سپس دو بردار بازنمایی تولید شده پس از عبور از مکانیزم توجه برای رده‌بندی به شبکه کاملاً متصل وارد می‌شوند. در نهایت با استفاده از بردار بازنمایی نهایی کلاس برنده مشخص می‌شود.

#### ۴-۳-۳ مدل‌های از قبل آموزش دیده

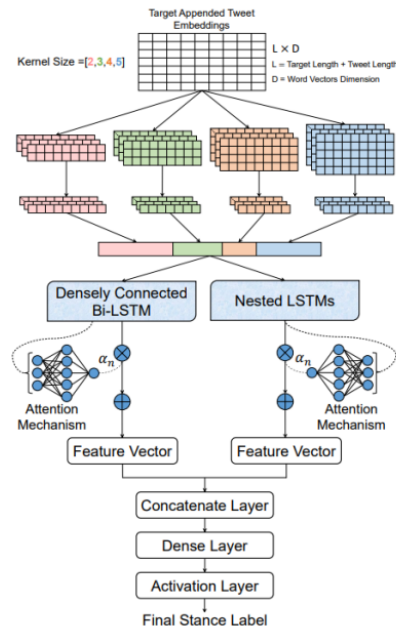
در سال‌های اخیر مدل‌های از پیش آموزش دیده در حوزه پردازش زبان طبیعی تعریف شده است. این مدل‌ها مبتنی بر ترنسفورمر هستند و به دلیل آموزش بر روی حجم بسیار زیادی از داده‌ها، قابلیت بازنمایی از کلمات دارند.

مدل TweetEval [۸] یک ارزیابی مبنایی<sup>۴۲</sup> Benchmark Evaluation برای شبکه اجتماعی تویتر تعریف می‌کند. TweetEval بر روی ۷ مجموعه داده جمع آوری شده از تویتر ارزیابی شده است. بهترین

<sup>۴۰</sup>Densely Connected BiLSTM

<sup>۴۱</sup>Nested LSTMs

<sup>۴۲</sup>text



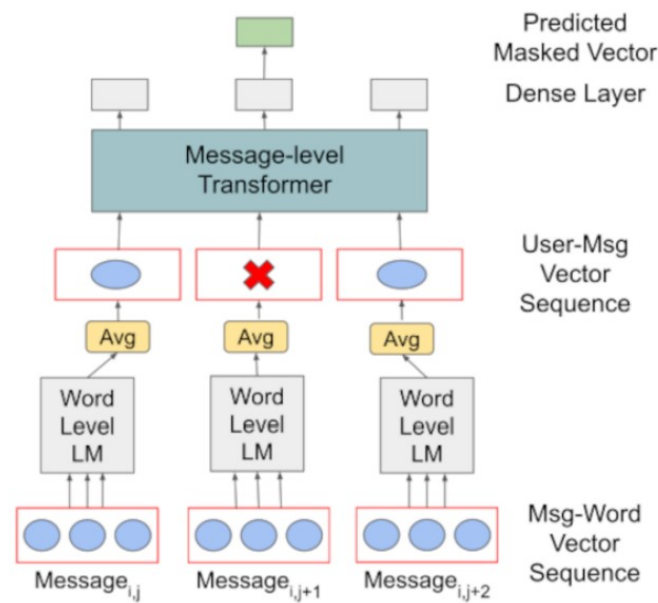
شکل ۴-۵: معماری مدل PNEM [۴۶]

نتایج با رویکرد استفاده از وزن‌های از پیش تعیین شده RoBERTa و آموزش مجدد با داده‌های توئیتر به دست می‌آید.

مدل TimeLMs [۳۵] یک مدل زبانی تعریف شده بر شبکه اجتماعی توئیتر است. مشابه با TweetEval مدل TimeLMs نیز از RoBERTa استفاده می‌کند. نحوه آموزش مدل نیز مشابه با روش انتخاب شده در TweetEval است. نکته جدیدی که در این پژوهش مدل نظر قرار گرفته این است که با توجه به تغییراتی که به صورت مداوم در نحوه نگارش در شبکه‌های اجتماعی رخ می‌دهد، نیاز است مدل به صورت ادامه دار آموزش ببیند. این مدل بازه‌های چهار ماهه را برای این کار انتخاب کرده است. به این صورت که هر چهار ماه به بار مدل پیشین بر روی داده‌هایی که اخیراً از شبکه‌های اجتماعی استخراج شده آموزش ببیند. بدین ترتیب مدل زبانی همواره به روز می‌ماند.

یکی دیگر از مدل‌های زبانی استفاده شده در تشخیص موضع MeLT [۳۷] می‌باشد. روش‌های قبلی مدل زبانی را بر روی یک پیام کاربر و با حذف توکن‌های آن آموزش می‌دادند. گاهی تنها با بررسی یک پیام کاربر نمی‌توان موضع پیام نسبت به موضوع خاصی را تشخیص داد. اما زمانی که سایر متن‌های نوشته شده از کاربر بررسی می‌شود، موضع پیام قبلی که تشخیص آن سخت بود نیز میسر می‌شود. بنابراین این ایده مطرح شد که

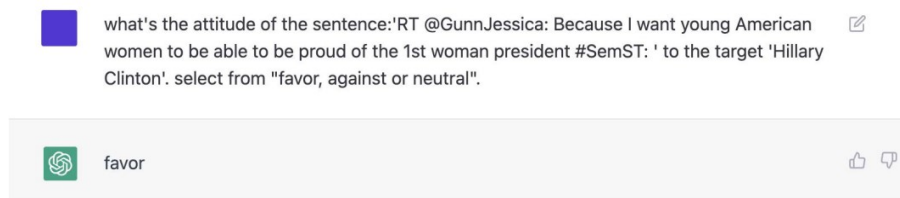
پیام‌های کاربر به صورت پشت سر هم دیده شود. به این منظور چهل پیام اخیر هر کاربر انتخاب می‌شود. و به کمک مدل زبانی از هر پیام یک بردار بازنمایی تولید می‌شود. در ادامه همان روش آموزش مدل‌های زبانی استفاده می‌شود. با این تفاوت که این بار به جای حذف یک توکن متن، بردار بازنمایی یک پیام کاربر حذف می‌شود (شکل ۴-۶) و مدل زبانی آموزش می‌بیند تا این بردار را مجدداً تولید کند.



شکل ۴-۶: معماری مدل MeLT [۳۷]

در پژوهش [۶۶] ژنگ و همکاران با ChatGPT و پیشنهاد پرامپت، مسئله تشخیص موضع را حل کردند.

شکل ۴-۷ پرامپت ارائه شده توسط این مقاله برای تعامل با ChatGPT نشان می‌دهد.



شکل ۴-۷: پرامپت پیشنهادی پژوهش [۶۶]

## ۴-۴ معیار ارزیابی مسئله تشخیص موضع

معیارهای Precision، Recall و F-Score معمولاً در مسائل بازیابی اطلاعات<sup>۴۳</sup> و استخراج اطلاعات<sup>۴۴</sup> مورد استفاده قرار می‌گیرد. معیار F-Score با وزن‌دهی بر دو معیار Precision و Recall محاسبه می‌شود. معیار ارزیابی F-Score همچنین در مسئله تشخیص موضع با سه کلاس نیز مورد قرار می‌گیرد. پرکاربردترین نسخه F-Score به صورت میانگین کلان<sup>۴۵</sup> است که صورت زیر محاسبه می‌شود.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (۱-۴)$$

$$Precision = \frac{TP}{TP + FP} \quad (۲-۴)$$

$$Recall = \frac{TP}{TP + FN} \quad (۳-۴)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (۴-۴)$$

<sup>۴۳</sup>Information Retrieval<sup>۴۴</sup>Information Extraction<sup>۴۵</sup>Macro Average

## فصل ۵

### روش پیشنهادی برای تشخیص موضع با نظارت

تغییرات اقلیمی از جمله مهم‌ترین وقایع محیط زیستی عصر حاضر می‌باشد که بر جوامع انسانی تاثیر مستقیمی می‌گذارد. از سوی دیگر با بررسی نظرات کاربران در شبکه‌های اجتماعی، می‌توان موضع کلی افراد جامعه نسبت به این موضوع را مورد ارزیابی قرار داد. با پیشرفت تکنیک‌های پردازش زبان طبیعی، امکان تشخیص موضع برای داده‌های شبکه اجتماعی فراهم شده است. برای پیشبرد تحقیقات در این حوزه، رویداد ClimateActivism (بخش ۴-۱-۶) مسئله تشخیص موضع بر داده‌های شبکه اجتماعی توییتر با موضوع تغییرات اقلیمی را برگزار کرد [۵۲].

در این پژوهش برای دستیابی به یک مدل هوش مصنوعی برای تشخیص موضع با نظارت، یک روش جستجو جامع به کار گرفته شد. در این جستجو، روش‌های مختلف پیش‌پردازش داده و معماری‌های متفاوت مورد بررسی قرار گرفته است. همچنین تکنیک‌های موثر بر مقابله با داده‌های نامتوازن<sup>۱</sup>، (مانند افزایش داده و انتخاب تابع ضرر متناسب) نیز استفاده شده است. در نهایت معماری پیشنهادی از ترکیب BERTweet و شبکه پیچشی (CNN) و Weighted Cross Entropy به عنوان تابع ضرر استفاده می‌کند. در ادامه این فصل، روش جستجو استفاده شده، فضا جستجو تعریف شده، جزییات معماری پیشنهادی، آزمایش‌های انجام شده و تحلیل نتایج به صورت مفصل بیان می‌شود.

---

<sup>1</sup> Imbalance data

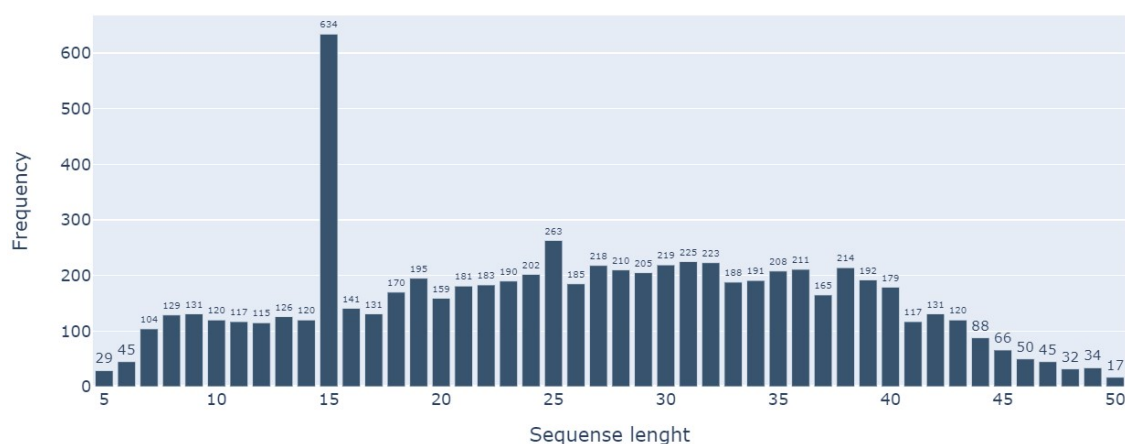


## ۵-۱ دادگان آموزشی

برای آموزش، ارزیابی و مقایسه مدل پیشنهادی از مجموعه داده ClimaConvo [۴۵] استفاده شد. این مجموعه داده شامل شش مسئله مختلف از جمله تشخیص موضع، تشخیص طنز<sup>۲</sup> و تشخیص سخنان نفرت انگیز<sup>۳</sup> می‌باشد. در اینجا تنها از بخش تشخیص موضع دادگان استفاده شده است. توزیع داده‌های هر کلاس در جدول ۵-۱ نشان داده شده است. بخش ۴-۲-۲ شامل توضیحات تکمیلی درباره مجموعه داده می‌باشد. شکل ۵-۱ توزیع تعداد تک واژه‌های مجموعه داده ClimaConvo را نشان می‌دهد.

جدول ۵-۱: توزیع کلاس‌ها در بخش‌های آموزش، توسعه و ارزیابی مجموعه داده ClimaConvo

	%	موافق	بدون نظر	مخالف
آموزش (Train)	۷۰%	۴۳۲۸	۲۲۵۶	۷۰۰
توسعه (Dev)	۱۵%	۸۹۷	۵۱۱	۱۵۳
ارزیابی (Test)	۱۵%	۹۲۱	۵۰۰	۱۴۱
مجموع	۱۰۰%	۶۱۴۶	۳۲۷۶	۹۹۴



شکل ۵-۱: توزیع تعداد تک واژه‌های مجموعه داده ClimaConvo

همچنین در قسمت پایانی این فصل برای ارزیابی عملکرد مدل بر مجموعه داده‌ای که بر روی آن آموزش ندیده، دو بخش "قانونی شدن سقط جنین" و "تغییرات اقلیمی" از مجموعه داده SemEval [۴۰] مورد استفاده قرار گرفت. جدول ۵-۲ توزیع کلاس‌های مجموعه داده به ازای هر موضوع را نشان می‌دهد.

<sup>2</sup>Humor Detection

<sup>3</sup>Hate Speech Detection

جدول ۵-۲: توزیع کلاس‌ها در بخش‌های ارزیابی مجموعه داده SemEval

موضوع	موافق	بدون نظر	مخالف
تغییرات اقلیمی (CC)	۱۱	۳۵	۱۲۳
قانونی شدن سقط جنین (LA)	۱۸۶	۴۵	۴۶

## ۵-۲ روش حل مسئله

رویکرد ارائه شده، از روش جستجو معماری شبکه عصبی الگو می‌گیرد. بدین منظور معماری شبکه به چهار بخش اصلی تقسیم شد. به ازای هر یک از بخش‌ها با توجه به دانش قبلی، فضای جستجو متناسب معرفی شد. در ادامه از جستجو تطبیقی برای یافتن بهترین پارامترهای مدل و ابرپارامترها استفاده شد. روش جستجو استفاده شده در مقایسه با جستجو تصادفی و جستجو شبکه‌ای مزایای زیادی دارد. این روش از پیشینه جستجو برای انتخاب پارامترهای بعدی استفاده می‌کند. همچنین در صورت ناموفق بودن یک جستجو، به صورت زودهنگام جستجو را متوقف می‌کند. بنابراین در زمان مناسب می‌تواند جستجو نسبتاً جامعی انجام دهد. از سوی دیگر داده‌های ورودی مدل نیز اهمیت ویژه‌ای دارند. از این رو هفت سطح از پیش‌پردازش داده معرفی شد تا تاثیر آن بر نتایج نهایی مورد بررسی قرار بگیرد. همچنین با توجه به نامتوازن بودن داده‌های آموزشی از روش‌های افزایش داده برا غلبه به این مشکل استفاده شد. در ادامه توضیحات مربوط به سطوح پیش‌پردازش داده، روش افزایش داده و رویکرد جستجو معماری به صورت مفصل بیان می‌شود.

