



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر



استفاده از روش های یادگیری ماشین برای تشخیص موضع کاربران در مورد شایعه در
شبکه های اجتماعی

پایان نامه برای دریافت درجه کارشناسی
در رشته مهندسی برق تمرکز الکترونیک

نام
سارینا جامی الاحمدی
شماره دانشجویی
۸۱۰۱۹۵۳۷۲

استاد راهنما:
دکتر سید پویا شریعت پناهی

شهریورماه ۱۴۰۱

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تعهدنامه اصالت اثر

باسمه تعالی

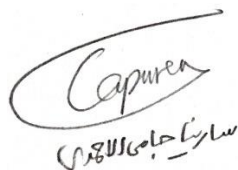
اینجانب سارینا جامی الاحمدی تأیید می‌کنم که مطالب مندرج در این پایان‌نامه حاصل تلاش اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشته از آنها استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارائه نشده است.

کلیه حقوق مادی و معنوی این اثر متعلق به دانشکده فنی دانشگاه تهران می‌باشد.

نام و نام خانوادگی دانشجو :

سارینا جامی الاحمدی

امضای دانشجو :



سارینا جامی الاحمدی

چکیده^۱

با پیشرفت تکنولوژی و دسترسی سریع و راحت عموم مردم به اینترنت، شبکه های اجتماعی^۲ بستر اصلی دریافت اطلاعات و اخبار در سراسر دنیا شده است. اگر چه این دسترسی سریع و آسان به اطلاعات از هر مرز و بومی باعث بالاتر رفتن سطح زندگی ها شده است و مزایای بسیاری دارد ولی از عیوب آن نمیتوان چشم پوشی کرد. همان طور که هر شخصی میتواند از طریق شبکه های اجتماعی به اخبار دسترسی پیدا کند، میتواند خود نیز اطلاعاتی را در اینترنت منتشر کند. این موضوع باعث شده افرادی به اشتباه یا با قصد سوءاستفاده اطلاعات نادرست را در شبکه های اجتماعی به اشتراک بگذارند و باعث آسیب رساندن به جامعه یا افراد شوند. این اطلاعات نادرست میتواند باعث ایجاد ترس و گمراهی در میان افراد جامعه شود و منجر به تاثیرگذاری مستقیم روی نتایج انتخابات داشته باشد یا زمینه ایجاد شورش را فراهم کند. به این دلیل نیاز است تا از پخش شدن اخبار جعلی^۳ و شایعه^۴ جلوگیری شود و هر چه زودتر از فراگیری این اخبار جلوگیری شود تا احتمال آسیب رساندن آن کمتر است. در این مقاله قصد داریم روش ها و ابزارهای موجود برای صحت سنجی^۵ شایعه در شبکه های اجتماعی که موفق عمل کرده اند را مرور کنیم و سپس یک روش یادگیری ماشین که جزو وضعیت روز^۶ مدل ها است را یاددهی و پیاده سازی کنیم. در نهایت تلاش میکنیم تا بهبودهایی روی مدل انجام دهیم و عملکرد آن را با نتایج مدل های مطرح شده در مقالات دیگر مقایسه کنیم.

کلمات کلیدی: شایعه، موضع کاربر، شبکه های اجتماعی، یادگیری ماشین، یادگیری عمیق

¹ Abstract

² Social media

³ Fake news

⁴ Rumor

⁵ Verification

⁶ State-of-the-art

فهرست مطالب

فصل ۱: مقدمه و بیان مساله ۱

۱-۱- مقدمه ۲

۱-۲- تاریخچه ۳

۱-۳- شرح مسئله تشخیص موضع کاربر در مورد شایعه ۵

۱-۴- تعریف مسئله ۶

۱-۵- اهداف و آرمان‌های کلی ۷

۱-۶- روش انجام تحقیق ۸

۱-۷- ساختار پایان‌نامه ۸

فصل ۲: ابزارها و مفاهیم لازم جهت تشخیص موضع کاربر در رابطه با شایعه ۹

۲-۱- پیش زمینه ۱۰

۲-۱-۱- یادگیری ماشین با روش های مرسوم ۱۲

۲-۱-۱-۱- یادگیری عمیق ۱۴

۲-۱-۱-۲- تشخیص زود هنگام ۱۶

۲-۱-۱-۳- یادگیری چند وظیفه ای ۱۷

۲-۱-۲- مجموعه ویژگی ها ۱۸

۲-۱-۳- مجموعه داده ها ۱۹

۲-۲- خلاصه و جمع بندی ۲۰

فصل ۳: مدل سازی و روش طراحی ۲۱

۳-۱- مقدمه ۲۲

۳-۲- روش های مرسوم در پردازش زبان طبیعی ۲۲

۲۲	۳-۲-۱- الگوریتم‌های یادگیری ماشین
۲۲	۳-۲-۱-۱- دسته‌بندی بیزین احمقانه
۲۳	۳-۲-۱-۲- نزدیک‌ترین همسایه
۲۳	۳-۲-۱-۳- ماشین بردار پشتیبان
۲۴	۳-۲-۱-۴- درخت تصمیم
۲۴	۳-۲-۲- الگوریتم‌های یادگیری عمیق
۲۴	۳-۲-۲-۱- شبکه‌های عصبی درهم‌پیچیده
۲۵	۳-۲-۲-۲- شبکه‌های عصبی برگشتی
۲۵	۳-۲-۲-۳- شبکه‌های عصبی مصنوعی
۲۶	۳-۲-۳- الگوریتم‌های مدرن
۲۶	۳-۲-۳-۱- LSTM
۲۷	۳-۲-۳-۲- GRU
۲۸	۳-۲-۳-۳- GPT
۲۹	۳-۲-۳-۴- BERT
۳۰	۳-۳- گزینش مدل
۳۱	۳-۴- تحلیل نتایج و ارائه مدل پیشنهادی
۳۱	۳-۵- خلاصه و جمع‌بندی
۳۲	فصل ۴: پیاده‌سازی مدل پیشنهادی
۳۳	۴-۱- مقدمه
۳۳	۴-۲- مدل پیشنهادی
۳۴	۴-۲-۱- معرفی مدل
۳۵	۴-۲-۲- مجموعه داده‌ها

۳۶	۴-۲-۳- نحوه استفاده از داده‌ها در مدل
۳۸	۴-۲-۴- معماری‌های استفاده شده
۳۸	۴-۲-۵- مراحل و روش پیاده سازی
۴۱	۴-۲-۶- نتایج و ارزیابی
۴۲	۴-۳- چالش‌ها و فرصت‌ها
۴۴	۴-۴- خلاصه و جمع‌بندی
۴۵	فصل ۵: جمع‌بندی، نتیجه‌گیری و پیشنهادها
۴۶	۵-۱- جمع‌بندی
۴۶	۵-۲- نتیجه‌گیری
۴۶	۵-۲-۱- دستاوردها
۴۶	۵-۲-۲- محدودیت‌ها
۴۷	۵-۲-۳- پیشنهادها
۴۸	فصل ۶: مراجع

فهرست علائم اختصاری

ML	Machine Learning
DL	Deep Learning
BERT	Bidirectional Encoder Representations from Transformers
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
AI	Artificial Intelligence
SVM	Support Vector Machines
NLP	Natural Language Processing
PTM	Pre-Trained Models
ANN	Artificial Neural Network
RL	Reinforcement Learning
DNN	Deep Neural Network
RNN	Recurrent Neural Network
GRU	Gated Recurrent Unit

فصل 1

فصل ۱:

مقدمه و بیان مسأله

در این فصل ابتدا وارد مقدمات شایعه و شبکه‌های اجتماعی و تاریخچه آن‌ها می‌شویم. سپس مسأله‌ی موجود در این زمینه را شرح می‌دهیم و اهداف کلی از انجام آن را بیان می‌کنیم. در نهایت در مورد روش انجام تحقیق صحبت کرده و سپس به ساختار پایان‌نامه اشاره خواهیم کرد.

۱-۱- مقدمه

با گسترش تعداد کاربران شبکه های اجتماعی برخط، این شبکه ها تبدیل به رسانه ی تاثیرگذاری به منظور به اشتراک گذاشتن اطلاعات شدند و افراد برای تصمیم گیری هایشان در مورد یک موضوع خاص مانند مباحث سیاسی، تحت تاثیر دوستان خود هستند. علاوه برآن نیز شبکه های اجتماعی از جمله منابع مهم خبر برای بیشتر مردم محسوب میشوند. از این رو اطلاعات نادرست موجود در این شبکه ها، باعث تاثیر غلط در جمعیت زیادی از مردم و انتقال آلودگی های مختلف بین دوستان خواهد شد. انتشار شایعه و اخبار نادرست از زمان های گذشته به عنوان یک مسئله ی اجتماعی و روانی با ابعاد گسترده بوده است که در مواقع بحرانی مانند جنگ بیشتر حائز اهمیت میشود. در هر زمان افراد و دولت ها ممکن است با اهداف مختلفی تصمیم به انتشار اخبار نادرستی بگیرند که از جمله این اهداف می توان به ایجاد بدبینی نسبت به نظام و مسئولین جامعه، افزایش نگرانی و اضطراب در بین مردم، ایجاد تقابل و صف بندی میان قشرهای مختلف و پاسخ گویی به حس انتقام، اشاره کرد. در گذشته شایعات بیشتر از طریق گفت و گو، روزنامه ها، رادیو و تلویزیون منتشر میشد اما امروزه به دلیل قابلیت های موجود در شبکه های اجتماعی، این شبکه ها نیز به عنوان ابزاری برای انتشار شایعه تبدیل شده اند. شایعه به دلیل گیرایی و کششی که دارد با سرعتی باور نکردنی در جامعه حرکت میکند و گسترش آن میتواند تاثیرات مخرب زیادی را بر جوامع انسانی بگذارد و گاهی عواقب جبران ناپذیری را نیز به بار آورد. با توجه به تاثیری که گسترش شایعات در جوامع از طریق شبکه های اجتماعی دارند، این موضوع و یافتن راهی به منظور مقابله با آن زمینه ی پژوهش های جدیدی در این حوزه شده است. با توجه به مسائل و مشکلات مطرح شده هنگام انتشار شایعات، پژوهش های زیادی در این زمینه انجام شده است. با این حال، حتی اگر این روش ها با سرعت خوبی رشد کنند، این مشکل پیچیدگی های زیادی دارد که نیازمند بررسی عمیق تر روی موضوع است [2]. هدف اصلی کار پیش رو، بررسی مدل های مختلف شبکه های عصبی که بر روی تشخیص و دسته بندی شایعه کار شده اند و همچنین مجموعه داده های¹ تهیه و استفاده شده در این

¹ Dataset

زمینه بوده و سپس با انتخاب یکی از روش‌های بروز آن در حوزه‌ی یادگیری عمیق^۱، می‌خواهیم آن را پیاده‌سازی کنیم و با بهبودهایی روی آن، بتوانیم نتایج خود را با مقاله‌های دیگر مقایسه کنیم.