### ۵-۲-۱ پیش‌پردازش داده‌ها

آماده‌سازی داده‌های مناسب و با کیفیت یکی از مهم‌ترین بخش‌های آموزش یک مدل یادگیری عمیق است. ابتدایی‌ترین گام در مسائل پردازش زبان طبیعی، پیش‌پردازش داده‌ها است که به ارتقاء کیفیت داده‌ها و استخراج معنای بهتر از آن‌ها کمک می‌کند. پیش‌پردازش داده‌ها به معنای تمیز کردن و سازماندهی داده‌های خام به شکل قابل فهم برای آموزش مدل‌های یادگیری عمیق است.

داده‌های متنی شبکه‌های اجتماعی (همانند توییتر) معمولاً شامل نویز و عباراتی هستند که ممکن است مدل را دچار خطا کند. در فرآیند پیش‌پردازش داده، رویکرد اصلی این است که هیچ داده‌ای دور ریخته نشود و حداکثر اطلاعات از داده‌های متنی موجود استخراج شود. اما به دلیل محدودیت در داده‌های آموزشی، وجود نویز در داده‌ها ممکن است آموزش مدل را با مشکل روبرو کند. در ادامه هفت سطح متفاوت برای

پیش‌پردازش داده معرفی می‌شود. روش‌های معرفی شده به صورت سلسله مراتبی هستند. بدین معنا که هر روش نسبت به روش قبلی، اطلاعات بیشتری را حذف می‌کند یا تغییر می‌دهد.

۱. نگهداشت داده‌های اصلی: در این رویکرد فرض کردیم بهترین روش عدم تغییر داده‌های اصلی می‌باشد. در این روش متن داده‌های موجود بدون هیچ تغییری به عنوان ورودی مدل در نظر گرفته می‌شود.

۲. حذف آدرس اینترنتی<sup>۴</sup>: با توجه به اینکه اغلب آدرس‌های اینترنتی به صورت مختصر وجود داشتند (به عنوان مثال <https://t.co/rs1vhBp2ax>) امکان یافتن الگو از آن‌ها ممکن نبود. بنابراین در این رویکرد فقط آدرس اینترنتی حذف شده و مابقی متن را تغییری نکرد.

۳. حذف نام کاربری<sup>۵</sup>: فرض ما این است وجود نام کاربری بدون اینکه اطلاعات دیگری از کاربران داشته باشیم بیشتر موجب گمراهی مدل می‌شود. برای بررسی صحت فرض مطرح شده، این رویکرد را نیز در آزمایش‌های خود قرار دادیم.

۴. حذف همزمان نام کاربری و آدرس اینترنتی: رویکرد دوم و سوم را برای بررسی اثربخشی همزمان آن‌ها با هم اعمال شد.

۵. حذف همزمان نام کاربری و آدرس اینترنتی و جدا کردن هشتگ: در شبکه‌های اجتماعی همچون توییتر برای بیشتر دیده شدن توییت‌ها و همراهی با رویدادی خاص، گاهی هشتگ‌هایی در متن اصلی توییت استفاده می‌شود. استفاده از هشتگ‌ها جستجو آن‌ها را نیز ساده‌تر می‌کند. در این رویکرد برای تمیز کردن داده‌ها هشتگ‌ها به صورت یک عبارت در آوردیم به گونه‌ای که #FridaysForFuture به Fridays For Future تبدیل می‌شود. انتظار می‌رود بعد از جداسازی هشتگ‌ها عبارت قابل فهم‌تری داشته باشیم.

۶. حذف همزمان نام کاربری و آدرس اینترنتی و جدا کردن هشتگ و کوچک کردن حروف: در زبان انگلیسی معمولاً کلمات در ابتدا جمله با حروف بزرگ شروع می‌شوند. از طرفی گاهی کلیه حروف یک کلمه بنا به دلایلی خاص از جمله تاکید، خشم یا عصبانیت به صورت بزرگ نوشته می‌شود. از سوی دیگر دو کلمه Book و book هم معنی هستند اما با توجه به اختلاف در حروف اول دو بازنمایی

<sup>4</sup>URL

<sup>5</sup>Username

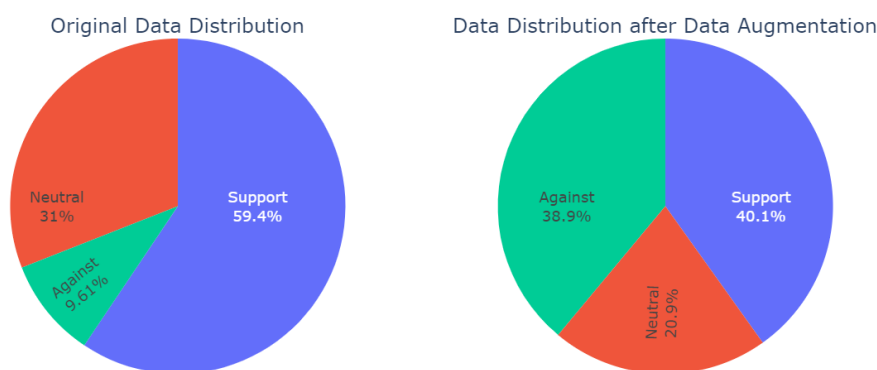
متفاوت در مدل‌های هوش مصنوعی پیدا می‌کنند. یکی از گام‌های پیش‌پردازش داده‌ها کوچک کردن همه حروف<sup>۶</sup> می‌باشد. در این سطح از پیش‌پردازش داده قصد داریم تاثیر این عمل بر نتیجه نهایی مدل را به دست بیاوریم.

۷. تمیز کردن کامل داده‌ها: این سطح شامل کامل‌ترین پیش‌پردازش می‌باشد و علاوه بر همه موارد رویکرد ششم، کلمه‌های توقف<sup>۷</sup> و علائم دستوری نیز حذف می‌شوند.

در بخش ۵-۳-۳-۱، عملکرد هر یک از روش‌های پیش‌پردازش مورد بررسی قرار می‌گیرد.

## ۵-۲-۲ افزایش داده

یکی از روش‌های استفاده شده برای غلبه بر مشکلات ناشی از مجموعه داده کوچک و نامتوازن، استفاده از تکنیک‌های افزایش داده می‌باشد. در این پژوهش برای افزایش داده، از مدل مبتنی بر T5 که بر روی مجموعه داده ChatGPT paraphrase [۵۸] آموزش دیده [۵۹]، برای افزایش داده استفاده می‌کنیم. این مدل آموزش دیده تا متن ورودی را به گونه‌ای دیگر بازنویسی<sup>۸</sup> کند. برای افزایش داده به کمک این روش، توییت‌هایی کلاس مخالف را به عنوان ورودی مدل دادیم. مدل به ازای هر متن، ۵ بازنویسی از آن را تولید کرده است. توزیع مجموعه داده آموزشی بعد از افزایش داده در شکل ۵-۲ قابل مشاهده است. در بخش ۵-۳-۳-۲، تاثیر افزایش داده بر نتایج به دست آمده مورد بررسی قرار می‌گیرد.



شکل ۵-۲: توزیع کلاس‌ها قبل و بعد از افزایش داده.

<sup>۶</sup>Lower Casing

<sup>۷</sup>Stop Words

<sup>۸</sup>Paraphrase

## ۵-۲-۳ جستجو معماری

در این پژوهش، یک روش جستجو قاعده‌مند<sup>۹</sup> برای یافتن بهترین معماری مدل استفاده شده است. برای تعریف فضا جستجو، ابتدا معماری پیشنهادی به چهار بخش اصلی تقسیم شد. سپس برای هر یک از چهار بخش اصلی فضای جستجو مناسب تعریف شد. برای تعیین مناسب‌ترین پارامترها برای هر بخش، آزمایش‌های متعددی با پیکربندی‌های مختلف انجام داده شد و مقادیر بهینه در فضای جستجوی تعریف شده جستجو شد (جدول ۵-۳). جستجو در فضای تعریف شده با استفاده از کتابخانه Optuna، که از نمونه‌بردار با استفاده از الگوریتم TPE (Tree-structured Parzen Estimator) استفاده می‌کند، انجام شد. پیکربندی مدل بهینه را بر اساس امتیاز Macro-F1 در مجموعه توسعه<sup>۱۰</sup> انتخاب شد. فضا جستجو تعریف شده برای هر بخش در جدول ۵-۳ قابل مشاهده است.

جدول ۵-۳: فضای جستجو معماری پیشنهادی

پارامتر	فضا جستجو
بازنمایی کلمات	[BERT, RoBERTa, BERTweet, XLM-RoBERTa, DEBERTA]
رده‌بند	[FNN, CNN]
N_last_layer	[1, 2, 3, 4, 5]
بهینه‌ساز	[Adam, AdamW, RMSprop, SGD]
تابع ضرر	[Weighted Cross Entropy, Focal]

۱. بازنمایی کلمات<sup>۱۱</sup>: برای این بخش تعدادی از مدل‌های کدگذار<sup>۱۲</sup> که نتایج بهتری در مسائل مشابه کسب کردند را انتخاب کردیم. استفاده از این مدل‌ها در مسائل رده‌بندی<sup>۱۳</sup>، متداول‌تر از مدل‌های کدگشا<sup>۱۴</sup> و کدگذار-کدگشا<sup>۱۵</sup> هستند (جدول ۵-۳).

۲. رده‌بند<sup>۱۶</sup>: در این بخش دو معماری متفاوت را بررسی کردیم.

- معماری شبکه عصبی تماماً متصل (FNN)<sup>۱۷</sup>

<sup>9</sup>Systematic

<sup>10</sup>Development Set

<sup>11</sup>Embedding

<sup>12</sup>Encoder-Only

<sup>13</sup>Classification Task

<sup>14</sup>Decoder-only

<sup>15</sup>Encoder-Decoder

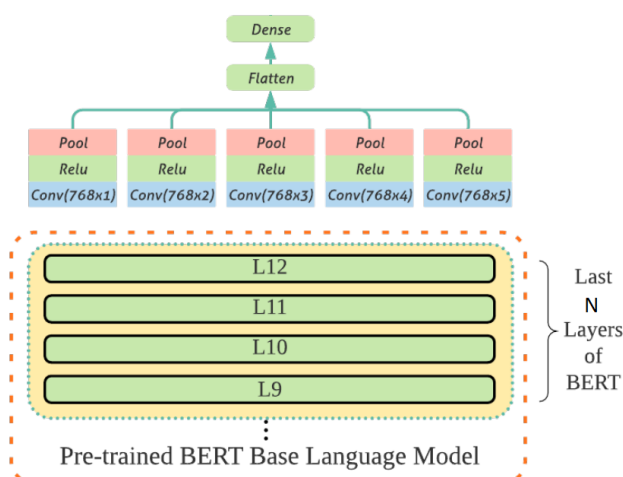
<sup>16</sup>Classifier

<sup>17</sup>Fully Connected Neural Networks

در این بخش ما از یک معماری سه لایه خطی، به همراه تابع فعال سازی ReLU و dropout استفاده می کنیم. در نهایت Softmax را بر روی خروجی اعمال می کنیم.

• معماری شبکه عصبی پیچشی (CNN)<sup>۱۸</sup>

ایده اصلی معماری استفاده شده مشابه کار سفایا و همکاران [۴۴] می باشد. تفاوت اصلی در این است که به جای ۴ لایه آخر، فضا جستجو برای تعداد لایه ها تعریف کرده و با آزمایش مقادیر مختلف، بهترین مقدار را به دست می آوریم. در این معماری Embedding به تعداد ۳۲ فیلتر کانولوشنال موازی با اندازه متفاوت وارد می شود (۱×۷۶۸، ۲×۷۸۶، ۳×۷۸۶، ۴×۷۸۶، ۵×۷۸۶). معماری پیشنهاد شده مشابه شبکه Inception در بینایی ماشین می باشد. بعد از اعمال فیلتر کانولوشنال، تابع فعال سازی ReLU و Global Max-Pooling بر خروجی مرحله قبل اعمال می شود. خروجی تولید شده بعد از Flatten شدن به یکدیگر Concat می شوند. در انتها با اعمال Softmax خروجی نهایی حاصل می شود. شکل [۴۴] معماری رده بند متشکل از شبکه عصبی پیچشی را نشان می دهد.



شکل ۵-۳: رده بند با شبکه عصبی پیچشی [۴۴]

۳. بهینه ساز<sup>۱۹</sup>: فضای جستجو تعریف شده برای بهینه سازها شامل چهار بهینه ساز معروف، که عملکرد

<sup>18</sup>Convolutional Neural Networks

<sup>19</sup>Optimizer

مثبتی از خود نشان دادند، می باشد (جدول ۵-۳).

۴. تابع ضرر <sup>۲۰</sup>: از آنجایی که با داده های نامتوازن و مسئله رده بندی مواجه هستیم دو تابع ضرر متفاوت با توجه به شرایط تعریف کردیم.

- تابع ضرر Focal: تابع ضرر Focal [۳۲] با کاهش وزن داده های کلاس اکثریت (که به آسانی توسط مدل یادگرفته می شوند) در زمان آموزش مدل، تاثیر نامتوازن بودن داده را کم می کند. این تابع ضرر به یادگیری صحیح داده های سخت تر (کلاس اقلیت) تاکید بیشتری دارد تا به این صورت عملکرد نهایی بهبود پیدا کند.

- تابع ضرر Weighted Cross Entropy: این تابع ضرر گونه ای از تابع ضرر استاندارد Cross Entropy می باشد به گونه ای که به کلاس های مختلف، وزن های متفاوتی اختصاص می دهد. وزن های هر کلاس با محاسبه نسبت معکوس تعداد داده های آموزشی هر کلاس به کل داده های آموزشی به دست می آید. به عبارت دیگر اشتباهات کلاس اقلیت نیز اهمیت پیدا می کند.

## ۵-۲-۴ انتخاب ابرپارامترها

انتخاب دقیق ابرپارامترهای استفاده شده در زمان آموزش مدل، از اهمیت ویژه ای برخوردار است. در این قسمت از کتابخانه Optuna و نمونه بردار TPE (Tree-structured Parzen Estimator) استفاده شد. در این شیوه جستجو ابرپارامترها در هر آزمایش، پارامترها در فضا جستجو انتخاب می شوند. این الگوریتم معیاری را برای مقایسه نتیجه آزمایش های انجام شده در نظر می گیرد. در این پژوهش معیار F1-Score به عنوان معیار اصلی تعریف شد. در نهایت ترکیبی از فرایارامترها که باعث افزایش معیار F1-Score شوند به عنوان پیشنهاد نهایی ارائه می شود. همچنین الگوریتم قابلیت خاتمه زودهنگام آزمایش هایی با نتایج نامناسب را دارد. جدول ۵-۴ فضای جستجو ابرپارامترهای مدل را نشان می دهد.