۲-۱- تاریخچه

امروزه رسانه‌های اجتماعی نقش بیشترین تکنولوژی مورد استفاده برای بدست آوردن اطلاعات، به اشتراک گذاشتن نظرات و انتشار خبر را دارند. استفاده وسیع از رسانه‌های اجتماعی مرزها و محدودیت‌های فرهنگی و جغرافیایی را جهت بدست آوردن اطلاعات از بین برده است. از این رو روزنامه نگار ها^۲ از شبکه‌های اجتماعی برای پیدا کردن اخبار نوظهور استفاده میکنند. همچنین، شبکه‌های اجتماعی برخط^۳ به طور گسترده‌ای باعث چگونگی بازخورد کاربران به ساخت و توزیع پیام‌های رسانه‌ای شده است [1]. با توجه به مقاله [2] رسانه‌های اجتماعی باعث کمرنگ شدن نقش متخصصان و نویسندگان معتبر شده‌اند به گونه‌ای که مردم معمولاً بیشتر خبر را از دوستان و همکاران خود دریافت میکنند به جای سازمان‌ها و ژورنال‌ها. بنابراین، بستر رسانه‌های اجتماعی فرصت اشتراک‌گذاری و تفسیر اخبار را بطور آزادانه برای هر شخصی فراهم کرده است. در نتیجه این عمل باعث افزایش حجم زیادی از اخبار تایید نشده و نامعتبر شده است که ممکن است باعث سردرگمی مردم شود و شک را در بین آن‌ها افزایش دهد. اگر چه این بستر برخط باعث رفاه بیشتری برای جامعه شده است ولی معایب آن از چشم پوشیده نیست. اخبار صحت سنجی نشده^۴ که در شبکه‌های اجتماعی معمولاً به عنوان شایعه^۵ شناخته میشود ممکن است در زمانی دیرتر به عنوان اطلاعات اشتباه معین گردد. شایعات اشتباه و فراگیر تا زمانی که هنوز صحت آن‌ها مشخص نشده است ممکن است تاثیر نامطلوبی بر افراد یا اجتماع داشته باشد [3]. این نوع از اطلاعات باعث گمراهی مردم میشود و حتی ممکن است بانی یک سری فعالیت‌های آسیب زنده در سطح جامعه شود. به عنوان مثال، در

¹ Deep Learning

² Journalist

³ Online

⁴ Unverified

⁵ Rumor

سال ۲۰۱۶ شایعه ای در مورد رخ دادن یک واقعه پدوفیلی در یک رستوران به نام پیزا^۱ واقع در واشینگتن به طور عمده ای پخش شده بود که باعث شد شخصی با تفنگی در دست به آن رستوران رود و به عنوان یک تهدید محسوب شود. اون نه تنها هیچ گواهی بر این موضوع نیافت بلکه تهدیدی که او به عمل آورد باعث دستگیر شدن این شخص شد [4]. مثالی دیگر از تاثیر منفی اطلاعات غلط سقوط سریع میانگین صنعتی داوجونز^۲ در سال ۲۰۱۳ بعد از توییت کردن شایعه نادرست حمله به کاخ سفید توسط حساب هک شده آن ها شد [5].

به دلیل افزایش روزمره محتوای برخط، صحت سنجی شایعات بصورت دستی کاری بسیار وقت گیر و سنگین است. این مسئله تعداد بسیاری از پژوهشگران را برای توسعه سیستم های اتوماتیک در راستای تشخیص شایعه در رسانه های اجتماعی و صحت سنجی آن قبل از شیوع وسیع آن ها به این سمت کشانده است. مطالعه ماهیت شایعه یکی از موضوعات کانونی در زمانی طولانی بین پژوهشگران با زمینه های مختلف از مطالعات روان شناختی تا بررسی های محاسباتی بوده است [6]. این مطالعات حقایق جالبی را در مورد پخش شایعه و مشارکت کاربران در آن آشکار کرده اند. به عنوان مثال، نقش یک سیستم صحت سنجی شایعه در هشدار دادن به کاربران از اطلاعات نامعتبر که در نتیجه باعث کمک به کاربران و بیشتر کردن توجه آن ها به اخبار نادرست میشود بسیار حیاتی است. همچنین، کاربران تمایل دارند درباره صحت اطلاعات به بحث و گفتگو در شبکه های اجتماعی بپردازند. استفاده از نظرات کاربران در تشخیص شایعه میتواند کمک شایانی کند. به عنوان مثال، اگر اکثر کاربران به نادرست بودن متن شایعه اشاره کنند میتوان این برداشت را کرد که این شایعه نادرست است. این مسئله با عنوان موضع^۳ کاربر شناخته میشود و تمرکز ما در این مقاله، شرح بیشتر و دقیق تر موضع کاربران در مورد شایعه در شبکه های اجتماعی، داده های موجود، تحلیل های مرتبط با این موضوع و روش های مطرح شده برای دستیابی به عملکرد بهتر در این زمینه است.

¹ Pizza restaurant

² Dow Jones Industrial Average

³ Stance

۳-۱- شرح مسئله تشخیص موضع کاربر در مورد شایعه^۱

با توجه به مقاله [6] شایعه به عنوان اطلاعاتی در حال توزیع شدن^۲ در نظر گرفته میشود که صحت آن مشخص نیست. به عبارتی دیگر، شایعات اطلاعاتی هستند که در بستر شبکه های اجتماعی پخش میشوند و به اشتراک گذاشته میشوند که اعتبار آن ها قطعی نیست. همچنان که در طول زمان شواهدی مرتبط به شایعه جمع میشود، صحت شایعه میتواند مشخص گردد. گرچه در بعضی موارد، شواهد موجود برای تعیین صحت شایعه کافی نیست و این شایعه به عنوان تایید نشده یا صحت سنجی نشده باقی میماند [7]. در نتیجه، پست های شامل شایعه معمولاً برچسب صحیح^۳، غلط^۴ و تایید نشده میخورند [7].

کاربران شبکه های اجتماعی تمایل دارند تا در مورد صحت یک شایعه گفتگو کنند و دیدگاه خود را بیان کنند. به عنوان مثال، آن ها ممکن است در مورد اعتبار شایعه سوال بپرسند و درخواست منبع معتبر کنند؛ ممکن است شایعه را تایید و یا رد کنند و منبعی را به عنوان گواهی بر حرفشان بیاورند؛ همچنین میتوانند شایعه را رد یا تایید نکنند و بدون جهت گیری در مورد آن نظر دهند [8]. توضیحاتی که بیان شد دسته بندی دیدگاه یا موضع کاربران را مشخص میکند که شامل تایید کردن^۵، رد کردن^۶، سوال پرسیدن^۷ و نظر دادن^۸ است [6].

¹ Rumor stance classification

² Diffusing

³ True

⁴ False

⁵ Support

⁶ Deny

⁷ Quert

⁸ Comment

در مقاله [6]، یک سیستم تشخیص شایعه^۱ به عنوان یک مجموعه ای که شامل چهار عنصر است تعریف میشود: شناسایی شایعه^۲، ردیابی شایعه^۳، دسته بندی موضع کاربر^۴ و صحت سنجی شایعه^۵. دو عنصر اول مازول های لازم جهت جمع آوری داده های شامل پست شایعه و نظرات کاربران هستند. به این صورت که عنصر شناسایی شایعه مشخص میکند که یک پست دارای متن شایعه هست یا خیر. عنصر ردیابی شایعه پست های شامل نظرات کاربران که مربوط به شایعه پیدا شده هست را مشخص میکند. عنصر سوم این پست های کاربران را طبق اینکه چه دیدگاهی نسبت به شایعه دارند دسته بندی میکند. در آخر، عنصر چهارم پست منبع که همان شایعه است را صحت سنجی میکند. در این عنصر، میتوان از نتایج عنصر سوم استفاده کرد و روند صحت سنجی یک شایعه را بهبود بخشید.

۴-۱- تعریف مسئله

در این مقاله میخواهیم روشی برای تشخیص موضع کاربران در رابطه با یک شایعه در شبکه های اجتماعی پیاده سازی کنیم. همانطور که بیان شد موضع کاربران میتواند کمک شایانی به صحت سنجی شایعه بکند و با بهبود روش های دسته بندی موضع کاربران میتوانیم دقت تشخیص شایعه را بالا ببریم. مجموعه داده^۶ در این مسئله به این صورت است که پست منبع و تمام اطلاعات مربوط به پروفایل کاربر و پست های کاربران دیگر در جواب به پست منبع به همراه اطلاعات این کاربران در اختیارمان قرار دارد. این مجموعه داده برچسب گذاری شده و به دو دسته یاددهی^۷ و تست^۸ تقسیم بندی شده است. هدف ما

¹ Rumor detection framework

² Rumor detection

³ Rumor tracking

⁴ Stance classification

⁵ Veracity classification

⁶ Dataset

⁷ Training

⁸ Testing

استفاده از مدل از پیش یاددهی شده^۱ BERT به عنوان یکی از مدل های یادگیری عمیق^۲، که عملکرد بسیار خوبی در مدل کردن و دسته بندی مسائل مختلف نشان داده است که در ادامه توضیح خواهیم داد، است و مقایسه خروجی آن با دیگر مدل های به روز است.

۵-۱- اهداف و آرمان های کلی

با توجه به اینکه مسئله ی تشخیص موضع کاربر و راه حل های ارائه شده در این زمینه رو به رشد است، نیاز هست تا بتوانیم یک دسته بندی از آن ها داشته باشیم و پیشرفت آن ها در اثر گذر زمان را دنبال کنیم. هدف از انجام این تحقیق، مرور و مقایسه تعدادی از مقاله هایی که در راستای دسته بندی موضع کاربران نسبت به شایعه در شبکه های اجتماعی و سپس انتخاب یک روش مناسب و به روز برای پیاده سازی و خروجی گرفتن است.