<sup>20</sup>Loss Function

جدول ۴-۵: فضا جستجو ابرپارامترها

فضای جستجو	ابرپارامتر
$[0.1 : 0.5]$	Dropout
$[1e^{-5} : 1e^{-2}]$	نرخ یادگیری
$[4, 8]$	تعداد دوره (Epoch)
$[4, 8]$	اندازه دسته (Batch Size)
$[1, 2, 3, 4, 5]$	پارامتر گاما تابع ضرر Focal

### ۳-۵ آزمایش‌ها و تحلیل نتایج

برای یافتن بهترین معماری و تحلیل اثرگذاری بخش‌های مختلف بر نتیجه اصلی اقداماتی صورت گرفته است. در این بخش به توضیح اقدامات انجام شده پرداخته می‌شود.

#### ۱-۳-۵ گام اول: تحلیل نحوه عملکرد الگوریتم جستجو

برای شناخت بیشتر از عملکرد ابزار optuna دو آزمایش طراحی و اجرا شد. هدف از انجام این آزمایش‌ها بررسی میزان تکرارپذیری و قابل اتکا بودن این ابزار بود.

۱. جستجو در فضای نمونه تعریف شده به ازای هر کدام از روش‌های تمیز کردن داده‌ها ۳ بار و هر کدام به تعداد ۱۰ تریال<sup>۲۱</sup> تکرار شد (یعنی ۱۰ بار پارامترهای مختلف در فضا جستجو انتخاب شد و مدل آموزش داده شد). هدف از این بخش بررسی بهترین پارامترهایی است که بعد از ۱۰ بار جستجو به دست می‌آید. آیا همواره به پارامترهای یکسانی می‌رسد؟

۲. جستجو در فضای نمونه به تعداد تریال‌های مختلف (۱۰، ۲۰، ۴۰) تکرار شد تا تاثیر افزایش تعداد تریال بر عملکرد نهایی مدل را به دست بیاید (با مقایسه F1-Score).

برای جستجو به تعداد ده تریال، نتایج به دست آمده نشان می‌دهد بهترین پارامترهای پیشنهادی هر کدام از جستجو‌ها تا حدودی با یکدیگر متفاوت بودند. اما F1-Score نهایی بهترین مدل تقریباً یکسان بودند. از

<sup>۲۱</sup> منظور از تریال همان trial می‌باشد. آزمایشی که در آن یک بار از فضای جستجو مقادیر پارامترها انتخاب شده و عملکرد مدل بر اساس آن مقادیر مورد بررسی قرار می‌گیرد. در این گزارش برای نزدیک بودن متن نوشته به توضیحات مربوط به optuna از همان اصطلاح تریال استفاده شده است.



طرفی با افزایش تعداد ترایال‌ها به ۲۰، نتیجه نهایی بهبود داشت و پارامترهای دقیق‌تری انتخاب شدند. منتها نتایج ۴۰ ترایال و ۲۰ ترایال اختلاف معناداری با یکدیگر نداشتند. با توجه به نتایج به دست آمده تصمیم گرفته شد آزمایش‌ها را به تعداد ۲۰ ترایال تکرار کرده و بهترین پیکربندی به دست آمده، به عنوان مدل پیشنهادی ارائه شود.

### ۵-۳-۲ گام دوم: به دست آوردن پارامترهای مدل با عملکرد مطلوب

در بخش ۵-۲-۱ هفت سطح مختلف برای پیش‌پردازش داده‌ها معرفی شد. همچنین در بخش ۵-۲-۲ نحوه افزایش داده به تفصیل توضیح داده شد. آزمایش‌ها در دو حالت با افزایش داده و بدون افزایش داده تکرار شدند. بنابراین چهارده پیکربندی مختلف تعریف شد. به ازای هر پیکربندی، ۲۰ بار در ترایال‌های مختلف با استفاده از الگوریتم جستجو Optuna، پارامترهای متفاوتی انتخاب شد و سپس مدل با آن پارامترها آموزش داده شد. همچنین مکانیزمی برای توقف ترایال‌هایی که نتایج آن‌ها در جهت بهبود پیش نمی‌رود نیز وجود دارد. همان‌طور که در بخش‌های قبل اشاره شد، به دلیل بزرگ بودن فضا جستجو، جستجو شبکه‌ای بسیار زمان‌بر و عملاً غیر ممکن خواهد بود. برای اینکه بتوان فضای جستجو نسبتاً وسیع‌تری را مورد بررسی قرار داد از جستجو تطابقی، که پیشینه جستجو را در انتخاب‌های بعدی در نظر می‌گیرد، استفاده شد. بنابراین در این آزمایش‌ها همه ترکیب‌های ممکن بررسی نشده است اما الگوریتم جستجو به صورت هوشمند در هر ترایال بهترین پارامترها را انتخاب می‌کند و در صورت لزوم ادامه آموزش یک ترایال را متوقف می‌کند. چنین فرآیندی سرعت رسیدن به پارامترهای بهینه را افزایش می‌دهد. نتایج به دست آمده از این جستجو در جدول ۵-۵ قابل مشاهده است. در این جدول تنها نتایجی که معیار F1-Score بالاتر از ۶۸ درصد کسب کردند، گزارش شدند (از آن جایی که فقط نتایج بیشتر از ۶۸ درصد در معیار F1-Score گزارش شده است، ممکن است به ازای برخی از حالات پیش‌پردازش داده، رده‌بند FNN وجود نداشته باشد).

بهترین پارامترهای به دست آمده از این جستجو جامع را به عنوان مدل پیشنهادی خود معرفی کردیم. مدل پیشنهادی از BERTweet و رده‌بند CNN و تابع ضرر Weighted Cross Entropy استفاده می‌کند. همچنین برای آموزش مدل، داده‌های آموزشی با تکنیک معرفی شده افزایش داده می‌شود و برای پیش‌پردازش داده‌ها نام کاربری و آدرس اینترنتی را حذف می‌شوند. این روش در رویداد ClimateActivismStance (بخش ۴-۱-۶) توانست بین ۱۹ شرکت‌کننده مسابقات رتبه سوم را کسب کند (جدول ۴-۱). جدول ۵-۶ نشان می‌دهد معماری طراحی شده توانست مدل پایه (Baseline) را به ۱۹،۹۷ درصد در معیار F1-Score بهبود دهد.

جدول ۵-۵: نتایج به دست آمده بر روی مجموعه داده ClimateCanvo Stance بخش ارزیابی. روش‌های پیش‌پردازش داده: (C1: نگهداشت داده‌های اصلی، C2: حذف آدرس اینترنتی، C3: حذف نام کاربری، C4: حذف نام کاربری و آدرس اینترنتی به صورت همزمان، C5: حذف نام کاربری و حذف آدرس اینترنتی به صورت همزمان و جدا کردن هشتک، C6: حذف نام کاربری و آدرس اینترنتی به صورت همزمان و جدا کردن هشتک و کوچک کردن حروف، C7: تمیز کردن کامل داده‌ها).

پیش‌پردازش	افزایش داده	کدگذار	رده‌بند	تابع ضرر	بهینه‌ساز	F1-Score	Recall	Precision	Accuracy
C1	-	RoBERTa	CNN(N=1)	WCE	SGD	71.74	69.94	74.83	68.82
C2	-	XLM-RoBERTa	CNN(N=3)	WCE	AdamW	69.80	69.38	74.16	64.91
C2	-	BERT	CNN(N=2)	WCE	SGD	70.28	68.64	73.82	66.00
C2	+	BERT	CNN(N=3)	WCE	SGD	68.75	66.27	75.26	70.16
C3	-	RoBERTa	FNN	WCE	SGD	71.89	68.89	79.55	73.81
C3	+	XLM-RoBERTa	CNN(N=3)	F(g=4)	SGD	68.59	68.51	75.93	69.84
C4	-	XLM-RoBERTa	CNN(N=2)	WCE	RMSprop	71.82	69.56	74.80	69.52
C4	-	BERT	CNN(N=5)	WCE	SGD	72.82	69.56	74.80	72.85
C4	-	XLM-RoBERTa	CNN(N=3)	F(g=1)	SGD	73.97	70.91	78.17	72.59
C4	-	RoBERTa	FNN	WCE	RMSprop	71.52	68.31	78.17	70.06
C4	-	XLM-RoBERTa	CNN(N=4)	WCE	SGD	72.72	69.38	78.85	73.17
C4	+	BERTweet	CNN(N=5)	WCE	SGD	74.47	70.31	79.31	73.11
C4	+	BERT	CNN(N=4)	WCE	SGD	70.64	67.75	75.63	70.01
C5	-	DEBERTa	FNN	WCE	Adam	71.33	68.78	75.43	67.73
C5	-	XLM-RoBERTa	CNN(N=2)	WCE	SGD	72.70	69.63	77.18	72.15
C5	-	BERT	FNN	F(g=1)	SGD	72.01	68.62	77.44	71.75
C5	+	DEBERTa	FNN	WCE	AdamW	71.20	68.51	77.68	72.72
C5	+	BERT	CNN(N=3)	WCE	SGD	70.85	70.01	73.41	67.22
C6	-	BERT	FNN	F(g=2)	SGD	71.83	68.38	74.48	72.21
C6	-	BERT	FNN	WCE	RMSProp	70.13	67.18	74.85	69.65
C6	+	XLM-RoBERTa	CNN(N=4)	WCE	SGD	72.70	69.43	78.56	72.85
C7	-	BERT	CNN(N=5)	F(g=1)	SGD	71.68	68.77	76.53	71.76
C7	+	BERTweet	CNN(N=2)	F(g=4)	AdamW	69.36	66.56	74.09	69.06

جدول ۵-۶: مقایسه بهترین معماری طراحی شده با مدل پایه. نتایج \* از [۴۵] گزارش شده است.

F1	ACC	Model
74.47	73.11	<sup>۲۲</sup> BERTTweets
73.97	72.59	<sup>۲۳</sup> XLM-RoBERTa
54.5	65.1	*ClimateBERT (Baseline)

شکل ۴-۵: تغییرات معیارهای ارزیابی را به ازای هر دوره از آموزش بر مجموعه داده آموزشی، توسعه و ارزیابی نشان می‌دهد.

<sup>۲۲</sup> رده‌بند CNN با ۵ لایه آخر BERTTweets همراه با افزایش داده، تابع ضرر Weight Cross Entropy و بهینه‌ساز SGD و مکانیزم پیش‌پردازش حذف همزمان نام کاربری و آدرس اینترنتی.

<sup>۲۳</sup> رده‌بند CNN با ۵ لایه آخر XLM-RoBERTa، تابع ضرر Focal و بهینه‌ساز SGD و مکانیزم پیش‌پردازش حذف همزمان نام کاربری و آدرس اینترنتی.



شکل ۵-۴: عملکرد مدل آموزش دیده در هر epoch آموزش

در گام بعدی اثربخشی بخش‌های مهم روش پیشنهادی مورد بررسی دقیق‌تر قرار می‌گیرد. به صورت دقیق‌تر تاثیر نوع پیش‌پردازش داده، افزایش دادگان آموزشی، رده‌بند و تابع ضرر بر نتایج نهایی بررسی می‌شود.

### ۳-۳-۵ گام سوم: ارزیابی تکمیلی مدل پیشنهادی

در گام دوم با جستجو کامل بر روی همه پارامترهای مدل و ابرپارامترها، بهترین ترکیب پارامترها به دست آمد. پارامترهای نهایی مدل پیشنهادی در جدول ۷-۵ قابل مشاهده است. در این بخش با طراحی آزمایش‌هایی سعی شده تا تاثیر هر یک از موارد در نتیجه نهایی به صورت جداگانه بررسی شود.

جدول ۷-۵: پارامترهای بهترین مدل پیشنهادی.

پارامتر	مقدار
تعداد دوره (Epoch)	8
اندازه دسته (Batch Size)	4
نرخ یادگیری	0.007903
طول عبارت ورودی	128
Dropout	0.5
Learning schedule	Linear Schedule With Warmup
کدگذار	BERTweet
رده‌بند	CNN
N_last_layer	5
بهینه‌ساز	SGD
تابع فعال‌سازی	Weighted Cross Entropy

### ۱-۳-۳-۵ اثربخشی نوع پیش‌پردازش داده

برای تعیین میزان اثربخشی نوع پیش‌پردازش بر نتایج به دست آمده، آزمایش‌ها با بهترین پیکربندی تکرار شد (طبق پارامترهای جدول ۷-۵). در این آزمایش‌ها تنها نوع پیش‌پردازش تغییر کرده و پارامترهای مدل و ابپارامترها ثابت هستند. برای هر تکنیک پیش‌پردازش، جهت اطمینان از نتایج به دست آمده، آزمایش‌ها ۱۰ بار تکرار شدند. جدول ۸-۵ نتایج به دست آمده را نشان می‌دهد.

جدول ۸-۵: نتایج بررسی اثربخشی نوع پیش‌پردازش داده. علامت † نشان می‌دهد نتایج به دست آمده در مقایسه با روش پیش‌پردازش "نگهداشت داده‌های اصلی" به صورت معنی‌دار ( $p < 0.005$ ) بهتر هستند. علامت \* نشان می‌دهد نتایج به دست آمده در مقایسه با روش پیش‌پردازش "تمیز کردن کامل داده‌ها" به صورت معنی‌دار ( $p < 0.005$ ) بهتر هستند.

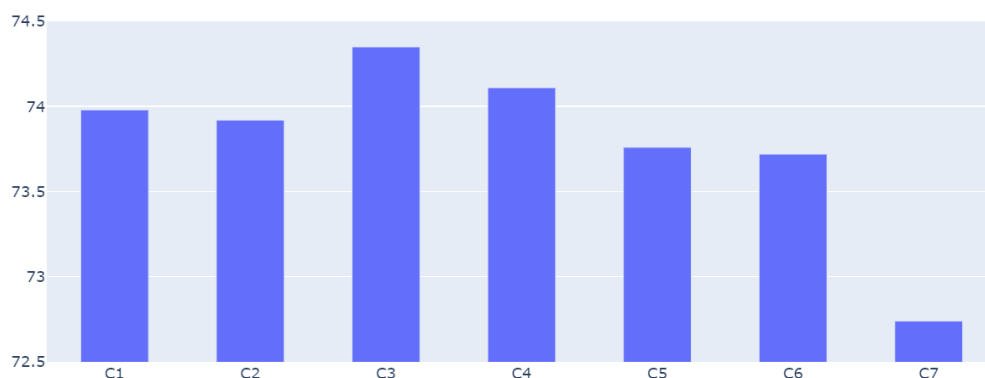
<i>F1 – Score</i>	روش پیش‌پردازش
$73.98 \pm 0.0012^*$	نگهداشت داده‌های اصلی
$73.92 \pm 0.0017^*$	حذف آدرس اینترنتی
$74.35 \pm 0.0015^{*\dagger}$	حذف نام کاربری
$74.11 \pm 0.0029^{*\dagger}$	حذف همزمان نام کاربری و آدرس اینترنتی
$73.76 \pm 0.0014^*$	حذف همزمان نام کاربری و آدرس اینترنتی و جدا کردن هشتک
$73.72 \pm 0.0009^*$	حذف همزمان نام کاربری و آدرس اینترنتی و جدا کردن هشتک و کوچک کردن حروف
$72.42 \pm 0.0020$	تمیز کردن کامل داده‌ها

با ارزیابی مقدار F1-Score به دست آمده می‌توان نتیجه گرفت روش "حذف نام کاربری" و "حذف همزمان نام کاربری و آدرس اینترنتی" در مقایسه با سایر روش‌ها بهترین نتیجه را کسب کردند. همچنین این روش‌ها به صورت معنادار از روش‌های "تمیز کردن کامل داده‌ها" و "نگهداشت داده‌های اصلی" بهتر است. مقدار  $p - value$  به دست آمده در مقایسه این روش‌ها با سایر تکنیک‌ها کمتر از 0.0005 می‌باشد که نشان می‌دهد بهبود معنادار<sup>۲۴</sup> است. همچنین موثرترین روش پیش‌پردازش در مجموعه داده استفاده شده، حذف نام کاربری می‌باشد. علاوه بر این می‌توان گفت همه روش‌ها نسبت به روش "تمیز کردن کامل داده‌ها" نتایج بهتری کسب کردند. این بدان معنی است در این روش ویژگی‌هایی<sup>۲۵</sup> از ورودی که می‌تواند تاثیر مثبتی بر عملکرد مدل داشته باشد حذف شده است. به عنوان مثال گاهی چندین علامت سوال یا علامت تعجب متوالی (!!!)، (؟؟؟؟؟) نشان دهنده یک حالت خاصی از تعجب یا عصبانیت و یا مخالفت است. اما در روش "تمیز کردن کامل داده‌ها" با حذف علائم نگارشی، فرصت یادگیری با استفاده از این چنین موارد از مدل سلب می‌شود.