آرمان این تحقیق این است که مطالعه جامعی درباره دسته بندی موضع کاربر در مورد شایعه ارائه کنیم که شامل روش های انجام شده تا به امروز، ویژگی هایی که اطلاعات بیشتری در مورد صحت شایعه میدهند، مجموعه داده هایی که تا بحال ارائه شده و مشکلات مربوط به آن ها، یافته های آزمایش های انجام شده و همچنین چالش های این مسئله است. با توجه به اینکه مطالعه جامعی در این زمینه صورت نگرفته است، مقاله ما میتواند کمکی در راستای این موضوع باشد. همچنین، با توجه به روش های انجام شده و نتایجی که از مطالعاتمان بدست آوردیم به پیاده سازی یک مدل یادگیری ماشین میپردازیم تا بتوانیم به بررسی بیشتر مسئله موضع کاربر و مشکلات طراحی برای آن بپردازیم. همچنین امیدواریم توسعه ی این ابزارها به حدی برسد که از آن ها بشود برای پیشگیری از تولید و شیوع شایعه های نادرست در رسانه ها و بسترهای شبکه های اجتماعی استفاده شود، چرا که فقط در این حالت هست که می توان از این خطر بزرگ، یعنی کنترل افکار عمومی، جلوگیری کرد و صرف تشخیص آن فقط قدم اول در راستای حل این مشکل است.

^۱ Pre-trained

^۲ Deep learning

۶-۱- روش انجام تحقیق

روش انجام تحقیق به این صورت بوده است که ابتدا با مطالعه مقالات مرور ساخته یافته^۱ فهم جامعی درباره اطلاعات غلط^۲ در رسانه های اجتماعی بدست آوردیم. سپس، به بررسی دقیق تر موضوعات مختلف آن پرداختیم و تصمیم گرفتیم مسئله شایعه را برای ادامه تحقیق انتخاب کنیم. با مطالعه بیشتر و دقیق تر درباره شایعه به مبحث مهمی که بسیار مورد توجه پژوهشگران است برخوردیم. این موضوع مطالعه موضع کاربران در مورد شایعه است که با چالش های بسیاری روبرو است و به عنوان یک روش مفید برای کمک به صحت سنجی شایعه تلقی میشود. با توجه به نبود مقاله جامعی در زمینه موضع کاربر تصمیم گرفتیم به مطالعه بیشتر این مهم بپردازیم و مقاله جامعی در این راستا منتشر کنیم که کمکی به جامعه پژوهشگران موضع کاربر در مورد شایعه در شبکه های اجتماعی شود. سپس، با توجه به روش های مطالعه شده مدلی پیشرفته و به روز را با استفاده از یک دیتاست شناخته شده پیاده سازی کردیم تا بتوانیم پژوهشمان را بصورت تکنیکال در این زمینه ادامه دهیم.

۷-۱- ساختار پایان نامه

در فصل دوم، به مطالعه دقیق کار های انجام شده و مدل های پیشنهاد شده در این حوزه میپردازیم. همچنین، مجموعه داده های موجود، ویژگی های مفید استخراج شده، نتایج تحلیل های انجام شده و چالش های این زمینه را معرفی میکنیم. فصل سوم دربرگیرنده توضیحات مربوط به انواع الگوریتم های مرسوم مورد استفاده در حوزه پردازش زبان طبیعی بوده و سپس بر اساس معیارهای ارائه شده، این مدل ها مقایسه و ارزیابی می شوند. در فصل چهارم مدل پیشنهادی خود را تشریح کرده و سپس به بیان مراحل و روش های پیاده سازی می پردازیم. سپس این فصل را با چالش ها و ناکامی هایی که در پیاده سازی پروژه به آن برخوردیم به پایان می رسانیم. در نهایت، در فصل پنجم، نتیجه گیری های کلی حاصل شده در این تحقیق را بیان کرده و نوآوری های موجود در آن را تبیین می کنیم. سپس محدودیت ها مورد بحث قرار گرفته و در نهایت پیشنهاداتی برای ادامه کار خواهیم داد.

¹ Systematic review

² Misinformation

فصل 2

فصل ۲: ابزارها و مفاهیم لازم جهت تشخیص موضع کاربر در رابطه با شایعه

در فصل پیش رو، ابتدا پیش زمینه تشخیص شایعه و موضع کاربر در شبکه های اجتماعی را بیان میکنیم. سپس به مطالعه روش های مطرح شده در این حوزه میپردازیم. در ادامه، ویژگی های استخراج شده و مجموعه داده ها و یافته های این زمینه را مطالعه میکنیم.

۱-۲- پیش زمینه^۱

مطالعه شایعات در گستره بزرگی از رشته‌ها انجام شده است مانند همانندی ریاضی پخش شایعه به شیوع عفونت [9]، بررسی جنبه روانی شایعه در یک انجمن و دلایل منشاء آن [10] و به کارگیری روش‌های محاسباتی برای تشخیص شایعه در اینترنت. تعداد بسیار زیادی کار پژوهشی متمرکز بر پخش اطلاعات نادرست در فضای برخط در سال‌های اخیر گزارش شده است. این مطالعات مسئله را با روش‌های مختلفی همچون انتخاب مجموعه ویژگی‌های مفید، استفاده از مدل‌های یادگیری با نظارت^۲ و نیمه نظارت شده^۳، بکارگیری مشخصه‌های زمانی، در نظرگیری مدل توزیع شبکه و غیره. نتایج مطالعات نشان داده که الگوهای خاصی در ساختار درختی^۴ گفتگوها مشاهده میشود که میتواند تعیین کننده درستی یا نادرستی یک شایعه باشد [11]. اولین کار بر روی تشخیص خودکار شایعه توسط مقاله [12] انجام شد که شامل جمع‌آوری مجموعه داده از شایعاتی که بطور طولانی مدت^۵ در شبکه‌های اجتماعی وجود دارند و استخراج ویژگی‌های بر پایه محتوا^۶، بر پایه شبکه^۷ و خاص توییتر^۸ مانند هشتگ است. آن‌ها از یک مدل regression استفاده کردند و دو برجسب برای موضع کاربران در نظر گرفتند شامل دسته‌های صحیح و غلط و سپس نتیجه مدل یاددهی شده بر روی توییت‌هایی از همان شایعه‌ها تست کردند. کار آن‌ها توسط دیگر پژوهشگران گسترش یافت و جنبه‌های مختلف آن بررسی شد [13]. با این حال، این روش جمع‌آوری داده نتوانست در تشخیص شایعاتی که در شرایط بحرانی پدید می‌آیند موفق عمل کند که این موضوع باعث محدود کردن دریافت مفهوم شایعه در رسانه‌ها اجتماعی شد. در سال ۲۰۱۶، رساله [7] روشی جدید را برای جمع داده

¹ Background² Supervised³ Semi-supervised⁴ Tree-structured⁵ Long-standing rumors⁶ Content-based⁷ Network-based⁸ Twitter-specific

برای شایعاتی که از قبل نمیشناسیم معرفی کرد. آن‌ها یک مجموعه داده بزرگ جمع‌آوری کردند و تحقیقات عددی بر روی شیوع شایعه و عکس‌العمل کاربران انجام دادند.

در ادامه رویکرد^۱ های متفاوتی برای دسته‌بندی موضع کاربران در مورد شایعه مطالعه میکنیم.

ما این روش‌ها را به چهار دسته تقسیم‌بندی میکنیم شامل:

- یادگیری ماشین با روش‌های مرسوم^۲ که به مدل‌هایی اشاره میکند که از تولید و توسعه ویژگی‌های بیشتر برای بهتر کردن عملکرد سیستم استفاده میکنند.
- یادگیری عمیق^۳ که از شبکه‌های عصبی بهره‌میلد و شامل روش‌هایی همانند CNN^۴ و LSTM^۵ است.
- تشخیص زودهنگام^۶ که تاکید بر اهمیت تشخیص سریع شایعه دارد و مدل‌ها را بر اساس عملکرد آن‌ها در ساعات اولیه پخش شدن شایعه ارزیابی میکند.
- یادگیری چندوظیفه‌ای^۷ که به بررسی یادگیری همزمان دسته‌بندی موضع کاربران و صحت‌سنجی شایعه میپردازد. این در حالی است که رویکرد‌های دیگر ابتدا موضع کاربر را دسته‌بندی میکنند و سپس از خروجی این مدل به عنوان ورودی برای سیستم صحت‌سنجی شایعه استفاده میکنند.

^۱ Approach

^۲ Conventional machine learning

^۳ Deep learning

^۴ Convolutional neural network

^۵ Long short-term memory

^۶ Early detection

^۷ Multi-task learning

۱-۱-۲- یادگیری ماشین با روش های مرسوم

تعداد زیادی پژوهش نشان داده اند که با بکار بردن تعداد بیشتری ویژگی های مفید در مدل های ساده یادگیری میتوانیم به عملکرد بسیار خوبی دست یابیم. در مقاله [14] ویژگی هایی را استخراج میکنند که نشان دهنده زبان توییت^۱، بر پایه گفتگو^۱ و ویژگی های موثر هستند که جنبه احساسی محتوای توییت را مشخص میکنند. آن ها این ویژگی ها را به یک مدل طبقه بندی به نام SVM میدهند و نشان میدهند که ویژگی های بر پایه گفتگو تاثیر به سزایی بر روی عملکرد مدل دارند بدون نیاز به پیچیدگی های روند یادگیری. علاوه بر این، مقاله [15] از مدل درختی^۲ J48 استفاده کردند و آن را برای تعدادی ویژگی های بطور خودکار شناسایی شده مربوط به مسئله همانند نمره محاسبه شده برای قطعیت و عدم قطعیت استفاده کردند. نویسندگان این رساله مشخص کردند که هر کدام از این مشخصه ها در بهتر شدن عملکرد مسئله نقش ایفا میکنند. در رساله دیگری [16] مولف های آن بر روی موضع کاربر در مورد شایعه هایی که به تداخل روانی مربوط هستند پژوهش کردند. آن ها نمونه کوچکی از توییت هایی مربوط به سلامت روان را جمع آوری کردند و همچنین داده های بیشتری از یک مجموعه داده شناخته شده که مربوط به سلامت روان نیستند اضافه کردند. آن ها داده های یاددهی و تست را به دو گروه تقسیم کردند: ۱. داده های تست هرگز در داده های یاددهی مشاهده نشده اند ۲. یک بخشی از داده های تست (۱۰ درصد تا ۶۰ درصد) در داده های یاددهی تلفیق میشوند. مدل های یادگیری SVM, J48, Random Forest, Naïve Bayes, Instance-based classifier استفاده شدند. آن ها ابتدا دو تکنیک بیان شده را بر روی داده های نامرتب به سلامت روان پیاده کردند سپس، داده های مربوط به سلامت روان را به عنوان داده های تست در نظر گرفتند و تمام مدل های طبقه بندی را بر روی داده های نامرتب به سلامت روان یاددهی کردند. در هر دو سناریو، مدل J48 بهترین عملکرد را نشان داد اگر چه وقتی که مورد های موجود در داده تست قبلا توسط مدل مشاهده نشده اند، دقت نهایی سیستم کاهش میابد. این موضوع بیان میکند که شایعات مختلف دارای ویژگی های متفاوتی از یکدیگر هستند که توسط مدل یادگیری ماشینی که آن ها را قبلا مشاهده نکرده است ممکن است قابل تشخیص نباشد.

¹ Conversation

² Decision tree

در رساله [17] و [18] از مدل طبقه بندی فرآیند گاوسی^۱ بر روی دو مجموعه داده استفاده شده است. آن‌ها از دو رویکرد استفاده کرده اند به نام های Leave-one-out و Leave-part-out که دومی همانند روش رساله قبلی است با این تفاوت که حداکثر ۷ درصد از شایعه هدف را در داده های یاددهی استفاده کرده اند. در مجموعه داده اول آن‌ها متوجه میشوند که مدل فرآیند گاوسی میتواند توزیع موضع های متمایز را بهتر نسبت به مدل های پایه^۲ تشخیص دهد. همین نتیجه برای مجموعه داده دوم نیز درست است با این تفاوت که نتیجه اندکی ضعیف تر است. دلیل آن ممکن است بخاطر وجود رویداد های مختلف در مجموعه داده دوم باشد که باعث شود مدل طبقه بندی ویژگی هایی را ببیند که قبلا در مرحله یاددهی با آن روبرو نشده است یا ممکن است به دلیل اینکه مجموعه داده دوم شامل توییت هایی است که در جواب به توییت های قبلی خود هستند و طول متن آن‌ها کوتاه است باعث شده باشد که مدل طبقه بندی قادر به فهم موضع آن توییت ها نشود. رساله [19] از مدل طبقه بندی فرآیند هاوکس^۳ با استفاده از اطلاعات زمانی بهره گرفته اند. مطالعات آن‌ها نشان دهنده تاثیر مثبت استفاده از متغیر های زمانی در دسته بندی موضع کاربران بوده است. رساله [20] موضع کاربران را در یک مدل HMM^۴ استفاده میکند تا تاثیر موضع کاربران را در صحت سنجی شایعه بررسی کند. در آزمایشی دیگر، مولفان از HMM با چندین فضا^۵ استفاده میکنند تا زمان را نیز در نظر بگیرند. آن‌ها نشان میدهند که استفاده از موضع کاربران بصورت جمعی به عنوان یه مشخصه که به ویژگی های استخراج شده در مدل اضافه شوند میتوانند در بهبود عملکرد یک مدل تشخیص شایعه کمک کنند.

رساله [21] ادعا میکند که الگوریتم های نیمه نظارت شده عملکرد بهتری نسبت به الگوریتم های نظارت شده در ابعاد بزرگ داده^۶ دارند. به دلیل محدودیت های الگوریتم های نظارت شده، این روش یک ابزار سودمند برای بدست آوردن دیدگاه یه پیام در مقابل یک شایعه است. علاوه بر این، الگوریتم های نظارت شده

¹ Gaussian process

² Baseline

³ Hawkes process

⁴ Hidden Markov Model

⁵ Multi-spaced

⁶ Large-scale data

هیبرید^۱ بسیار زمان بر و پر هزینه هستند به دلیل اینکه مدل باید برای هر شایعه جدید دوباره یاددهی شود، درحالیکه مدل های نظارت نشده ناکافی هستند زیرا که نمیتوانیم مشخصه هر طبقه بندی را از خروجی الگوریتم بدست بیاوریم. الگوریتم های نیمه نظارت شده نیاز به کار دستی کمتری دارند و همچنین سریع تر و مختص شایعه هستند بدین معنی که میتوانند متن را بر اساس محتوای هر شایعه طبقه بندی کنند. همچنان که داده های برچسب زده بیشتری در دسترس قرار میگیرد، مولفان از دو روش Label spreading و Label propagation که یک سیستم سریع تر و دقیق تر را تولید میکند استفاده میکنند. آن ها مدل را بر اساس چهار مجموعه ویژگی های متمایز ارزیابی میکنند تا بهترین سیستم را انتخاب نمایند.

۱-۱-۱-۲- یادگیری عمیق

یادگیری عمیق به عنوان یک روش استخراج ویژگی بسیار محبوبیت پیدا کرده است به این دلیل که کار دستی کمتری نسبت به روش های سنتی یادگیری ماشین لازم دارد. مقاله [22] مدل های طبقه بندی ترتیبی^۲ را تنها با ویژگی های محلی^۳ برای بدست آوردن تاثیر رشته های گفتگو^۴ بر عملکرد مدل توسعه داده است. آن ها مدل های با ساختار درختی ترتیبی و کامل را با استفاده از زنجیره خطی و درختی^۵ CRF مقایسه کردند. در رساله دیگری [23]، برگرفته از تاثیر توییت ها بر یکدیگر، آن ها مطالعات قبلی را با ترکیب کردن مدل های طبقه بندی ترتیبی بیشتر و بهره گیری از ساختار دیالوگی در توییت ها گسترش میدهند تا تاثیر رشته گفتگو توییت را بر هر دو مدل طبقه بندی ترتیبی و غیرترتیبی تحلیل کنند. آن ها متوجه میشوند که مشخصه های متن به طور محسوسی عملکرد مدل های طبقه بندی غیرترتیبی را بهبود میبخشند ولی تاثیر کمی بر مدل های طبقه بندی ترتیبی دارند زیرا که این مدل ها میتوانند بطور ذاتی مفهوم را در رشته گفتگو یاد بگیرند. در نتیجه، با اینکه اضافه کردن داده های مربوط به متن به مدل های طبقه بندی ترتیبی ممکن است باعث افزایش فراوانی شود ولی این مدل ها عملکرد بهتری نسبت به مدل های طبقه بندی غیرترتیبی بدون ویژگی های متن دارند.

¹ Hybrid supervised algorithms

² Sequential classifiers

³ Local features

⁴ Conversational thread

⁵ Conditional random fields

رساله [24] یک سیستم بی درنگ^۱ پیشنهاد داده اند که در آن از CNN برای بدست آوردن اطلاعات متن از توییت ها استفاده میشود، در حالیکه از RNN^۲ به همراه یک مکانیزم توجه^۳ در سطح توییت برای در نظر گرفتن ترتیب توییت در یک درخت گفتگو استفاده میشود. آن ها از اطلاعات زمانی^۴ برای پیش بینی کردن موضع کاربران و گرفتن میانگین خروجی هر دسته موضع برای کل درخت در جهت پیش بینی صحت شایعه استفاده کرده اند. رساله [8] ویژگی های متنی را به همراه صفت های زمانی و ارتباطی استخراج کرده است. آن ها با کمک توییت های قبلی و بعدی هر توییتی برای ویژگی های زمانی کمک میگیرند سپس پنجره ای از شش توییت استفاده میکنند، که سه تا برای توییت های قبلی و سه تا برای توییت های بعدی است، و یک مدل CNN برای استخراج ویژگی از این پیکربندی^۵ پیاده میکنند. رساله [25] یک شبکه عصبی بازگشتی با ساختار درختی به همراه تلفیقی از اطلاعات محتوایی و ساختاری برای قوی یا ضعیف کردن موضع یک گره^۶ از طریق بازگشت در درخت توزیع طراحی کرده اند. آن ها دو مدل پیشنهاد کرده اند: ۱. درخت پایین به بالا^۷ که جهت حرکت از گره بچه به گره والد است ۲. درخت بالا به پایین^۸ که جهت حرکت از گره والد به بچه است. برای پست های منبع شایعه مدل بالا به پایین بقیه مدل ها را شکست میدهد اگرچه برای پست هایی که شایعه نیستند این دو مدل نمیتوانند عملکرد خوبی داشته باشند و بقیه مدل هایی که ویژگی های آن ها مهندسی شده است بسیار بهتر عمل میکنند. این موضوع نشان میدهد که اطلاعات بیشتری نسبت به محتوا و ساختار مانند اطلاعات کاربر لازم است.

در رساله [26] یک مکانیزم توجه stochastic با استفاده از CNN برای استخراج ویژگی طرح شده است. توجه شود که مکانیزم توجه سنتی نمیتوانست اطلاعات حیاتی را در پیام کاربران به دلیل تفاوت ها در عادت خواندن و نوشتن شناسایی کند. مدل این مولفان میتواند کلمات کلیدی مربوط به هر طبقه را تشخیص دهد

¹ Real-time system

² Recurrent neural network

³ Attention mechanism

⁴ timestamp

⁵ Configuration

⁶ Node

⁷ Bottom-up tree

⁸ Top-down tree

و کلمات نامرتبط را از تمام لایه های متصل شده خارج کند. مقاله [27] از خروجی دو شبکه عصبی برای پیش بینی موضع کاربر استفاده میکند: شبکه عصبی کپسولی^۱ که متن را به عنوان بردار هایی با طول ثابت میگیرد و شبکه عصبی پرسپترون چند لایه^۲ که ویژگی های استخراج شده از متن را میگیرد. آن ها نشان دادند که تعداد رخداد های کلاس پرسش تا حد خوبی درست پیش بینی شده اند که نشان دهنده ارزش شبکه عصبی پرسپترون چند لایه است. آن ها همچنین عملکرد شبکه کپسولی را با CNN مقایسه کردند طوری که بجای شبکه کپسولی از CNN در همان تنظیمات قبلی استفاده کردند و مشاهده کردند که وقتی شبکه کپسولی استفاده میشود دقت سیستم بالاتر میرود.