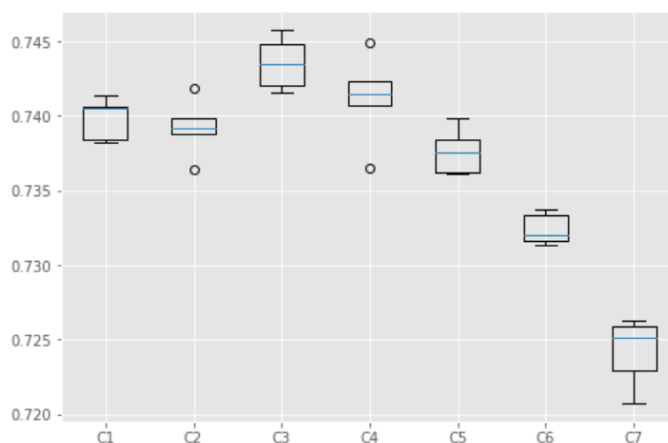
<sup>24</sup>Significance

<sup>25</sup>Feature

نمودار ۵-۵ نشان می‌دهد از روش اول (نگهداشت داده‌های اصلی) تا روش سوم (حذف نام‌کاربری) تکنیک‌های پیش‌پردازش باعث بهبود مدل می‌شود. اما بعد از آن هر چه موارد بیشتری از متن ورودی حذف شود، عملکرد مدل به صورت معناداری کاهش می‌یابد. میانگین و بازه اطمینان مدل پیشنهادی در شکل ۵-۶ را برای انواع تکنیک‌های پیش‌پردازش نشان می‌دهد.



شکل ۵-۵: مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده



شکل ۵-۶: بازه اطمینان مدل پیشنهادی بر اساس روش‌های پیش‌پردازش داده بر حسب معیار F1-Score

### ۵-۳-۳-۲ اثربخشی افزایش داده

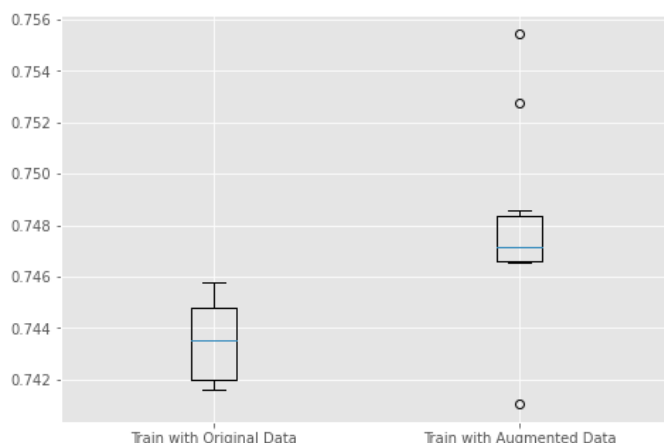
افزایش داده یکی از تکنیک‌های بهبود عملکرد مدل در مواجهه با مجموعه داده نامتوازن است. در این آزمایش همه پارامترها و فرآیندهای مدل یکسان است و تنها داده‌های آموزشی تفاوت دارند. آزمایش‌ها به ازای

## فصل ۵. روش پیشنهادی برای تشخیص موضع با نظارت ۳-۵. آزمایش‌ها و تحلیل نتایج

داده‌های افزایش یافته و داده‌های اصلی ۱۰ بار تکرار شدند. جدول ۹-۵ تاثیر افزایش داده‌های آموزشی بر عملکرد نهایی مدل را نشان می‌دهد. بازه اطمینان مدل آموزش دیده با داده‌های افزایش یافته در شکل ۷-۵ نمایش داده شده است. همچنین برای بررسی دقیق‌تر، میزان تغییر معیارهای ارزیابی به ازای هر کلاس در جدول ۱۰-۵ نشان داده شده است.

جدول ۹-۵: نتایج بررسی اثربخشی افزایش داده. علامت † نشان می‌دهد نتایج به دست آمده در مقایسه با حالت بدون افزایش داده به صورت معنی‌دار ( $p < 0.05$ ) بهتر هستند.

$F1 - Score$	افزایش داده	روش پیش‌پردازش
$74.11 \pm 0.0029$	-	حذف نام کاربری
$74.79 \pm 0.0036^\dagger$	+	



شکل ۷-۵: بازه اطمینان مدل پیشنهادی آموزش دیده با داده‌های افزایش داده شده بر حسب معیار F1-Score

جدول ۱۰-۵: نتایج بررسی اثربخشی افزایش داده به ازای هر کلاس

$F1 - Score$	$Recal$	$Precision$	کلاس	افزایش داده
77.25	78.19	76.33	موافق	-
83.48	93.80	75.17	مخالف	
59.61	56.90	62.6	بدون نظر	
<u>79.51</u>	77.58	81.54	موافق	+
<u>85.39</u>	95.57	76.59	مخالف	
<u>60.95</u>	61.16	59.80	بدون نظر	

با بررسی نتایج می‌توان گفت افزایش داده‌های آموزشی عملکرد نهایی مدل را به صورت معناداری بهبود

## فصل ۵. روش پیشنهادی برای تشخیص موضع با نظارت ۳-۵. آزمایش‌ها و تحلیل نتایج

داده است. همچنین با بررسی F1-Score به ازای هر کلاس می‌توان گفت پیش‌بینی همه کلاس‌ها نیز بهبود قابل قبولی داشتند. بدین ترتیب مدل با دیدن توزیع مناسبی از داده‌ها در زمان آموزش، عملکرد مناسب‌تری در زمان ارزیابی از خود نشان داده است. همچنین می‌توان نتیجه گرفت تکنیک انتخاب شده برای افزایش داده (بخش ۲-۲-۵) داده‌های با کیفیتی تولید کرده است.

### ۳-۳-۳-۵ اثربخشی رده‌بند

در این بخش میزان اثربخشی نوع رده‌بند در عملکرد مدل را مورد بررسی قرار گرفت. از آنجا که طبق نتایج جدول ۸-۵ دو سطح از پیش‌پردازش به صورت معنادار از بقیه روش‌ها بهتر بودند، آزمایش‌های این بخش تنها با دو روش پیش‌پردازش تکرار شد. همچون قسمت قبل، پارامترهای مدل و ابرپارامترها ثابت هستند و تنها نوع رده‌بند تغییر می‌کند. علاوه بر این جهت اطمینان از نتایج به دست آمده، آزمایش‌ها ۱۰ بار تکرار شد. جدول ۱۱-۵ نتایج به دست آمده را نشان می‌دهد.

جدول ۱۱-۵: نتایج بررسی اثربخشی رده‌بند. علامت † نشان می‌دهد نتایج به دست آمده در مقایسه با رده‌بند FNN به صورت معنی‌دار ( $p < 0.05$ ) بهتر هستند.

رده‌بند	روش پیش‌پردازش	$F1 - Score$
CNN	حذف نام‌کاربری	$74.35 \pm 0.0015^\dagger$
	حذف همزمان نام کاربری و آدرس اینترنتی	$74.11 \pm 0.0029$
FNN	حذف نام‌کاربری	$73.73 \pm 0.0058$
	حذف همزمان نام کاربری و آدرس اینترنتی	$73.91 \pm 0.0045$

نتایج نشان می‌دهد رده‌بند ساخته شده با شبکه عصبی پیچشی (CNN) به صورت معناداری از رده‌بند شبکه عصبی تماماً متصل عملکرد بهتری دارد. از آنجا که رده‌بند CNN کانولوشن‌هایی با سایزهای مختلف و به صورت موازی اعمال می‌کند، احتمال یادگیری از همسایگی با سایزهای مختلف را پیدا کرده و عملکرد بهتری از خود نشان می‌دهد.

### ۴-۳-۳-۵ اثربخشی تابع ضرر

جدول ۱۲-۵ تاثیر تغییر تابع ضرر بر نتایج مدل نهایی را نمایش می‌دهد. آزمایش‌های این بخش به ازای هر یک از تنظیمات ۱۰ بار تکرار شدند. سایر پارامترها و ابرپارامترهای مدل ثابت می‌باشد (طبق جدول ۷-۵).

## فصل ۵. روش پیشنهادی برای تشخیص موضع با نظارت ۵-۳. آزمایش‌ها و تحلیل نتایج

جدول ۵-۱۲: نتایج بررسی اثربخشی تابع ضرر بر حسب F1-Score. علامت † نشان می‌دهد نتایج به دست آمده در مقایسه با تابع ضرر Weighted Cross Entropy به صورت معنی‌دار ( $p < 0.05$ ) بهتر هستند. Cross Entropy به اختصار CE نوشته شده است.

روش پیش‌پردازش	Focal	Weighted CE	Weighted CE + Focal
نگهداشت داده‌های اصلی	$73.95 \pm 0.0034$	$73.98 \pm 0.0012$	$74.15 \pm 0.0035$
حذف آدرس اینترنتی	$74.03 \pm 0.0011$	$73.92 \pm 0.0017$	$74.01 \pm 0.0021$
حذف نام کاربری	$74.13 \pm 0.0015$	$74.35 \pm 0.0015$	$74.31 \pm 0.0031$
حذف نام کاربری و آدرس اینترنتی	$73.14 \pm 0.0032$	$74.11 \pm 0.0029$	$74.35 \pm 0.0025$
تمیز کردن کامل داده‌ها	$73.12 \pm 0.0028$	$72.42 \pm 0.0020$	$73.03 \pm 0.0035^\dagger$

همان‌طور که مشخص است استفاده از جمع دو تابع ضرر Focal و Weighted Cross Entropy عملکرد بهتری نسبت به استفاده از یکی از آن‌ها دارد. برای حالت پیش‌پردازش "تمیز کردن کامل داده‌ها" این بهبود معنادار است. در این حالت مقدار  $p$ -value به دست آمده کوچکتر 0.05 می‌باشد. در سایر موارد با مقایسه مقدار F1-Score می‌توان گفت مدل بهبود داشته اما این بهبود معنی‌دار نمی‌باشد.

### ۵-۳-۳-۵ بررسی تاثیر دامنه بر نتایج مدل آموزش دیده

یکی دیگر از بررسی‌های صورت گرفته میزان Domain Specific بودن مدل آموزش دیده می‌باشد. برای بررسی دقیق‌تر، دو موضوع تغییرات آب و هوایی و قانونی شدن سقط جنین از مجموعه داده SemEval-2016 انتخاب شدند. به این منظور مدل‌های آموزش دیده بر روی مجموعه داده ClimaConvo بر دو قسمت از مجموعه داده SemEval2016 مورد ارزیابی قرار گرفتند. نتایج به دست آمده در جدول ۵-۱۳ قابل مشاهده است.

جدول ۵-۱۳: بررسی تاثیر دامنه بر نتایج مدل آموزش دیده بر اساس معیار F1-Score.

روش پیش‌پردازش	ClimaConvo	SemEval-CC	SemEval-LA
نگهداشت داده‌های اصلی	$73.98 \pm 0.0012$	$30.89 \pm 0.0216$	$29.65 \pm 0.0142$
حذف آدرس اینترنتی	$73.92 \pm 0.0017$	$29.97 \pm 0.0231$	$31.05 \pm 0.0074$
حذف نام کاربری	$74.35 \pm 0.0015$	$30.53 \pm 0.0215$	$29.11 \pm 0.0104$
حذف همزمان نام کاربری و آدرس اینترنتی	$74.11 \pm 0.0029$	$30.34 \pm 0.0187$	$30.40 \pm 0.0088$
تمیز کردن کامل داده‌ها	$72.42 \pm 0.0020$	$30.18 \pm 0.0328$	$29.05 \pm 0.0142$

در مجموعه داده SemEval تاثیر نوع پیش‌پردازش کمتر می‌باشد. همچنین موضوع مجموعه داده ClimaConvo و SemEval-CC یکسان است اما عملکرد مدل بر روی آن بسیار ضعیف است. به نظر می‌رسد گذر زمان باعث شده مدل آموزش دیده بر روی مجموعه داده ClimaConvo بر مجموعه داده SemEval-CC عملکرد مناسبی



## فصل ۵. روش پیشنهادی برای تشخیص موضع با نظارت ۵-۳. آزمایش‌ها و تحلیل نتایج

نداشته باشد. به عبارت دیگر گذر زمان موجب تغییر در نوع داده‌ها شده و بنابراین مدل آموزش دیده بر داده‌های جدید (جمع‌آوری شده در سال ۲۰۲۲) مناسب داده‌های قدیمی جمع‌آوری شده (در سال ۲۰۱۶) نمی‌باشد. از طرف دیگر عملکرد مدل بر موضوع قانونی شدن سقط جنین چندان مناسب نیست. به طور کلی می‌توان گفت مدل برای داده‌های با موضوع و توزیع متفاوت نمی‌تواند عملکرد مناسبی داشته باشد.