مولفان مقاله [28] پیشنهاد داده اند که از مدل بین زبانی^۳ برای طبقه بندی موضع کاربران در چهار زبان شامل انگلیسی، آلمانی، دانمارکی و روسی استفاده کنند. آن ها از مدل چند زبانه BERT (که عدم توازن در داده ها را از طریق تنظیم آستانه رفع میکند) و ترجمه ماشین^۴ در چهار زبان مختلف استفاده میکنند: ۱. مدل تک زبانه که برای هر مجموعه داده بطور جداگانه مدل MBERT یاددهی میشود ۲. مدلی که در آن از MBERT که یاددهی شده بر روی زبان انگلیسی است برای دیگر زبان ها استفاده میکنند ۳. مدل MBERT که بر روی زبان انگلیسی یاددهی شده با دیگر زبان ها تنظیم دقیق^۵ شود ۴. مدلی که بر روی تمام زبان ها یاددهی میشود. مولفان به این نتیجه رسیدند که مدل سوم عملکرد بهتری نسبت به بقیه مدل ها دارد که ممکن است دلیل آن این باشد که این مدل زبان هدف را به طور خاص یاد میگیرد.

۲-۱-۱-۲- تشخیص زودهنگام

یک مسئله مهم در تشخیص شایعه این است که در سریع ترین زمان ممکن قبل از اینکه به فرد یا جامعه آسیب وارد شود شایعه تشخیص داده شود. در زمان اولیه شیوع شایعه، اطلاعات لازم برای صحت سنجی آن کم است بنابراین به جستجوی بیشتر برای رسیدن به این هدف نیاز داریم.

¹ Capsule neural network

² Multi-layer perceptron network

³ Cross-lingual

⁴ Machine translation

⁵ Fine-tune

به منظور صحت سنجی شایعه، مولفان رساله [29] تنها به متن توییت و نظرات زود هنگام کاربران اتکا میکنند. آن‌ها از CNN و مدل‌های شبکه عصبی عمیق بر پایه BERT استفاده میکنند تا موضع کاربران را از طریق یادگیری انتقالی^۱ بدست آورند با در نظر گرفتن اینکه اطلاعات برچسب برای داده‌های موضع کاربر در دست ندارند. سپس، آن‌ها عملکرد سیستم خود را با دیگر مدل‌هایی که از یادگیری انتقالی استفاده نکرده اند مقایسه میکنند. مدل بر پایه BERT دیگر سیستم‌ها را مخصوصاً در زمانی نزدیک تر به پست شدن متن شایعه اصلی شکست میدهد. رساله [30] مدل درختی شیوع شایعه را مدل میکند و یک مدل کرنل برای پیدا کردن شباهت‌های بین این درخت‌ها استفاده میکند که از اطلاعات کاربر، متن و ویژگی‌های زمانی بهره میبرد. آن‌ها مدل‌های خود را در ساعات اولیه پخش شایعه ارزیابی میکنند و با وجود اینکه اطلاعات شیوع نسبت به ساعات دیرتر بسیار کم است همچنان مدل آن‌ها مدل‌های پایه جایگزینی را که ممکن است از اطلاعات شیوع استفاده کنند ولی نتوانند ویژگی‌های آن‌ها را به خوبی یاد بگیرند شکست میدهد. با توجه به یافته‌هایشان، شایعات در ساعات اولیه ساختار توزیع پیچیده تری نسبت به خبرهای غیر شایعه دارند. مولفان رساله [25] مدلی برای تشخیص شایعه طرح میکنند که در آن از شبکه عصبی بازگشتی با ساختار درختی استفاده میکنند و نشان میدهند که الگوهای قابل قیاسی از شایعات نادرست در ساعات اولیه پخش شایعه مورد توجه قرار میگیرد.

۳-۱-۱-۲- یادگیری چند وظیفه‌ای

یادگیری چند وظیفه‌ای میتواند یک انتخاب مناسب برای صحت سنجی شایعه باشد زیرا که موضع کاربران در مورد شایعه اطلاعات مفیدی را نمایان میکند [31]. یادگیری چند وظیفه‌ای مزایای مختلفی مانند دریافت داده‌های بیشتر برای یاددهی و کاهش احتمال بیش‌برازش^۲ یک مجموعه داده تک را دارد [32].

مقاله [32] یک مدل BranchLSTM برای یادگیری بطور مشترک و چند منظوره به سه روش ارائه کرد: طبقه بندی موضع و صحت سنجی شایعه، تشخیص شایعه و صحت سنجی شایعه و تشخیص موضع و شایعه و صحت سنجی شایعه. آن‌ها نشان دادند که مدل یاددهی چند منظوره مدل‌های دیگر و مدل‌های پایه را در هر سه روش شکست میدهند. اگر چه این مدل‌ها عملکرد متفاوتی برای هر کلاس دارند. به عنوان مثال،

¹ Transfer learning

² Overfitting

مدل‌های یادگیری معمولی در پیش‌بینی شایعات غلط بهتر عمل کردند در حالی که مدل‌های یادگیری چند منظوره در تشخیص شایعه‌های تایید نشده بهتر عمل کردند.

رساله [24] از سه روش برای استفاده از نتیجه پیش‌بینی موضع کاربر در مسئله صحت‌سنجی شایعه استفاده کرده است: مدل خط‌لوله^۱ که خروجی مدل پیش‌بینی موضع را به عنوان ورودی برای تشخیص صحت شایعه استفاده می‌کند، مدل مشترک که یک تابع خطای چند منظوره برای یک مدل استفاده می‌کند و مدل انتقالی که وزن‌های مدل پیش‌بینی موضع را در کل مدل استفاده می‌کند. نتایج آن نشان می‌دهند که مدل انتقالی دقت بالاتری دارند که می‌تواند به دو علت باشد: ارتباط بین مسئله‌های موضع کاربر و صحت‌سنجی شایعه یا عدم توازن در سباز مجموعه داده‌ها. رساله [53] یک مدل Longformer Transformer با یک مکانیزم توجه برای جمع‌آوری موضع کاربران و یاددهی توسعه موضع کاربران در یک دیالوگ استفاده می‌کند. مدل طبقه‌بندی موضع کاربر و صحت‌سنجی شایعه با یکدیگر یاددهی شدند و از یک رویکرد گروهی^۲ برای بدست آوردن حداکثر دقت از بین تعداد زیادی مدل با پیکربندی‌های متفاوت استفاده کردند.

۲-۱-۲- مجموعه ویژگی‌ها^۳

در طبقه‌بندی موضع کاربران، ویژگی‌ها بسیار مهم هستند. ویژگی‌های مفید می‌توانند تاثیر به سزایی بر عملکرد مدل بگذارند [15] [14]. در این بخش به بررسی ویژگی‌های متفاوت می‌پردازیم. دسته‌بندی و مشخصه‌های آن‌ها در ادامه توضیح داده شده است.

ما پنج دسته‌بندی برای ویژگی‌ها بر اساس مطالعات انجام شده در این زمینه معرفی می‌کنیم: متنی، ساختاری، مفهومی، بر پایه گفتگو و پروفایل کاربر. در چندین پژوهش، دسته‌بندی‌های متنوعی برای ویژگی‌ها معرفی شده است. این دسته‌بندی‌ها ممکن است شبیه هم باشند و مرز مشخصی نداشته باشند که این موضوع بیان می‌کند ممکن است ویژگی‌های یکسانی در دسته‌های متفاوتی (بسته به درک آن پژوهشگر از مسئله) دیده شوند. در نتیجه، ما دسته‌بندی خودمان را از ویژگی‌ها براساس پژوهش‌های موجود پیشنهاد

¹ Pipeline model

² Ensemble

³ Feature sets

میدهیم. یک مجموعه ویژگی قوی میتواند عملکرد مدل را بسیار بهبود بخشد به همین دلیل پژوهشگران تشویق میشوند که به مطالعه شایعه و موضع کاربر و مشخصه های شبکه های اجتماعی بپردازند تا بتوانند ویژگی های باارزشی استخراج کنند.

ویژگی های متنی شامل اطلاعات زبانی مانند استفاده از کلمات با بار منفی یا نقش کلمه در جمله هستند. این ویژگی ها تنها از متن استخراج میشوند و کمترین اطلاعاتی هستند که میتوان از یک توییت بدست آورد. ویژگی های ساختاری شامل مشخصه های خاص بستر مورد نظر مانند هشتگ در توییت هستند. این ویژگی ها اطلاعات مهمی درباره موضع کاربر و همچنین اینکه آیا شایعه صحت دارد یا خیر به ما میدهند. ویژگی های مفهومی که به مفهوم یک توییت یا رابطه بین یک توییت و توییت های دیگر نگاه میکنند. این ویژگی ها شامل شباهت بین محتوای توییت های مجزا و موضع کاربران در یک رشته گفتگو (بچه یا والد بودن نود) هستند. ویژگی های بر پایه گفتگو مشخصه های یک درخت گفتگو را تعیین میکنند. این ویژگی ها میتوانند عمق یک توییت در یک ترتیبی از توییت ها و شباهت بین توییت ها را نشان دهند. اگرچه شباهت بین توییت ها میتواند یک ویژگی مفهومی نیز در نظر گرفته شود، در برخی مطالعات یک ویژگی بر پایه گفتگو در نظر گرفته شده است.

۳-۱-۲- مجموعه داده ها

مطالعه تشخیص شایعه بعد از در دسترس قرار گرفتن مجموعه داده ها به صورت چشم گیری افزایش یافت. در سال ۲۰۱۷، مسابقات تحلیل معنا^۱ مسئله تشخیص شایعه را در خود جای دادند و دو مجموعه داده بصورت جداگانه برای تشخیص شایعه و طبقه بندی موضع کاربر در اختیار شرکت کنندگان قرار دادند. در ادامه به مطالعه تعدادی از مجموعه داده های عمومی میپردازیم.

پروژه ای به نام PHEME در سال ۲۰۱۴ با هدف بهبود بخشیدن به صحت سنجی شایعه در شبکه های اجتماعی شکل گرفت [34]. مولفان این مقاله مجموعه داده برای شایعه را جمع آوری و برچسب گذاری کردند تا پژوهش ها در این زمینه را سرعت بخشند. آن ها تعداد ۹ موضوع خبر فوری^۲ را از توییت جمع آوری کردند. این مجموعه داده شامل ۴۸۴۲ توییت به انگلیسی و ۲۸۲ توییت به آلمانی است که در مجموع

^۱ Sematic Evaluation (SemEval)

^۲ Emergine news

۳۳۰ دیالوگ مربوط به شایعه را در فرمت ساختار درختی شامل میشود. بر اساس شواهد موجود در توییت های که به پست منبع پاسخ داده بودند به برچسب گذاری شایعه ها پرداختند. برای برخی از این شایعات هیچ توییتی که بتواند آن خبر را صحت سنجی کند در بازه زمانی ای که داده ها را جمع آوری کرده بودند موجود نبود. داده های مربوط به شایعه در SemEval-2017 مشمول قسمتی از داده های برچسب گذاری شده PHEME و دو رویداد جدید که به آن اضافه کردند است. داده های یاددهی شامل ۲۹۷ رشته گفتگو مربوط به ۸ رویداد است، در حالی که داده های تست شامل ۲۸ رشته گفتگو است که به مجموعه داده PHEME اضافه شده بود. افزایش پژوهش در این زمینه و ناکافی بودن مجموعه داده ها و لزوم تشخیص خبر های نادرست در فضای مجازی باعث شد مسابقات ارزیابی معنا در سال ۲۰۱۹ مجموعه داده دیگری مرتبط با شایعه در اختیار شرکت کنندگان بگذارند. این مجموعه داده جدید بزرگ تر از قبلی است زیرا که داده هایی از بستر Reddit و داده های جدیدی از توییتر به داده های قبلی اضافه کرده است. در این مسئله داده های مربوط به موضع کاربر شامل ۸۵۷۴ و داده های مربوط به شایعه شامل ۴۴۶ نمونه است.

۲-۲- خلاصه و جمع بندی

در این فصل با مفاهیم اولیه و پیش‌زمینه هایی که جهت درک هرچه بهتر موضوع مسئله، یعنی طبقه بندی موضع کاربر با استفاده از رویکرد های مختلف یادگیری ماشین و یادگیری عمیق، مورد نیاز بوده است را بیان کردیم. ابتدا به موضوع یادگیری و انواع آن اشاره کردیم و سپس به بررسی ویژگی ها و مجموعه داده های موجود پرداختیم. اکنون با داشتن این پیش‌زمینه از موضوع، می‌توانیم به سراغ ادامه‌ی این مقاله یعنی بررسی مدل های یادگیری ماشین و یادگیری عمیق برویم.

فصل 3

فصل ۳: مدل سازی و روش طراحی

در این فصل به مطالعه و بررسی بیشتر مدل های یادگیری ماشین و یادگیری عمیق میپردازیم تا از طریق آشنایی بیشتر با آن ها بتوانیم یک مدل مناسب برای مسئله این پژوهش بیابیم.

۱-۳- مقدمه

برای داشتن مدل سازی صحیحی از مسئله‌ی دسته بندی موضع کاربران، نیاز هست تا به مدل هایی که در زمینه‌ی پردازش زبان طبیعی فعالیت می کنند اشراف داشته باشیم تا بتوانیم بهترین مدلی که در چارچوب مسئله و نیازمندی ما قرار می گیرد را ارائه دهیم. در این فصل ابتدا به سراغ معرفی برخی از معروف ترین مدل های موجود در این حوزه می رویم و سپس با در نظر گرفتن معیار هایی، آن ها را مقایسه و ارزیابی می کنیم.

۲-۳- روش های مرسوم در پردازش زبان طبیعی

در این بخش به بررسی برخی از روش های مرسوم در پردازش زبان طبیعی برای مسئله‌ی دسته بندی موضع کاربر می پردازیم. در حالت کلی ما سه دسته از الگوریتم ها را در نظر گرفتیم که در ادامه به تشریح آن ها می پردازیم.

۱-۲-۳- الگوریتم های یادگیری ماشین

این دسته از الگوریتم ها از مدل های دسته بندی یادگیری ماشین در حوزه پردازش زبان طبیعی استفاده می کنند. در اینجا چهار الگوریتم معروف را مورد بررسی قرار می دهیم.

۱-۱-۲-۳- دسته بندی بیزین احمقانه^۱

این روش که مدل یادگیری تولیدکننده^۲ نیز نامیده می شود، یک مدل یادگیری ماشین احتمالی است و

^۱ Naïve Bayes Classifier

^۲ Generative Learning Model

این کار را با قانون تصمیم بیشترین پسامد^۱ انجام می دهد. این مدل برگرفته از تئوری بیزین است که در کارهای دسته بندی استفاده می شود. این مدل که از الگوریتم های قدیمی تر برای تصمیم گیریست، مقیاس پذیر^۲ بوده و با کاربرد های دنیای واقعی نیز ارتباط دارد. این کاربرد ها پاسخ سریع به کاربران می دهند. این الگوریتم برای فیلتر کردن اسپم ها در ایمیل یا تشخیص احساسات کاربرد زیادی دارد [35].

۲-۱-۲-۳- نزدیک ترین همسایه^۳

الگوریتم چندمین نزدیک ترین همسایه یک الگوریتم یادگیری ماشین تحت نظارت سریع و راحت است که مسائل دسته بندی و برگشتی^۴ را حل می کند. به عنوان مثال، تعدادی نقطه ی برچسب گذاری شده را در نظر بگیرید و فاصله ی بین هر دو نقطه را محاسبه و سپس از بزرگ به کوچک مرتب کنید و اولین نقطه را k برچسب گذاری کنید. این برچسب های k ، میانگین و مد مدل های دسته بندی و برگشتی را برمی گردانند. این یک مدل حساس است و مسائل دسته بندی خروجی های گسسته ای دارند. این الگوریتم برای شناسایی الگوها و یا مدل تشخیص نفوذ نیز کاربرد دارد [35].

۳-۱-۲-۳- ماشین بردار پشتیبان^۵

این یک الگوریتم یادگیری تحت نظارت است و در مقایسه با الگوریتم های دیگر، یکی از الگوریتم های بهترین سازگاری^۶ است. ماشین بردار پشتیبان از زبان شناسی محاسبه ای برای تشخیص اخبار جعلی استفاده می کند. این الگوریتم برای تبدیل مدل های یادگیری که کاربردهای خاص خودش را دارد استفاده می شود. در

¹ Max Posteriori Decision Rule

² Scalable

³ Nearest Neighbour

⁴ Regression

⁵ Support Vector Machine

⁶ Best Fit Algorithm

این روش همسان سازی^۱ روی مجموعه داده ها استفاده می شود، که در بخش داده های یادگیری نقش خوبی را ایفا کرده است. ماشین بردار پشتیبان یکی از پرکاربردترین الگوریتم های دسته بندی است و غالباً برای استخراج حجم زیادی از داده ها و تعداد کمی از ابعاد عکس ها استفاده می شود [35].

۴-۱-۲-۳- درخت تصمیم^۲

درخت تصمیم یک مدل یادگیری تحت نظارت است که به صورت فلوچارت نمایش داده می شود. این الگوریتم یک روش بدون متغیر است که برای مسائل دسته بندی و برگشتی استفاده می شود. در درخت تصمیم، ما داده های منبع را به زیردسته هایی بر اساس مقادیر آنها تقسیم بندی می کنیم و این فرآیند به صورت بازگشتی روی هر زیردسته تکرار می شود تا کامل شود و به گره های برگ^۳ برسیم. این الگوریتم می تواند بر روی داده های طبقه بندی شده و یا اعداد کار کند [35].

۴-۲-۳- الگوریتم های یادگیری عمیق

در این بخش به بررسی برخی الگوریتم های معروف در زمینه یادگیری عمیق می پردازیم.

۱-۲-۲-۳- شبکه های عصبی درهم پیچیده^۴

شبکه های عصبی درهم پیچیده یک شبکه ای از نورون های بهم متصل در لایه های مختلف هستند که وظیفه ی آنها دریافت ورودی و سپس خروجی دادن بر اساس آنهاست. این شبکه ها نوعی از مدل شبکه های روبه جلو^۵ هستند که برای تشخیص اشیاء و تحلیل تصویر مناسب هستند. در این شبکه ها سه مرحله ی اصلی

¹ Normalization

² Decision Tree

³ Leaf Nodes

⁴ Convolutional Neural Network

⁵ Feed Forward Network

وجود دارد که هر کدام وظیفه خود را دارد: لایه‌ی درهم پیچیده، لایه‌ی تشخیص دهنده، و لایه‌ی جمع کننده. در لایه‌ی درهم پیچیده، ساختار ویژگی‌های درهم پیچیده ساخته می‌شود. در لایه‌ی تشخیص دهنده، المان‌های غیرخطی از ساختار ویژگی‌ها ایجاد می‌شود. لایه‌ی جمع کننده نیز اطلاعات قبلی را تجمیع کرده و سپس خروجی می‌دهد. شبکه‌ی CNN کاراکترهای ساکت را در مطالب خبری می‌تواند پیدا کند. مهم‌ترین استفاده‌ی این مدل، حجم داده‌ها و داده‌های یاددهی است و بر اساس سرعت و عملکرد مورد بررسی قرار می‌گیرد [35].

۳-۲-۲-۲- شبکه‌های عصبی برگشتی

شبکه‌های عصبی برگشتی بوسیله‌ی انتشار به عقب یاددهی می‌شوند و جزو دسته‌ی شبکه‌های عصبی مصنوعی محسوب می‌شوند. این شبکه‌ها همچنین روبه‌جلو حساب می‌شوند چرا که ورودی را از حلقه‌های تکرار شونده دریافت می‌کنند. شبکه RNN تحلیل داده‌ها را به صورت ترتیبی انجام می‌دهند، مانند تحلیل احساسات، تشخیص صوت، و سایر کارها. این شبکه یک مدل با حافظه است که ورودی‌های گذشته را دریافت می‌کند. این مدل درک خوبی از زبان انسان و پاسخ‌های متناسب با آن را دارد. نمونه بارزی از RNN، برنامه‌ی Siri از کمپانی Apple و یا Alexa از کمپانی Amazon است. این مدل نمی‌تواند آینده را بر اساس داده‌های گذشته پیش‌بینی کند، بلکه اطلاعات گذشته را به خاطر می‌سپارد و از همان پارامترها برای ورودی‌ها یا لایه‌های پنهان بیشتر برای تولید خروجی بهره می‌گیرد [35].