### ۵-۳-۴ تحلیل خطا

با تحلیل خروجی به دست آمده می‌توان گفت مدل در پیشبینی کلاس بدون نظر ضعیف است. شکل ۵-۸ نشان می‌دهد که در ۲۰۹ نمونه کلاس موافق به اشتباه بدون نظر پیشبینی شده است. همچنین در ۱۴۹ نمونه کلاس بدون نظر به اشتباه موافق پیشبینی شده است. همچنین گاهی در عبارات کلاس مخالف که در قالب طعنه یا کنایه مطرح شده، مدل در تشخیص کلاس درست دچار خطا میشد. در این نوع توییت، نویسنده متن با کنایه مخالفت خود را نسبت به موضوع اهمیت تغییرات اقلیمی بیان می‌کند. در حالی که مدل نتوانسته بود این موضوع را به درستی درک کند. در جدول ۵-۱۴ نمونه‌ای از اشتباهات در خروجی قابل مشاهده است.



شکل ۵-۸: ماتریس درهم‌ریختگی برای داده ارزیابی مجموعه ClimConv

جدول ۵-۱۴: چند نمونه پیشبینی‌های اشتباه مدل

متن توییت	نوع خطا	کلاس اصلی	کلاس پیش‌بینی شده
#FridaysForFuture #ClimateChange#ExtinctionRebellion #GlobalWarming What are we saving?	کنایه	مخالف	بدون نظر

برای بررسی کامل‌تر علت خطاهای مدل، با استفاده از ابزار Transformers Interpret<sup>۲۶</sup> تحلیلی از خروجی تولید شده توسط مدل صورت گرفت. شکل ۵-۹ دو مثال از پیشبینی صحیح مدل را نشان می‌دهد. رنگ سبز نشان می‌دهد که برای پیش‌بینی کلاس نهایی، توجه بیشتری به آن کلمه شده است. نکته قابل توجه این است که برای هر دو خروجی مدل توانسته به کلمات تاثیرگذار توجه کند. به عبارت دیگر مدل کلمات مهم و تاثیرگذار برای فهم متن را توانسته به درستی تشخیص دهد. بنابراین همان‌طور که انتظار می‌رفت پیشبینی مدل نیز صحیح می‌باشد. در مقابل گاهی کلمات مورد توجه مدل برای پیشبینی صحیح کافی نیست. به عبارت دیگر نیاز به فهم بیشتری دارد. در شکل ۵-۱۰ دو نمونه از پیشبینی نادرست مدل مشاهده می‌شود. در مثال دوم اگر کلمات dying و don't care نیز مورد توجه مدل قرار می‌گرفت، احتمال پیشبینی درست توسط مدل افزایش می‌یافت.

Actual Label : **Support** [CLS] young people are power ##ing # climate ##act ##ion from every corner of the world , it 's time to give them the right sits on the  
predicted Label : **Support** table [SEP]

Actual Label : **Against**  
Predicted Label : **Against** [CLS] you 've been fooled by greta th ##un ##berg [SEP]

شکل ۵-۹: خروجی میزان توجه مدل به کلمات ورودی برای پیشبینی‌های درست

Actual Label : **Support** [CLS] even @ sec ##na ##v ##75 agrees that # climate ##cr ##isi ##s is " one of the great national security issues we face " ... then , why have  
Predicted label : **Nuetral** our leaders continuously turned a blind eye to the crisis ? were , and are , they sleep ##walk ##ing ? [SEP]

Actual Label : **Against**  
Predicted Label : **Nuetral** [CLS] yes , the earth is dying . no , i don 't care , i 'm old # fridays ##for ##tu ##ture [SEP]

شکل ۵-۱۰: خروجی میزان توجه مدل به کلمات ورودی برای پیشبینی‌های اشتباه

### ۵-۳-۵ مزایا و معایب روش پیشنهادی

روش پیشنهادی حل مسئله شامل یک تکنیک گام به گام برای طراحی معماری پیشنهادی و بررسی اثربخشی آن است. با توجه به فضای جستجو بررسی شده، می‌توان گفت این روش توانست به نتایج بسیار قابل قبولی دست پیدا کند. همچنین تعریف فضای جستجو به صورت مناسب نیز از دیگر دلایل موفقیت روش پیشنهادی است (استفاده از CNN در رده‌بند). از سوی دیگر بررسی انواع پیش‌پردازش متن در اکثر پژوهش‌ها مورد غفلت واقع می‌شود. اما مزیت این پژوهش توجه به روش پیش‌پردازش و بررسی انواع آن است.

<sup>26</sup><https://github.com/cdpierce/transformers-interpret>

از جمله معایب روش ارائه شده نیاز به داده آموزشی کافی برای آموزش مدل می‌باشد. از طرف دیگر طبق بررسی‌های انجام شده، مدل آموزش دیده Domain Specific می‌باشد و عملکرد مناسبی در سایر دامنه‌ها که در آن آموزش داده نشده، از خود نشان نداده است. وجود GPU با رم مناسب برای آموزش مدل‌ها ضروری است.

### ۵-۳-۶ نحوه پیاده‌سازی و اجرا آزمایش‌ها

پیاده‌سازی پروژه در اینجا<sup>۲۷</sup> قابل مشاهده می‌باشد که با زبان پایتون و فریم‌ورک پارتورچ انجام شده است. برای اجرا آزمایش‌ها لازم است با دستور `pip3 install -r requirements.txt` پکیج‌های مورد نیاز برای اجرا کد را نصب کنید. با توجه به طولانی بودن زمان اجرا کل آزمایش‌ها بهتر است با دستور `screen` یک اسکرین جدید برای اجرا بسازید و به شکل معمول دستورات را اجرا کنید. در این صورت حتی با بستن ترمینال اجرا ادامه پیدا می‌کند. با `screen -r` در صورت باز کردن مجدد ترمینال می‌توانید فرآیند و میزان پیشرفت اجرا را ببینید. راه حل دیگر برای تداوم اجرا در صورت بستن ترمینال استفاده از دستور `nohup` است. کافیست `nohup bash stance_run.bash` را اجرا کنید. خروجی‌های اجرا در `nohup.out` قابل مشاهده می‌باشد. در صورتی که یک `screen` جدید برای خود ساختید، با دستور `bash stance_run.bash` کلیه آزمایش‌ها به صورت متوالی اجرا می‌شود. در نهایت نتایج به دست آمده در پوشه `result` ذخیره می‌شود. همچنین وزن‌های شبکه در پوشه `models` قابل مشاهده می‌باشند.

برای اجرا از سخت‌افزار `GPU.1080Ti.xlarge` با رم `31.3GB` استفاده شد. دستور `nvidia-smi` میزان استفاده از GPU را نمایش می‌دهد. خروجی این دستور در هنگام آموزش مدل به صورت شکل ۵-۱۱ شد.

<sup>27</sup>[https://github.com/ghazaleh-mahmoodi/Climate\\_Activism\\_Stance\\_Detection](https://github.com/ghazaleh-mahmoodi/Climate_Activism_Stance_Detection)

```
Thu Feb 1 13:15:22 2024
```

NVIDIA-SMI 455.32.00				Driver Version: 455.32.00				CUDA Version: 11.1			
GPU	Name	Persistence-M	Bus-Id	Disp.A	Volatile	Uncorr.	ECC				
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute	M.				
						MIG	M.				
0	GeForce GTX 108...	Off	00000000:00:05.0	Off			N/A				
79%	83C	P2	251W / 250W	2674MiB / 11178MiB	87%	Default	N/A				

Processes:											
GPU	GI	CI	PID	Type	Process name			GPU Memory			
	ID	ID						Usage			
0	N/A	N/A	843	C	python			217MiB			
0	N/A	N/A	1118	C	python			2455MiB			

شکل ۵-۱۱: میزان مصرف GPU در زمان آموزش مدل پیشنهادی

## ۴-۵ جمع‌بندی

در این فصل به مسئله تشخیص موضع با نظارت پرداخته شد. ابتدا مجموعه داده‌های مورد استفاده معرفی شدند. در ادامه روش حل مسئله شامل تعریف اتولع حالات پیش‌پردازش داده، نحوه افزایش داده، روش جستجو برای معماری و ابرپارامترها و مثادیر فضای جستجو به صورت مفصل بیان شد. در انتها برای بررسی تاثیر نحوه پیش‌پردازش، افزایش داده، رده‌بند و تابع ضرر آزمایش‌های اصولی طراحی و انجام شد. در فصل بعد به مسئله تشخیص موضع بدون داده آموزشی پرداخته می‌شود.

## فصل ۶

# روش پیشنهادی برای تشخیص موضع بدون داده آموزشی

امروزه استفاده از مدل‌های زبانی بزرگ برای حل مسائل مختلف پردازش زبان طبیعی بسیار فراگیر شده است. با استفاده از این مدل‌ها می‌توان بدون داده آموزشی یا با داده آموزشی کم و با طراحی یک پرامپت مناسب<sup>۱</sup> نتایج قابل قبولی به دست آورد. در این پژوهش، به منظور استفاده از مدل‌های زبانی بزرگ در مسئله تشخیص موضع، چهار پرامپت با رویکردهای متفاوت پیشنهاد شد. پرامپت‌ها به صورت سلسله مراتبی تعریف شدند. بدین معنا که هر پرامپت اطلاعات بیشتری نسبت به پرامپت قبلی به عنوان ورودی دریافت می‌کنند. در ادامه هر یک از رویکردهای پیشنهادی به تفصیل توضیح داده می‌شود.

### ۱-۶ روش حل مسئله

در این پژوهش ۴ پرامپت در سطح‌های متفاوت تعریف شدند. در طراحی پرامپت از قواعد پیشنهاد شده توسط بزرگ‌ترین توسعه دهندگان مدل‌های زبانی بزرگ (همچون OpenAI) استفاده کردیم<sup>۲</sup>. اصول استفاده شده در طراحی پرامپت عبارتند از:

۱. نوشتن درخواست به صورت ساده و واضح و دقیق: باید ورودی و خروجی و مسئله‌ای که مدل قصد حل کردن آن را دارد به صورت کاملاً واضح و شفاف توضیح داده شود. هر چه قدر توضیحات دقیق‌تر و کامل‌تری در پرامپت باشد، احتمال رسیدن به پاسخ صحیح بیشتر می‌شود. همچنین مشخص کردن

---

<sup>۱</sup>Prompt Engineering

<sup>۲</sup><https://platform.openai.com/docs/guides/prompt-engineering/six-strategies-for-getting-better-results>

دقیق خروجی درخواستی از مدل، باعث بهبود نتایج می‌شود. وجود چند مثال از حل مسئله (رویکرد Few Shot [۹]) نیز می‌تواند ایده مناسبی برای بهره‌مندی بیشتر مدل از پرامپت ورودی باشد.

۲. تقسیم کردن مسئله اصلی به چند زیر مسئله ساده‌تر: نرخ پاسخگویی اشتباه در مسائل پیچیده بسیار بالاست. بنابراین یک راه حل بهینه این است که مسئله اولیه، در قالب چند زیر مسئله ساده‌تر تعریف شود. در ادامه گام‌های حل هر مسئله به صورت شفاف توضیح داده شود تا احتمال رسیدن به پاسخ مطلوب افزایش یابد (رویکرد Chain of Thought از این قاعده استفاده می‌کند). در این صورت به مدل زمان لازم برای استنتاج داده می‌شود و پاسخ‌های قابل اعتمادتری تولید می‌شود.

با توجه به اصول مطرح شده به عنوان راهنما برای طراحی پرامپت، چهار پرامپت در سطوح مختلف با رویکردهای متفاوت تعریف شدند. در ادامه به توضیح چهار پرامپت پیشنهادی این پژوهش پرداخته می‌شود.

### ۱-۱-۶ پرامپت Zero Shot

در این پرامپت ساده‌ترین روش در نظر گرفته شد. پرامپت متشکل از توضیح مسئله، وظیفه خواسته شده از مدل، ورودی‌های مسئله تشخیص موضع شامل موضوع<sup>۳</sup> و متن توییت و مشخص نمودن کلاس‌های خروجی می‌باشد. در نوشتن پرامپت توضیحات به صورت دقیق و ساده ارائه شدند. شکل ۱-۶ قالب پیشنهادی را نشان می‌دهد.

The following statements (tweet) are twitter posts about target target. Each statement can support, oppose, or be neutral toward its associated target belief and Each statement has the reason for that stance.

Now, classify the following statement as to whether support, oppose, and neutral toward the belief below being true, and give your explanation for the classification.

Target: <target>

Tweet: <tweet>

Stance : select 'Stance' ['support', 'oppose', 'neutral']

Explanation: gen 'explanation'

شکل ۱-۶: الگو پرامپت Zero Shot

<sup>3</sup>Target

## ۲-۱-۶ پرامپت Zero Shot + Chain of Thought

این رویکرد، در ادامه پرامپت Zero Shot پیشنهاد شده است. در این پرامپت توضیحات مربوط به نحوه حل مسئله گام به گام توضیح داده شده به طوری که مدل زمان لازم برای استدلال را داشته باشد. همچنین با تقسیم مسئله به زیر مسئله‌های ساده‌تر مدل برای دستیابی به پاسخ صحیح راهنمایی می‌شود. شکل ۲-۶ قالب پیشنهادی را نشان می‌دهد. قسمت مربوط به Chain of Thought (CoT) با رنگ آبی مشخص شده است.

The following statements (tweet) are twitter posts about target target. Each statement can support, oppose, or be neutral toward its associated target belief and Each statement has the reason for that stance.

Now, classify the following statement as to whether support, oppose, and neutral toward the belief below being true, and give your explanation for the classification.

let's think step by step.

First, check if the author of the tweet has commented on climate change. Has he commented at all?

Then check if the author supports that climate change is a concern and should be addressed. Or is the author's opinion that this matter does not require special attention?

Target: < target >

Tweet: < tweet >

Stance : select 'Stance' ['support', 'oppose', 'neutral']

Explanation: gen 'explanation'

شکل ۲-۶: الگو پرامپت Zero Shot + Chain of Thought

## ۳-۱-۶ پرامپت Zero Shot + Chain of Thought + Context Description

این پرامپت، در ادامه پرامپت Zero Shot + Chain of Thought پیشنهاد شد. نقطه ضعف پرامپت‌های قبلی عدم وجود توضیحاتی راجع به موضوع تغییرات اقلیمی است و تنها تعریف مسئله تشخیص موضع در متن پرامپت وجود دارد. در این پرامپت سعی شده توضیحاتی راجع به موضوع مورد بحث نیز به عنوان ورودی در نظر گرفته شود. توضیحات برگرفته از صفحه ویکی‌پدیا مربوطه می‌باشد. شکل ۳-۶ قالب پیشنهادی را نشان می‌دهد. قسمت مربوط به Context با رنگ سبز مشخص شده است.

The following statements (tweet) are twitter posts about Climate Change. Climate change describes global warming—the ongoing increase in global average temperature—and its effects on Earth’s climate system. Climate change in a broader sense also includes previous long-term changes to Earth’s climate. Fossil fuel use, deforestation, and some agricultural and industrial practices add to greenhouse gases, notably carbon dioxide and methane. Greenhouse gases absorb some of the heat that the Earth radiates after it warms from sunlight. Each statement can support, oppose, or be neutral toward its associated target belief and Each statement has the reason for that stance.

Now, classify the following statement as to whether support, oppose, and neutral toward the belief below being true, and give your explanation for the classification.

let’s think step by step.