۳-۲-۲-۳- شبکه‌های عصبی مصنوعی

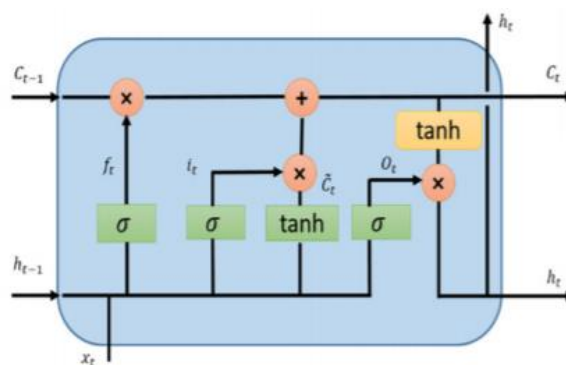
شبکه‌های عصبی مصنوعی از دسته الگوریتم‌های محاسباتی هستند. ساختار موضعی آن از الگوهای پیچیده و غیرخطی تقلید شده است. در پردازش تصاویر ماهواره‌ای از این نوع شبکه‌ها استفاده می‌شود. شبکه ANN به نوروهای انسان شباهت زیادی دارد، به این صورت که سیگنال الکتریکی ارسال می‌کند، و یا تعداد زیادی واحدهای پردازشی پیوسته به هم دارد که با یکدیگر کار می‌کنند. این شبکه نیز روبه‌جلو حساب می‌شود که یک لایه‌ی ورودی، یک لایه‌ی خروجی، و تعدادی لایه‌ی پنهان دارد. هدف از لایه‌ی ورودی، دریافت ورودی و ارتباط گرفتن با لایه‌های مخفی است. به همین ترتیب، لایه‌های پنهان، ورودی‌های دریافت شده از لایه‌ی ورودی را با هم ترکیب می‌کنند و پاسخ آن را به لایه‌ی خروجی می‌فرستند. در این شبکه‌ها، ما مدل را با تعدادی داده ورودی تغذیه می‌کنیم و خروجی‌های آن را با خروجی‌های واقعی مقایسه می‌کنیم [35].

۳-۲-۳- الگوریتم های مدرن

در این بخش به بررسی دسته ی دیگری از الگوریتم ها می پردازیم که در حوزه ی پردازش زبان طبیعی کاربرد بسیاری دارند و مقاله های زیادی از این روش ها در پیاده سازی مدل خود استفاده کرده اند. در این بخش، دو مدل از شبکه های عصبی برگشتی و سپس دو مدل از شبکه های یادگیری انتقالی را تشریح می کنیم.

LSTM -۳-۲-۳-۱

شبکه های LSTM نوعی از شبکه های عصبی برگشتی، و سلول های LSTM نوعی از نورو ن های برگشتی هستند که نتایج بسیار جالبی را در مسئله های مدل های ترتیبی نشان داده اند که به دلیل این قابلیت است که این مدل می تواند اطلاعات را از گذشته به یاد بیاورد. واحدهای LSTM ترکیبی از تعدادی درگاه هستند که مسئولیت نگهداری از وضعیت یک سلول پنهان را دارد که به آن ها اجازه می دهد تا بخشی از مسئله ی شیب مخفی شونده^۱ را حل کنند و به این ترتیب آن ها توانایی یادآوری اطلاعات بسیار دورتری نسبت به واحدهای برگشتی معمولی دارند. این ویژگی مخصوصاً در پردازش زبان طبیعی اهمیت پیدا می کند، چرا که در جملات متنی برای فهمیدن آن نیاز به هم بافت کلمات داریم و هر چقدر به عقب تر بتوانیم برویم، می توانیم تاثیر آن ها را بر کلمات فعلی دریابیم و نتیجتاً درک بهتری از متن داشته باشیم [36]. در شکل (۳-۱) ساختار این مدل را می توانیم ببینیم.



شکل (۳-۱) ساختار مدل حافظه های کوتاه مدت بلند [37]

¹ Vanishing Gradient Problem

به طور دقیق تر، این معماری می تواند از لایه های LSTM دو طرفه^۱ نیز بهره ببرد، که یعنی در یک توالی (مانند یک متن) هم به جلو و هم به عقب می تواند حرکت کند. این طراحی به این دلیل است که کلمات آینده نیز می توانند معنی کلمات فعلی را عوض کنند [36]. مثلاً در کلمه ی «شیر» ما برای فهمیدن معنی استفاده شده ی آن در جمله نیاز به هم بافت بیشتری داریم.

نوع ترکیبی^۲ دیگری از این الگوریتم به نام ConvLSTM وجود دارد که با شبکه های عصبی درهم پیچیده ادغام شده است که مشکل اصلی مدل های LSTM تمام متصل^۳، یعنی بکاربردن داده های فضایی-زمانی، را برطرف می کند. این مدل کارایی خوبی برای تحلیل داده های موقتی در مجموعه داده های ویدیویی دارد [37].

۴-۳-۲-۳-۲ GRU

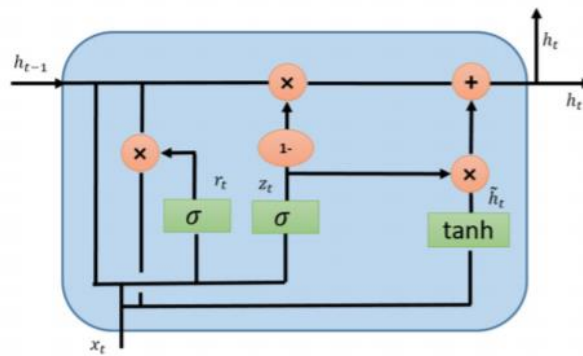
مدل GRU که در انجمن دانشمندان داده ای که با شبکه های تکرارشونده کار می کنند معروف است، نوعی از LSTM محسوب می شود. مهمترین دلیل معروفیت این مدل، هزینه ی محاسباتی و سادگی کار با آن است، که در شکل (۳-۲) ساختار آن آمده است. مدل GRU ورژن سبک تری نسبت به LSTM از رویکرد RNN در حوزه ی ساختاری، هزینه ای و پیچیدگی است. این مدل، درگاه های فراموشی و ورودی را با هم ترکیب کرده و به یک درگاه بروزرسان تبدیل می کند و سپس وضعیت سلول ها و سلول های پنهان را با هم ادغام می کند. حتی مدل های ساده تر از GRU نیز به سرعت در حال معروف شدن هستند [37].

¹ Bidirectional LSTM

² Ensemble

³ Fully Connected

⁴ Gated Recurrent Units



شکل (۳-۲) ساختار روش واحدهای درگاهی تکرارشونده [37]

۳-۲-۳-۲ GPT^۱

همان طور که در فصل قبلی گفته شد، مدل های از پیش یاددهی شده از دو مرحله تشکیل می شوند: مرحله ی pre-training و مرحله ی fine-tuning. مدل GPT که یکی از مدل های یادگیری انتقالی است، با مجهز بودن به یک تبدیل کننده^۲ رمزگشا^۳ در زیرساخت آن، در مرحله ی اول به صورت تولیدکننده و در مرحله ی دوم به صورت جداکننده عمل می کند. این مدل در مقایسه با مدل های قبلی خود، اولین مدلی است که معماری تبدیل کننده ی مدرن را با هدف پیش یاددهی تحت نظارت ترکیب می کند. به صورت تجربی، GPT موفقیت بسیاری در تقریباً تمامی کارهای NLP داشته است، که این شامل تحلیل زبان طبیعی، پرسش و پاسخ، استدلال منطقی، تشابه احساسات، و انواع دسته بندی ها است. با داشتن متون بدون برچسب گذاری با حجم زیاد، این مدل از روش مدل سازی استاندارد زبان خودبرگشتی^۴ استفاده می کند، به این صورت که احتمالات شرطی تمام کلمات با در نظر گرفتن کلمات قبل از آن در قالب همبافت را پیشینه می کند. در مرحله ی pre-training احتمالات شرطی هر کلمه توسط تبدیل کننده مدل می شود. برای هر کلمه، GPT توزیع احتمالاتی آن را با استفاده از انجام عملیات خودتوجهی چندسر^۵ روی کلمه های قبلی و با تمرکز بر موقعیت آن ها محاسبه می کند [38].

¹ Generative Pre-trained Transformer

² Transformer

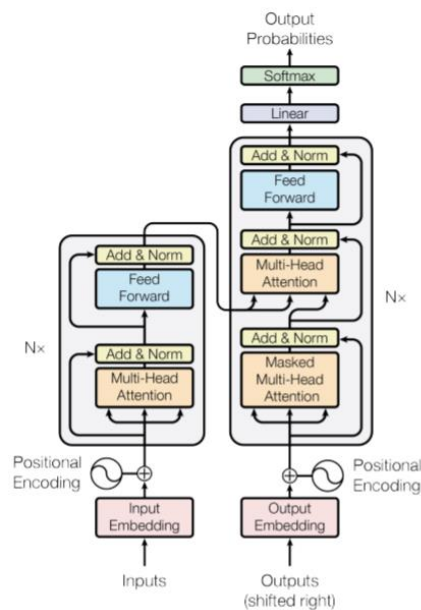
³ Decoder

⁴ Standard Autoregressive Language Modeling

⁵ Multi-head self-attention operations

BERT - ۳-۲-۳-۴

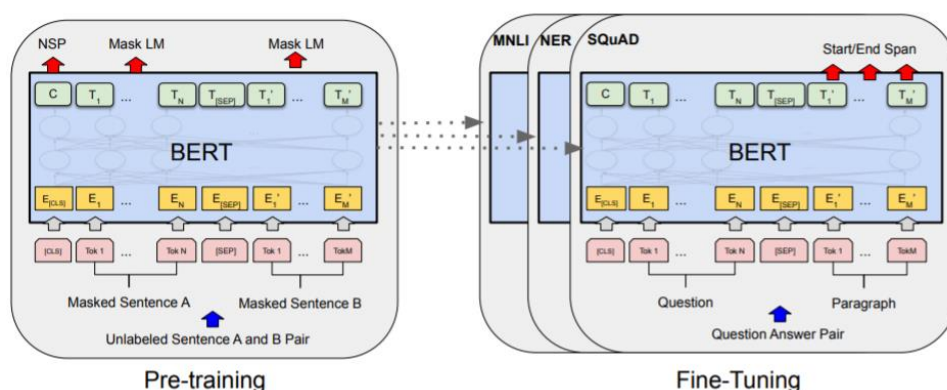
مدل BERT که اختصاری برای نمایش رمزکننده‌های دوطرفه از تبدیل‌کننده‌ها است، یک مدل زبانی بر اساس تبدیل‌کننده‌ها است که توسط گوگل^۱ ایجاد شده است. به‌طور خلاصه، ترکیبی از بلوک‌های تبدیل‌کننده‌ی رمزکننده است که بر روی همدیگر سوار شده‌اند. این مدل که در شاخه‌ی یادگیری انتقالی قرار می‌گیرد، با متون زیادی از منابع مختلف مانند کتاب‌ها، ویکی‌پدیا^۲، و... یاددهی شده است با این هدف که کاربر نهایی بتواند در مرحله‌ی fine-tuning مدل را با مسئله‌ی به‌خصوص خود وفق دهد. گوگل مدل‌های ازپیش‌یاددهی‌شده‌ی متعددی ارائه داده‌است، مانند مدل BASE یا مدل LARGE، که در پارامترهایشان با یکدیگر تفاوت دارند؛ مثلاً در مدل اول از ۱۲ بلوک استفاده می‌کند در حالی که مدل دوم ۲۴ بلوک دارد. به دلیل محدودیت‌های محاسباتی، معمولاً افراد از مدل BASE استفاده می‌کنند [36]. در شکل (۳-۳) ساختار بلوک‌های تبدیل‌کننده را می‌بینیم. همچنین در شکل (۴-۳) ساختار کلی دو مرحله‌ی مدل BERT به نمایش گذاشته شده است.