First, check if the author of the tweet has commented on climate change. Has he commented at all?

Then check if the author supports that climate change is a concern and should be addressed. Or is the author’s opinion that this matter does not require special attention?

Target: < *target* >

Tweet: < *tweet* >

Stance : select ‘Stance’ [‘support’, ‘oppose’, ‘neutral’]

Explanation: gen ‘explanation’

شکل ۳-۶: الگو پرامپت Zero Shot + Chain of Thought + Context Description

#### ۴-۱-۶ پرامپت Few Shot + Chain of Thought + Context Description

تا این مرحله اطلاعات مناسبی به متن پرامپت ورودی اضافه شد. پرامپت پیشنهاد شده در این مرحله از رویکرد Few Shot + Chain of Thought + Context Description پیروی می‌کند. در این پرامپت علاوه بر ورودی‌های قبلی، چند مثال از متن توییت و موضع توییت آورده شده است. در این روش سعی شده با دیدن چند نمونه آموزشی، بهبودی در عملکرد مدل ایجاد کنیم. شکل ۴-۶ قالب پیشنهادی را نشان می‌دهد. نمونه‌های آموزشی در این پرامپت به صورت تصادفی انتخاب شدند. قسمت مربوط به Few Shot با رنگ **قهوه‌ای** مشخص شده است.



The following statements (tweet) are twitter posts about Climate Change. Climate change describes global warming—the ongoing increase in global average temperature—and its effects on Earth’s climate system. Climate change in a broader sense also includes previous long-term changes to Earth’s climate. Fossil fuel use, deforestation, and some agricultural and industrial practices add to greenhouse gases, notably carbon dioxide and methane. Greenhouse gases absorb some of the heat that the Earth radiates after it warms from sunlight. Each statement can support, oppose, or be neutral toward its associated target belief and Each statement has the reason for that stance.

Let’s see some example.

the tweet < *tweet* > stance on Climate change is supportive.

the tweet < *tweet* > stance on Climate change is opposite.

the tweet < *tweet* > stance on Climate change is neutral.

Now, classify the following statement as to whether support, oppose, and neutral toward the belief below being true, and give your explanation for the classification.

let’s think step by step

First, check if the author of the tweet has commented on climate change. Has he commented at all? Then check if the author supports that climate change is a concern and should be addressed. Or is the author’s opinion that this matter does not require special attention?

Target: < *target* >

Tweet: < *tweet* >

Stance : select ‘Stance’ [‘support’, ‘oppose’, ‘neutral’]

Explanation: gen ‘explanation’

شکل ۴-۶: الگو پرامپت Few Shot + Chain of Thought + Context Description

## ۲-۶ آزمایش‌ها و تحلیل نتایج

برای انجام آزمایش‌ها از مجموعه داده ClimaConvo [۴۵] و SemEval [۴۰] و مدل Orca [۳۸] استفاده شد. به دلیل عملکرد مناسب مدل Orca در مقایسه با سایر مدل‌هایی با تعداد پارامتر مشابه از این مدل برای اجرا آزمایش‌ها استفاده شد (با توجه به سخت‌افزار موجود امکان استفاده از مدل‌های زبانی با پارامترهای بیشتر میسر نیست). در اجرا آزمایش‌ها پارامتر دما<sup>۴</sup> برابر با صفر قرار گرفت تا همواره توکن با بیشترین احتمال را تولید کند. از آنجایی که نتایج مدل‌های Zero Shot برای مجموعه داده SemEval موجود بود، ابتدا مقایسه‌ای از روش‌های موجود شامل تکنیک‌های یادگیری عمیق و پرامپت با روش معرفی شده انجام شد. در جدول ۱-۶ پرامپت‌های پیشنهادی با سایر روش‌های یادگیری عمیق با رویکرد یادگیری بدون داده آموزشی مورد مقایسه قرار گرفت. نتایج به دست آمده نشان می‌دهد، در موضوع ”قانونی شدن سقط جنین (LA)” پرامپت Zero Shot از سایر روش‌های قبلی به صورت معناداری بهتر است. با توجه به نوع متن ورودی پرامپت، تنها رویکرد Zero Shot برای این موضوع قابل استفاده بود. در موضوع ”تغییرات اقلیمی (CC)” پرامپت Zero Shot عملکرد بهتری از خود نشان داده است و نسبت به نتایج Bicond و GPT3.5 به صورت معنادار بهتر است اما نسبت به سایرین عملکرد ضعیف‌تری دارد. به طور کلی قسمت CC در مجموعه داده SemEval در مقایسه با LA در اکثر مدل‌ها نتایج ضعیفی از خود نشان داده است. این نتایج نشان دهنده پیچیده‌تر بودن داده‌های این موضوع از مجموعه داده SemEval می‌باشد.

<sup>4</sup>Temperature

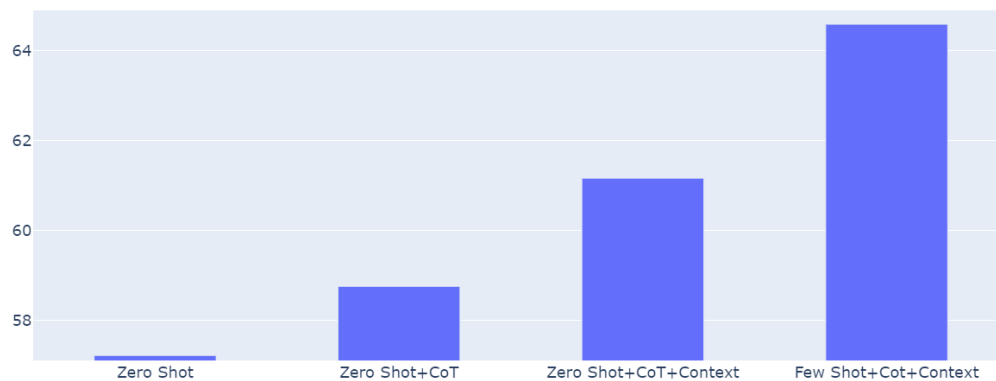
جدول ۶-۱: مقایسه نتایج بر اساس معیار F1-Score برای پرامپت‌های معرفی شده با سایر مدل‌های Zero-Shot بر مجموعه داده SemEval.

<i>CC</i>	<i>LA</i>	رویکرد
15.0	34.4	[۷] Bicond
-	34.4	[۶۳] CrossNet
-	44.6	[۶۷] SEKT
-	46.5	[۳۱] TPDG
37.3	44.8	[۱۸] Bert_Spc
35.5	44.2	[۳۳] Bert-GCN
38.9	50.9	[۳۰] PT-HCL
24.7	52.0	[۶۶] GPT-3.5
25.2	55.3	GPT-3.5+COT
23.65	57.33†	پرامپت Zero Shot
28.80†	-	پرامپت Zero Shot + CoT
19.30	-	پرامپت Zero Shot + CoT + Context
20.85	-	پرامپت Few Shot + Cot + Context

در ادامه عملکرد پرامپت‌های معرفی شده بر روی مجموعه داده ClimaConvo مورد بررسی قرار گرفت. هنگام انجام آزمایش‌های این قسمت، هیچ‌گونه پیش‌پردازشی در داده‌های متنی صورت نگرفته است. گزارش‌ها نشان می‌دهد (جدول ۶-۲) همه نتایج به دست آمده نسبت به روش قبلی خود در سلسله مراتب به صورت معنی‌دار بهبود داشتند. همان‌طور که از نتایج مشخص است، پرامپت با رویکرد Few Shot توانسته نتایج بسیار قابل قبولی به دست بیاورد. اضافه کردن توضیحات راجع به تغییرات اقلیمی نیز بهبود قابل توجهی در خروجی ایجاد کرده است که با مقایسه سطر ۲ و ۳ جدول ۶-۲ می‌توان به آن پی برد.

جدول ۶-۲: نتایج به دست آمده با استفاده از پرامپت‌های معرفی شده روی مجموعه داده ClimaConvo.

<i>F1 - Score</i>	رویکرد
57.21	پرامپت Zero Shot
58.75	پرامپت Zero Shot + CoT
61.16	پرامپت Zero Shot + CoT + Context
64.59	پرامپت Few Shot + CoT + Context

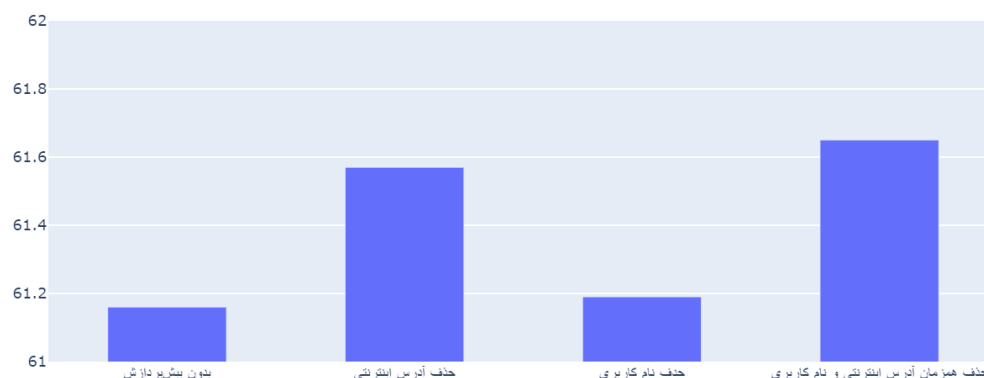


شکل ۶-۵: مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده

در ادامه برای بررسی اثربخشی پیش‌پردازش متن ورودی بر خروجی مدل زبانی بزرگ، آزمایش‌ها برای پرامپت Zero Shot + CoT + Context تکرار شد. در این آزمایش‌ها سه روش متفاوت برای پیش‌پردازش داده‌ها را که در فصل قبل عملکرد بهتری از خود نشان دادند مورد بررسی قرار گرفت. نتایج به دست آمده در جدول ۶-۳ گزارش شده است. در نمودار ۶-۶ نتایج به دست آمده به صورت شهودی قابل مقایسه است. همان‌طور که نتایج نشان می‌دهد، با تغییر روش پیش‌پردازش داده‌ها تغییر معنی‌داری در نتایج به دست آمده به وجود نمی‌آید.

جدول ۶-۳: تاثیر پیش‌پردازش داده‌ها در نتایج به دست آمده

$F1 - Score$	پیش‌پردازش
61.16	بدون پیش‌پردازش
61.57	حذف آدرس اینترنتی
61.19	حذف نام کاربری
61.65	حذف همزمان آدرس اینترنتی و نام کاربری



شکل ۶-۶: مقدار F1-Score هر یک از تکنیک‌های پیش‌پردازش داده

## ۱-۲-۶ مزایا و معایب روش پیشنهادی

از جمله مزایای روش پیشنهادی، پیش‌بینی خروجی بدون نیاز به داده آموزشی یا با وجود داده‌های آموزشی کم می‌باشد. بر خلاف روش فصل پنجم، استفاده از پرامپت Domain Specific نیست و بر دامنه‌های مختلف می‌توان نتایج قابل قبولی دریافت کرد.

از طرفی نیاز به منابع محاسباتی GPU با رم بالا از جمله معایب این روش می‌باشد. همچنین به دلیل اینکه پرامپت توسط نیروی انسانی طراحی می‌شود، ممکن است پرامپت کاملاً بهینه‌ای طراحی نشده باشد. علاوه امکان fine-tuning پرامپت‌های پیشنهادی برای مدل با توجه به منابع محاسباتی محدود، مقدور نمی‌باشد.

## ۲-۲-۶ نحوه پیاده‌سازی و اجرا آزمایش‌ها

پیاده‌سازی پروژه در اینجا<sup>۵</sup> قابل مشاهده می‌باشد که با زبان پایتون و فریم‌ورک پایتورچ انجام شده است. برای تعامل با مدل زبانی بزرگ از guidance<sup>۶</sup> استفاده شده است. هر بار پیش‌بینی برای کل داده‌های ارزیابی با توجه به نوع پرامپت حدود ۴۰ الی ۱۶۰ دقیقه زمان می‌برد. زمان تقریبی تولید خروجی برای یک ورودی حدود ۴ الی ۱۵ ثانیه می‌باشد.

برای اجرا از سخت‌افزار GPU.1080Ti.xlarge با رم 31.3GB استفاده شد. دستور nvidia-smi میزان استفاده از GPU را نمایش می‌دهد. خروجی این دستور در هنگام اجرا آزمایشات به صورت شکل ۷-۶ شد.

```
Thu Feb 8 17:16:51 2024
```

NVIDIA-SMI 455.32.00				Driver Version: 455.32.00				CUDA Version: 11.1			
GPU	Name	Persistence-M	Bus-Id	Disp.A	Volatile	Uncorr.	ECC				
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute	M.				
						MIG	M.				
0	GeForce GTX 108...	Off	00000000:00:05.0 Off		64%		N/A				
62%	74C	P2	156W / 250W	8447MiB / 11178MiB		Default	N/A				

Processes:							
GPU	GI	CI	PID	Type	Process name	GPU Memory	
	ID	ID				Usage	
0	N/A	N/A	29568	C	python	8445MiB	

شکل ۷-۶: میزان مصرف GPU در زمان اجرا آزمایش‌های مدل‌های زبانی بزرگ

<sup>5</sup>[https://github.com/ghazaleh-mahmoodi/Climate\\_Activism\\_Stance\\_Detection](https://github.com/ghazaleh-mahmoodi/Climate_Activism_Stance_Detection)

<sup>6</sup><https://github.com/guidance-ai/guidance>

## فصل ۷

## نتیجه‌گیری و کارهای آینده

## ۷-۱ نتیجه گیری

در این پژوهش روش‌های حل مسئله تشخیص موضع مورد بررسی قرار گرفت. در فصل پنجم، با الهام‌گیری از رویکرد جستجو معماری عصبی، برای طراحی مدل تشخیص موضع یک فضا جستجو تعریف شد. سپس یک جستجو قاعده‌مند صورت گرفت. همچنین اثربخشی روش پیش‌پردازش داده، افزایش داده، رده‌بند و تابع ضرر به صورت جداگانه مورد بررسی قرار گرفت. در نهایت کدگذار BERTweet و رده‌بند CNN به عنوان مدل پیشنهادی معرفی شد. افزایش داده و روش پیش‌پردازش با رویکرد حذف نام کاربری توانستند نتایج مدل را به صورت معناداری بهبود بدهند. رده‌بند CNN به صورت معناداری از FNN بهتر عمل می‌کند. روش جستجو ارائه شده عملکرد خوبی در طراحی یک معماری مناسب از خود نشان داد.

در فصل ششم، چهار پرامپت با رویکردهای متفاوت و سلسله مراتبی برای حل مسئله تشخیص موضع بدون داده آموزشی معرفی شد. در مجموعه داده ClimaConvo پرامپت Few Shot + Chain of Thought + Context بهترین نتیجه را در مقایسه با سایر رویکردها کسب کرد. در مجموعه داده SemEval پرامپت Zero Shot پیشنهادی توانست در موضوع ”قانونی شدن سقط جنین (LA)” مقدار 2.03 درصد در مقایسه با روش‌های دیگر Zero Shot بهبود داشته باشد. آزمایش‌ها نشان دادند، نحوه پیش‌پردازش داده در عملکرد مدل زبانی بزرگ تاثیر معناداری ایجاد نمی‌کند. از آن‌جا که جمع‌آوری داده به ازای همه موضوعات به راحتی میسر نمی‌باشد، با ادامه پژوهش در این زمینه (تشخیص موضع بدون داده آموزشی) می‌توان گامی در جهت استفاده از تشخیص موضع در کاربردهای زندگی روزمره برداشت.

## ۲-۷ پیشنهادها و کارهای آینده

برای ادامه پژوهش در مسئله تشخیص موضوع با نظارت موارد زیر پیشنهاد می شود.

۱. در این پژوهش تنها مدل‌های کدگذار (همچون BRET) مورد بررسی قرار گرفتند. در ادامه پیشنهاد می‌شود عملکرد مدل‌های کدگشا همانند GPT-2 و کدگذار-کدگشا همچون BART, T5 نیز مورد مقایسه قرار بگیرد.

۲. مدل‌های مبتنی بر درخت همچون XGBoost اخیراً به نتایج خوبی در مسائل رده‌بندی دست یافتند. پیشنهاد می‌شود در ادامه تحقیقات عملکرد این مدل نیز مورد بررسی قرار بگیرد. از جمله مزایای مدل‌های مبتنی بر درخت تفسیرپذیر بودن می‌باشد.

۳. برای داده‌های نامتوازن از تابع ضررهای متناسب استفاده می‌شود. در این پژوهش دو تابع ضرر معروف این حوزه مورد بررسی قرار گرفت. بررسی سایر تابع‌های ضرر از جمله dice پیشنهاد می‌شود. همچنین ترکیب تابع ضررهای متفاوت می‌تواند نتایج خوبی به ارمغان بیاورد.

۴. در بخش کدگذار می‌توان از چندین مدل از پیش‌آموزش دیده و Concat بازنمایی کلمات مربوط به آن‌ها استفاده کرد.

در رویکرد بدون داده آموزشی برای ادامه تحقیقات پیشنهادات زیر ارائه می‌شود.

۱. در پژوهش فعلی در پرامپت Few Shot نمونه‌های آموزشی به صورت تصادفی انتخاب شدند. پیشنهاد می‌شود با استفاده بازنمایی کلمات، به جای انتخاب تصادفی نمونه‌های ورودی نمونه‌هایی که بیشترین شباهت را با ورودی فعلی دارند به عنوان نمونه آموزشی انتخاب شوند.

۲. با استفاده از روش‌های Prefix-Tuning می‌توان با ثابت نگه‌داشتن وزن‌های مدل و تنها آموزش ۱ درصد وزن‌ها، نتایج بهتری با پرامپت تولیدی به دست آورد.

۳. تغییر پرامپت به صورتی که به ازای پیش‌بینی کلاس نهایی به صورت مستقیم، احتمال تعلق ورودی به هر کلاس را پیش‌بینی کند. در ادامه با ایده‌های سلسه مراتبی دیگر می‌توان کلاس نهایی را به دست آورد. به عنوان مثال می‌توان بین خروجی چند مدل زبانی بزرگ اجتماع گرفت.

- [1] ADDAWOOD, A., SCHNEIDER, J., AND BASHIR, M. Stance classification of twitter debates: The encryption debate as a use case. in *Proceedings of the 8th International Conference on Social Media & Society* (New York, NY, USA, 2017), Association for Computing Machinery.
- [2] AGERRI, R., CENTENO, R., ESPINOSA, M., DE LANDA, J. F., AND RODRIGO, A. Vaxxstance@iberlef 2021: overview of the task on going beyond text in cross-lingual stance detection. *Procesamiento del Lenguaje Natural* 67 (2021), 173–181.
- [3] AKIBA, T., SANO, S., YANASE, T., OHTA, T., AND KOYAMA, M. Optuna: A next-generation hyperparameter optimization framework. in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (New York, NY, USA, 2019), KDD '19, Association for Computing Machinery, p. 2623–2631.
- [4] ALDAYEL, A., AND MAGDY, W. Stance detection on social media: State of the art and trends. *Information Processing and Management* 58, 4 (2021), 102597.
- [5] ALLAWAY, E., AND MCKEOWN, K. Zero-shot stance detection: A dataset and model using generalized topic representations. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (2020).
- [6] ALTURAYEIF, N., LUQMAN, H., AND AHMED, M. A systematic review of machine learning techniques for stance detection and its applications. *Neural Computing and Applications* 35, 7 (Mar 2023), 5113–5144.
- [7] AUGENSTEIN, I., ROCKTÄSCHEL, T., VLACHOS, A., AND BONTCHEVA, K. Stance detection with bidirectional conditional encoding. in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (Austin, Texas, Nov. 2016), J. Su, K. Duh, and X. Carreras, eds. , Association for Computational Linguistics, pp. 876–885.

- [8] BARBIERI, F., CAMACHO-COLLADOS, J., ESPINOSA ANKE, L., AND NEVES, L. TweetEval: Unified benchmark and comparative evaluation for tweet classification. in *Findings of the Association for Computational Linguistics: EMNLP 2020* (Online, Nov. 2020), T. Cohn, Y. He, and Y. Liu, eds. , Association for Computational Linguistics, pp. 1644–1650.
- [9] BROWN, T. B., MANN, B., RYDER, N., SUBBIAH, M., KAPLAN, J., DHARIWAL, P., NEELAKANTAN, A., SHYAM, P., SASTRY, G., ASKELL, A., AGARWAL, S., HERBERT-VOSS, A., KRUEGER, G., HENIGHAN, T., CHILD, R., RAMESH, A., ZIEGLER, D. M., WU, J., WINTER, C., HESSE, C., CHEN, M., SIGLER, E., LITWIN, M., GRAY, S., CHESS, B., CLARK, J., BERNER, C., MCCANDLISH, S., RADFORD, A., SUTSKEVER, I., AND AMODEI, D. Language models are few-shot learners. in *Proceedings of the 34th International Conference on Neural Information Processing Systems* (Red Hook, NY, USA, 2020), NIPS’20, Curran Associates Inc.
- [10] CAN, U., AND ALATAS, B. A novel approach for efficient stance detection in online social networks with metaheuristic optimization. *Technology in Society* 64 (2021), 101501.
- [11] CHAUHAN, D. S., KUMAR, R., AND EKBAL, A. Attention based shared representation for multi-task stance detection and sentiment analysis. in *Neural Information Processing* (Cham, 2019), T. Gedeon, K. W. Wong, and M. Lee, eds. , Springer International Publishing, pp. 661–669.
- [12] CHEN, P., YE, K., AND CUI, X. Integrating n-gram features into pre-trained model: A novel ensemble model for multi-target stance detection. in *Artificial Neural Networks and Machine Learning – ICANN 2021* (Cham, 2021), I. Farkas, P. Masulli, S. Otte, and S. Wermter, eds. , Springer International Publishing, pp. 269–279.
- [13] CHOWDHARY, K. R. *Natural Language Processing*. Springer India, New Delhi, 2020, pp. 603–649.
- [14] CHUNG, J., GULCEHRE, C., CHO, K., AND BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. in *NIPS 2014 Workshop on Deep Learning, December 2014* (2014).
- [15] CIGNARELLA, A. T., LAI, M., BOSCO, C., PATTI, V., PAOLO, R., ET AL. Sardistance@evalita2020: Overview of the task on stance detection in italian tweets. in *Proceedings of the Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2020)* (2020), Ceur, pp. 1–10.



- [16] CONFORTI, C., BERNDT, J., PILEHVAR, M. T., GIANNITSAROU, C., TOXVAERD, F., AND COLLIER, N. Will-they-won't-they: A very large dataset for stance detection on Twitter. in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Online, July 2020), D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, eds. , Association for Computational Linguistics, pp. 1715–1724.
- [17] CONNEAU, A., KHADELWAL, K., GOYAL, N., CHAUDHARY, V., WENZKE, G., GUZMÁN, F., GRAVE, E., OTT, M., ZETTEMAYER, L., AND STOYANOV, V. Unsupervised cross-lingual representation learning at scale. *arXiv preprint arXiv:1911.02116* (2019).
- [18] DEVLIN, J., CHANG, M.-W., LEE, K., AND TOUTANOVA, K. BERT: Pre-training of deep bidirectional transformers for language understanding. in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (Minneapolis, Minnesota, June 2019), J. Burstein, C. Doran, and T. Solorio, eds. , Association for Computational Linguistics, pp. 4171–4186.
- [19] EBRAHIMI, J., DOU, D., AND LOWD, D. A joint sentiment-target-stance model for stance classification in tweets. in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (Osaka, Japan, Dec. 2016), Y. Matsumoto and R. Prasad, eds. , The COLING 2016 Organizing Committee, pp. 2656–2665.
- [20] FRAISIER, O., CABANAC, G., PITARCH, Y., BESANÇON, R., AND BOUGHANEM, M. Stance classification through proximity-based community detection. in *Proceedings of the 29th on Hypertext and Social Media* (New York, NY, USA, 2018), HT '18, Association for Computing Machinery, p. 220–228.
- [21] HARDALOV, M., ARORA, A., NAKOV, P., AND AUGENSTEIN, I. A survey on stance detection for mis- and disinformation identification. in *Findings of the Association for Computational Linguistics: NAACL 2022* (Seattle, United States, July 2022), M. Carpuat, M.-C. de Marneffe, and I. V. Meza Ruiz, eds. , Association for Computational Linguistics, pp. 1259–1277.
- [22] HOCHREITER, S., AND SCHMIDHUBER, J. Long short-term memory. *Neural Computation* 9, 8 (1997), 1735–1780.
- [23] KENTON, J. D. M.-W. C., AND TOUTANOVA, L. K. Bert: Pre-training of deep bidirectional transformers for language understanding. in *Proceedings of naacL-HLT* (2019), volume 1, p. 2.

- [24] KIM, Y. Convolutional neural networks for sentence classification. in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Doha, Qatar, Oct. 2014), A. Moschitti, B. Pang, and W. Daelemans, eds. , Association for Computational Linguistics, pp. 1746–1751.
- [25] KÜÇÜK, D., AND CAN, F. Stance detection: A survey. *ACM Comput. Surv.* 53, 1 (feb 2020).
- [26] KUCHER, K., MARTINS, R. M., PARADIS, C., AND KERREN, A. Stancevis prime: visual analysis of sentiment and stance in social media texts. *Journal of Visualization* 23, 6 (Dec 2020), 1015–1034.
- [27] KUCHER, K., PARADIS, C., SAHLGREN, M., AND KERREN, A. Active learning and visual analytics for stance classification with alva. *ACM Trans. Interact. Intell. Syst.* 7, 3 (oct 2017).
- [28] LI, Y., AND CARAGEA, C. Multi-task stance detection with sentiment and stance lexicons. in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (Hong Kong, China, Nov. 2019), K. Inui, J. Jiang, V. Ng, and X. Wan, eds. , Association for Computational Linguistics, pp. 6299–6305.
- [29] LI, Y., SOSEA, T., SAWANT, A., NAIR, A. J., INKPEN, D., AND CARAGEA, C. P-stance: A large dataset for stance detection in political domain. in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021* (Online, Aug. 2021), C. Zong, F. Xia, W. Li, and R. Navigli, eds. , Association for Computational Linguistics, pp. 2355–2365.
- [30] LIANG, B., CHEN, Z., GUI, L., HE, Y., YANG, M., AND XU, R. Zero-shot stance detection via contrastive learning. in *Proceedings of the ACM Web Conference 2022* (New York, NY, USA, 2022), WWW '22, Association for Computing Machinery, p. 2738–2747.
- [31] LIANG, B., FU, Y., GUI, L., YANG, M., DU, J., HE, Y., AND XU, R. Target-adaptive graph for cross-target stance detection. in *Proceedings of the Web Conference 2021* (New York, NY, USA, 2021), WWW '21, Association for Computing Machinery, p. 3453–3464.
- [32] LIN, T.-Y., GOYAL, P., GIRSHICK, R., HE, K., AND DOLLAR, P. Focal loss for dense object detection. in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (Oct 2017).

- [33] LIN, Y., MENG, Y., SUN, X., HAN, Q., KUANG, K., LI, J., AND WU, F. BertGCN: Transductive text classification by combining GNN and BERT. in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021* (Online, Aug. 2021), C. Zong, F. Xia, W. Li, and R. Navigli, eds. , Association for Computational Linguistics, pp. 1456–1462.
- [34] LIU, Y., OTT, M., GOYAL, N., DU, J., JOSHI, M., CHEN, D., LEVY, O., LEWIS, M., ZETTLEMOYER, L., AND STOYANOV, V. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).
- [35] LOUREIRO, D., BARBIERI, F., NEVES, L., ESPINOSA ANKE, L., AND CAMACHO-COLLADOS, J. TimeLMs: Diachronic language models from Twitter. in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (Dublin, Ireland, May 2022), Association for Computational Linguistics, pp. 251–260.
- [36] LOZHNIKOV, N., DERCZYNSKI, L., AND MAZZARA, M. Stance prediction for russian: Data and analysis. in *Proceedings of 6th International Conference in Software Engineering for Defence Applications* (Cham, 2020), P. Ciancarini, M. Mazzara, A. Messina, A. Sillitti, and G. Succi, eds. , Springer International Publishing, pp. 176–186.
- [37] MATERO, M., SONI, N., BALASUBRAMANIAN, N., AND SCHWARTZ, H. A. MeLT: Message-level transformer with masked document representations as pre-training for stance detection. in *Findings of the Association for Computational Linguistics: EMNLP 2021* (Punta Cana, Dominican Republic, Nov. 2021), Association for Computational Linguistics, pp. 2959–2966.
- [38] MATHUR, P. orca mini v3 7b: An explain tuned llama2-7b model. ”[https://huggingface.co/psmathur/orca\\_mini\\_v3\\_7b](https://huggingface.co/psmathur/orca_mini_v3_7b)”, 2023. Online; accessed 10-February-2024.
- [39] MIKOLOV, T., CHEN, K., CORRADO, G. S., AND DEAN, J. Efficient estimation of word representations in vector space, 2013.
- [40] MOHAMMAD, S., KIRITCHENKO, S., SOBHANI, P., ZHU, X., AND CHERRY, C. SemEval-2016 task 6: Detecting stance in tweets. in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (San Diego, California, June 2016), S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, and T. Zesch, eds. , Association for Computational Linguistics, pp. 31–41.

- [41] NGUYEN, D. Q., VU, T., AND NGUYEN, A. T. BERTweet: A pre-trained language model for English Tweets. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (2020), pp. 9–14.
- [42] PENNINGTON, J., SOCHER, R., AND MANNING, C. GloVe: Global vectors for word representation. in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Doha, Qatar, Oct. 2014), A. Moschitti, B. Pang, and W. Daelemans, eds. , Association for Computational Linguistics, pp. 1532–1543.
- [43] POPAT, K., MUKHERJEE, S., YATES, A., AND WEIKUM, G. DeClarE: Debunking fake news and false claims using evidence-aware deep learning. in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (Brussels, Belgium, Oct.-Nov. 2018), E. Riloff, D. Chiang, J. Hockenmaier, and J. Tsujii, eds. , Association for Computational Linguistics, pp. 22–32.
- [44] SAFAYA, A., ABDULLATIF, M., AND YURET, D. KUISAIL at SemEval-2020 task 12: BERT-CNN for offensive speech identification in social media. in *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (Barcelona (online), Dec. 2020), A. Herbelot, X. Zhu, A. Palmer, N. Schneider, J. May, and E. Shutova, eds. , International Committee for Computational Linguistics, pp. 2054–2059.
- [45] SHIWAKOTI, S., THAPA, S., RAUNIYAR, K., SHAH, A., BHANDARI, A., AND NASEEM, U. Analyzing the dynamics of climate change discourse on twitter: A new annotated corpus and multi-aspect classification. *Preprint* (2024).
- [46] SIDDIQUA, U. A., CHY, A. N., AND AONO, M. Tweet stance detection using an attention based neural ensemble model. in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (Minneapolis, Minnesota, June 2019), J. Burstein, C. Doran, and T. Solorio, eds. , Association for Computational Linguistics, pp. 1868–1873.
- [47] SIMAKI, V., PARADIS, C., AND KERREN, A. Stance classification in texts from blogs on the 2016 british referendum. in *Speech and Computer* (Cham, 2017), A. Karpov, R. Potapova, and I. Mporas, eds. , Springer International Publishing, pp. 700–709.
- [48] SOBHANI, P., INKPEN, D., AND ZHU, X. A dataset for multi-target stance detection. in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational*

*Linguistics: Volume 2, Short Papers* (Valencia, Spain, Apr. 2017), M. Lapata, P. Blunsom, and A. Koller, eds. , Association for Computational Linguistics, pp. 551–557.

- [49] SOBHANI, P., MOHAMMAD, S. M., AND KIRITCHENKO, S. Detecting stance in tweets and analyzing its interaction with sentiment. in *Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics* (Berlin, Germany, 2016), Association for Computational Linguistics, pp. 159–169.
- [50] SUN, Q., WANG, Z., ZHU, Q., AND ZHOU, G. Stance detection with hierarchical attention network. in *Proceedings of the 27th International Conference on Computational Linguistics* (Santa Fe, New Mexico, USA, Aug. 2018), E. M. Bender, L. Derczynski, and P. Isabelle, eds. , Association for Computational Linguistics, pp. 2399–2409.
- [51] TAULÉ, M., MARTÍ, M. A., RANGEL, F. M., ROSSO, P., BOSCO, C., PATTI, V., ET AL. Overview of the task on stance and gender detection in tweets on catalan independence at ibereval 2017. in *CEUR Workshop Proceedings* (2017), volume 1881, CEUR-WS, pp. 157–177.
- [52] THAPA, S., RAUNIYAR, K., JAFRI, F. A., SHIWAKOTI, S., VEERAMANI, H., JAIN, R., KOHLI, G. S., HÜRRİYETOĞLU, A., AND NASEEM, U. Stance and hate event detection in tweets related to climate activism - shared task at case 2024. in *Proceedings of the 7th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)* (2024).
- [53] TOUVRON, H., LAVRIL, T., IZACARD, G., MARTINET, X., LACHAUX, M.-A., LACROIX, T., ROZ-  
IÈRE, B., GOYAL, N., HAMBRO, E., AZHAR, F., RODRIGUEZ, A., JOULIN, A., GRAVE, E., AND  
LAMPLE, G. Llama: Open and efficient foundation language models, 2023.
- [54] TSAKALIDIS, A., ALETRAS, N., CRISTEA, A. I., AND LIAKATA, M. Nowcasting the stance of social media users in a sudden vote: The case of the greek referendum. in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (New York, NY, USA, 2018), CIKM 18, Association for Computing Machinery, p. 367–376.
- [55] TUTEK, M., SEKULIĆ, I., GOMBAR, P., PALJAK, I., ČULINOVIĆ, F., BOLTUŽIĆ, F., KARAN, M.,  
ALAGIĆ, D., AND ŠNAJDER, J. TakeLab at SemEval-2016 task 6: Stance classification in tweets using a genetic algorithm based ensemble. in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (San Diego, California, June 2016), S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, and T. Zesch, eds. , Association for Computational Linguistics, pp. 464–468.

- [56] VAMVAS, J., AND SENNRICH, R. X-Stance: A multilingual multi-target dataset for stance detection. in *Proceedings of the 5th Swiss Text Analytics Conference (SwissText) & 16th Conference on Natural Language Processing (KONVENS)* (Zurich, Switzerland, jun 2020).
- [57] VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. U., AND POLOSUKHIN, I. Attention is all you need. in *Advances in Neural Information Processing Systems* (2017), I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds. , volume 30, Curran Associates, Inc.
- [58] VLADIMIR VOROBEOV, M. K. Chatgpt paraphrases dataset. "<https://huggingface.co/datasets/humarin/chatgpt-paraphrases>", 2023. Online; accessed 10-February-2024.
- [59] VLADIMIR VOROBEOV, M. K. A paraphrasing model based on chatgpt paraphrases. "[https://huggingface.co/humarin/chatgpt\\_paraphraser\\_on\\_T5\\_base](https://huggingface.co/humarin/chatgpt_paraphraser_on_T5_base)", 2023. Online; accessed 10-February-2024.
- [60] WEI, P., MAO, W., AND ZENG, D. A target-guided neural memory model for stance detection in twitter. in *2018 International Joint Conference on Neural Networks (IJCNN)* (2018), pp. 1–8.
- [61] WEI, W., ZHANG, X., LIU, X., CHEN, W., AND WANG, T. pkudblab at SemEval-2016 task 6 : A specific convolutional neural network system for effective stance detection. in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (San Diego, California, June 2016), S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, and T. Zesch, eds. , Association for Computational Linguistics, pp. 384–388.
- [62] WOJATZKI, M., AND ZESCH, T. ltl.uni-due at SemEval-2016 task 6: Stance detection in social media using stacked classifiers. in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (San Diego, California, June 2016), S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, and T. Zesch, eds. , Association for Computational Linguistics, pp. 428–433.
- [63] XU, C., PARIS, C., NEPAL, S., AND SPARKS, R. Cross-target stance classification with self-attention networks. in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (Melbourne, Australia, July 2018), I. Gurevych and Y. Miyao, eds. , Association for Computational Linguistics, pp. 778–783.

- [64] XU, R., ZHOU, Y., WU, D., GUI, L., DU, J., AND XUE, Y. Overview of nlpcc shared task 4: Stance detection in chinese microblogs. in *Natural Language Understanding and Intelligent Applications* (Cham, 2016), C.-Y. Lin, N. Xue, D. Zhao, X. Huang, and Y. Feng, eds. , Springer International Publishing, pp. 907–916.
- [65] ZARRELLA, G., AND MARSH, A. MITRE at SemEval-2016 task 6: Transfer learning for stance detection. in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (San Diego, California, June 2016), S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, and T. Zesch, eds. , Association for Computational Linguistics, pp. 458–463.
- [66] ZHANG, B., DING, D., AND JING, L. How would stance detection techniques evolve after the launch of chatgpt? *arXiv preprint arXiv:2212.14548* (2022).
- [67] ZHANG, B., YANG, M., LI, X., YE, Y., XU, X., AND DAI, K. Enhancing cross-target stance detection with transferable semantic-emotion knowledge. in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Online, July 2020), D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, eds. , Association for Computational Linguistics, pp. 3188–3197.
- [68] ZHAO, G., AND YANG, P. Pretrained embeddings for stance detection with hierarchical capsule network on social media. *ACM Transactions on Information Systems (TOIS)* 39, 1 (2020), 1–32.
- [69] ZHOU, Y., CRISTEA, A. I., AND SHI, L. Connecting targets to tweets: Semantic attention-based model for target-specific stance detection. in *Web Information Systems Engineering – WISE 2017* (Cham, 2017), A. Bouguettaya, Y. Gao, A. Klimenko, L. Chen, X. Zhang, F. Dzerzhinskiy, W. Jia, S. V. Klimenko, and Q. Li, eds. , Springer International Publishing, pp. 18–32.

# واژه‌نامه فارسی به انگلیسی

Pre-train	آموزش اولیه
Train	آموزشی
Hyperparameter	ابریارامتر
Adaptive Selection	انتخاب تطبیقی
Embedding	بازنمایی کلمات
Paraphrase	بازنویسی
Information Retrieval	بازیابی اطلاعات
None/Neither/Neutral	بدون نظر
Context vector	بردار زمینه
Query vector	بردار پرس‌وجو
Optimizer	بهینه‌ساز
Prompt	پرامپت
Attention	توجه
Token	توکن
Loss Function	تابع ضرر
Humor Detection	تشخیص طنز
Stance Detection	تشخیص موضع
Hate Speech Detection	تشخیص سخنان نفرت انگیز
Aspect-Oriented Sentiment Analysis	تحلیل احساسات جنبه گرا
Self-attention	توجه به خود
Early Stopping	توقف زود هنگام
Grid Search	جستجو شبکه‌ای
URL	حذف آدرس اینترنتی
Imbalance data	داده‌های نامتوازن
Temperature	دما
Epoch	دوره



Classifier	رده‌بند
Social Media	شبکه‌های اجتماعی
Recurrent Neural Network (RNN)	شبکه عصبی بازگشتی
Fully Connected Neural Networks	شبکه عصبی تماماً متصل
Convolutional neural network	شبکه عصبی پیچشی
Systematic	قاعده‌مند
Legalization of Abortion	قانونی شدن سقط جنین
encoder	کدگذار
Decoder	کدگشا
Encoder-Decoder	کدگذار-کدگشا
Stop Words	کلمه‌های توقف
Lower Casing	کوچک کردن حروف
Classification Task	مسائل رده‌بندی
Development Set	مجموعه توسعه
Vanishing Gradient	محو شدگی گرادیان
Against/Oppose	مخالف
Large Language Models (LLMs)	مدل‌های زبانی بزرگ
Significance	معنادار
Support/Favor	موافق
Prompt Engineering	مهندسی پرامپت
Stance	موضع‌گیری
Macro Average	میانگین کلان
Username	نام کاربری
Curse of Dimensionality	نفرین ابعاد
Climate Change is Concern	نگرانی از تغییرات آب و هوایی
Feature	ویژگی
Target	هدف
In-context learning	یادگیری درون متنی
Adversarial Learning	یادگیری تقابلی

# واژه‌نامه انگلیسی به فارسی

Adaptive Selection	انتخاب تطبیقی
Adversarial Learning	یادگیری تقابلی
Against	مخالف
Aspect-Oriented Sentiment Analysis	تحلیل احساسات جنبه گرا
Attention	توجه
Classification Task	مسائل رده‌بندی
Classifier	رده‌بند
Climate Change is Concern	نگرانی از تغییرات آب و هوایی
Context vector	بردار زمینه
Convolutional neural network	شبکه عصبی پیچشی
Curse of Dimensionality	نفرین ابعاد
Decoder	کدگشا
Development Set	مجموعه توسعه
Early Stopping	توقف زود هنگام
Embedding	بازنمایی کلمات
Encoder	کدگذار
Encoder-Decoder	کدگذار-کدگشا
Epoch	دوره
Favor	موافق
Feature	ویژگی
Fully Connected Neural Networks	شبکه عصبی تماماً متصل
Grid Search	جستجو شبکه‌ای
Hate Speech Detection	تشخیص سخنان نفرت انگیز
Humor Detection	تشخیص طنز
Hyperparameter	ابریارامتر
Imbalance data	داده‌های نامتوازن

In-context learning	یادگیری درون متنی
Information Retrieval	بازیابی اطلاعات
Large Language Models (LLMs)	مدل‌های زبانی بزرگ
Legalization of Abortion	قانونی شدن سقط جنین
Loss Function	تابع ضرر
Lower Casing	کوچک کردن حروف
Macro Average	میانگین کلان
None/Neither/Neutral	بدون نظر
Optimizer	بهینه‌ساز
Oppose	مخالف
Paraphrase	بازنویسی
Pre-train	آموزش اولیه
Prompt	پرامپت
Prompt Engineering	مهندسی پرامپت
Query vector	بردار پرس‌وجو
Recurrent Neural Network (RNN)	شبکه عصبی بازگشتی
Self-attention	توجه به خود
Significance	معنادار
Social Media	شبکه‌های اجتماعی
Stance	موضع‌گیری
Stance Detection	تشخیص موضع
Stop Words	کلمه‌های توقف
Support	موافق
Systematic	قاعده‌مند
Target	هدف
Temperature	دما
Train	آموزشی
Token	توکن
URL	حذف آدرس اینترنتی
Username	نام کاربری
Vanishing Gradient	محو شدگی گرادیان

**Abstract:**

Nowadays, social media is a platform for freely expressing and sharing opinions and thoughts. This leads to the fact that by analyzing the data available on social media, a broad and comprehensive perspective on various users' opinions and sides about different topics could be gained. These topics include political, economic, social, and cultural issues. In Natural Language Processing, stance detection is the process of automatically recognizing the side and stance of a given text about a specific target.

In natural language processing tasks, the way text data is preprocessed significantly affects the performance of the trained model. In this research, seven different levels of preprocessing are introduced and examined. Additionally, to find the architecture of the stance detection model, the idea of Neural Architecture Search (NAS) was inspired. In this method, the model architecture is divided into four main parts, a search space is defined for each part, and adaptive search algorithms are used to design the final architecture. The best proposed model ultimately utilizes BERTweet as the encoder and a CNN classifier. The proposed architecture achieved an F1-Score of 74.47%, showing a 19.97% improvement over the Baseline model. Furthermore, the proposed method ranked third among 19 participants in a climate change stance detection event. Additionally, due to the lack of training data for different topics, stance detection without training data was also investigated. This approach, which uses large language models and prompt engineering, introduces four approaches based on different prompt types. Then, the performance of the proposed prompts was compared with other methods for stance detection without training data. The introduced approach achieved an F1-Score of 57.33%, showing a 2.03% improvement over similar approaches.

**Keywords:** Stance Detection for Textual Content in Social Media



**Iran University of Science and Technology  
Computer Engineering Department**

# **Stance Detection for Textual Content in Social Media**

**A Thesis Submitted in Partial Fulfillment of the Requirement for the Degree  
of Master of Science in Computer Engineering**

**By:**

**Ghazaleh Mahmoudi**

**Supervisor:**

**Dr. Sayyed Sauleh Eetemadi**

**February 2024**