شکل (۳-۳) ساختار بلوک‌های تبدیل‌کننده [36]

¹ Google

² Wikipedia



شکل (۳-۴) ساختار کلی دو مرحله‌ی مدل BERT [39]

۳-۳- گزینش مدل

مدل‌های فوق‌الذکر تنها مدل‌های موجود در این حوزه‌ها نیستند. مدل‌های دیگری برای انجام کارهای مشابه نیز وجود دارند، مانند RBM^۱، AE^۲، و یا GAN. همچنین مدل‌هایی که بالاتر به آن‌ها پرداختیم، حالت‌های مختلفی را نیز پشتیبانی می‌کنند که بعضی از آن‌ها مدل بروزشده‌ی مدل اصلی هستند و بعضی دیگر برای کارهای خاصی شخصی‌سازی شده‌اند؛ مثلاً برای BERT، مدل RoBERTa و DeBERTa و یا برای مدل RNN حالت GCRNN نیز وجود دارند.

ما در این پژوهش، بر اساس میزان استفاده، محبوبیت، زمان انتشار، و نتایج‌های بدست آمده از آن‌ها در مقاله‌ها، ۱۱ مدلی که در بخش قبلی به تشریح آن‌ها پرداختیم را از میان صدها مدلی که وجود دارند انتخاب کردیم و با توجه به مطالعه انجام شده و بسیار به روز بودن و عملکرد بسیار خوب مدل BERT تصمیم به پیاده سازی این مدل گرفتیم.

^۱ Restricted Boltzmann Machine

^۲ Auto Encoder

۴-۳- تحلیل نتایج و ارائه مدل پیشنهادی

بر اساس معیارهای ارائه شده و همچنین ارزیابی‌های انجام شده بر اساس آن معیارها، مدل BERT از نظر ما مدل برگزیده بود و در این تحقیق می‌خواهیم آن را پیاده‌سازی کنیم. در این راستا، با مطالعه بیشتر این مدل و پارامترهای تلاش میکنیم که پیاده‌سازی مناسبی از این مدل داشته باشیم.

۵-۳- خلاصه و جمع‌بندی

فصل سوم به طور عمده در برگیرنده‌ی مدل‌های مختلف در حوزه‌ی پردازش زبان طبیعی بوده که پس از توضیحات کامل آن‌ها، معیارهایی برای ارزیابی ارائه شد و سپس بر اساس آن معیارها، مدل‌ها و الگوریتم‌ها مقایسه شدند و در نهایت مدل پیشنهادی ما بر اساس آن‌ها ارائه شد. در بخش بعدی به سراغ پیاده‌سازی مدل پیشنهادی رفته و نتایج خود را ارائه می‌کنیم.

فصل 4

فصل ۴: پیاده سازی مدل پیشنهادی

پس از بررسی های انجام شده روی مدل های مختلف در فصل قبل، در این فصل می خواهیم جزئیات بیشتری از مدل پیشنهادی خود ارائه دهیم و سپس بر اساس آن جزئیات، روند پیاده سازی انجام شده در طول این پروژه را تشریح کنیم.

۱-۴- مقدمه

آنچه تا اینجا دیدیم، مباحث تئوری مورد نیاز برای آشنایی با مسئله و انتخاب و پیاده سازی یک مدل بود که شامل بررسی مدل های مختلف، نقاط مثبت و منفی و سپس ارزیابی آن ها بود. در این فصل می خواهیم ابتدا به جزئیات مدل پیشنهادی خود پرداخته و سپس مراحل و روند پیاده سازی که انجام شده است را ارائه دهیم. در پایان نیز خروجی های خود را ارزیابی کرده و با مدل های دیگر مقایسه میکنیم.

قبل از بررسی مدل پیاده سازی شده ابتدا دو مفهوم را برای درک بهتر این مدل توضیح میدهم:

۱. نمایش زبان بدون در نظر گرفتن مفهوم متن^۱: هر کلمه بدون تاثیری از کلمات

اطرافش در نظر گرفته میشود.

۲. نمایش زبان با در نظر گرفتن مفهوم متن: هر کلمه در جمله با تاثیر پذیری از دیگر

کلمات جمله ارائه میشود.

از آنجایی که مورد دوم یعنی نمایش زبان با در نظر گرفتن مفهوم متن رابطه هر کلمه با دیگر کلمات را در نظر میگیرد پس جهت گیری آن مهم است:

۱. جهت گیری یک طرفه^۲: کلمه بعدی را بر اساس کلمات قبلی پیش بینی میکند.

۲. جهت گیری دو طرفه^۳: کلمه را بر اساس کلمات قبل و بعد از آن پیش بینی میکند.

۲-۴- مدل پیشنهادی

در این مقاله می خواهیم از یک مدل BERT که هم مفهوم متن را در نظر میگیرد و هم جهت گیری دو طرفه دارد استفاده کنیم. BERT یک مدل پردازش زبان طبیعی منبع باز^۴ است که از

¹ Context-free language representation

² Unidirectionality

³ Bidirectionality

⁴ Open-source

پیش یاددهی شده و توسط تیم زبان هوش مصنوعی شرکت گوگل در سال ۲۰۱۸ توسعه یافته است. این مدل به عنوان یک توسعه پیشگامانه در زمینه پردازش زبان طبیعی به حساب می آید. مزیت این مدل بر دیگر مدل ها به دلیل استفاده از یادگیری دو طرفه (که اولین بار در این مدل پیاده سازی شد) این است که فهم بیشتر و عمیق تری از متن میتواند داشته باشد.

۱-۲-۴- معرفی مدل

در این مقاله ما میخواهیم یک مدل طبقه بندی موضع کاربران در مورد شایعه در شبکه های اجتماعی معرفی کنیم که مجموعه داده SemEval-2017 را تحلیل کنیم. یک روش مورد استفاده در شبکه های عصبی که سال های اخیر بسیار مورد استقبال قرار گرفته است استفاده از یادگیری انتقالی است که ابتدا یک شبکه عصبی بر روی مجموعه داده مربوط به یک مسئله خاص یاددهی میشود و سپس تنظیم دقیق انجام میشود که در آن از این شبکه عصبی یاددهی شده برای یک مدل با هدفی جدید استفاده میشود. مدل BERT از انتقال که یک مکانیزم توجه است که ارتباط مفهومی بین کلمات یک متن را یاد میگیرد استفاده میکند. روش انتقالی یک مدل دو طرفه محسوب میشود زیرا که تمام کلمات یک متن را به یکباره میخواند در حالی که مدل های دیگر متن را با ترتیب میخوانند (راست به چپ یا برعکس). این ویژگی مدل انتقالی اجازه میدهد که مفهوم یک کلمه را بر اساس کلمات اطرافش یاد بگیریم. در یاددهی یک مدل پردازش زبان طبیعی، بسیاری از مدل ها کلمه بعدی را در یک ترتیب کلمات پیش بینی میکنند که این روش یک طرفه محسوب میشود و یادگیری مفهوم را برای مدل محدود میکند. مدل BERT برای حل این چالش یک استراتژی یادگیری به نام Masked Language Model (MLM) را بکار میبرد که به این صورت است که نیاز است تا تعدادی از کلمات بطور تصادفی پنهان شوند و سپس مدل آن ها را حدس بزند که مدل BERT با پنهان شدگی ۱۵ درصد یاددهی شده است. قبل اینکه ترتیب کلمات به BERT داده شوند، ۱۵ درصد کلمات در هر ترتیب با یک token جایگزین میشوند. سپس، مدل سعی میکند مقادیر اصلی این کلمات را بر اساس کلمات دیگر موجود در ترتیب حدس بزند. مدل BERT یک شبکه عصبی عمیق است که تعداد زیادی لایه دارد و داده و زمان زیادی لازم است تا یادگیری انجام شود. برای برطرف کردن این مشکل، از مجموعه داده های بزرگی استفاده شده و یادگیری در لایه های اولیه آن انجام شده است که در واقع زبان انگلیسی را در این

لایه ها فرا گرفته است. به همین دلیل، این مدل عنوان pre-trained BERT را میگیرد و ما از مجموعه داده مورد نیاز خودمان برای تنظیم دقیق این مدل استفاده میکنیم که بدین معنی است که یادگیری لایه های آخر بر روی مجموعه داده ما انجام میشود. پس این مدل در واقع وزن هایی از چندین مجموعه داده کلی یاد گرفته است و ما با اجرا کردن این مدل بر روی مجموعه داده خودمان از وزن های قبلی استفاده میکنیم و وزن های جدیدی مرتبط با این مجموعه داده بدست می آوریم.

۴-۲-۲- مجموعه داده ها

در این مقاله از مجموعه داده RumourEval-2017 که شامل دو دسته داده یاددهی و توسعه است استفاده میکنیم. داده های توسعه تنها ۶.۲۲ درصد کل مجموعه داده را مشمول میشوند که از آن برای اعتبار سنجی^۱ نهایی استفاده میشود. به منظور حل این مسئله، داده های یاددهی و توسعه را یکی میکنیم و سپس فرآیند جداسازی داده ها را با نسبت ۱۵ درصد برای اعتبار سنجی به دو دسته داده تقسیم میکنیم.

داده های RumourEval-2017 از یک مجموعه داده بسیار بزرگ به نام PHEME گرفته شده است که در فصل دو توضیح داده ایم. از این داده ها در سال ۲۰۱۷ برای برگزاری مسابقات SemEval مربوط به طراحی مدل برای پردازش زبان طبیعی استفاده شده است که یکی از مسئله های مطرح شده در آن مرتبط به شایعه بوده است. سپس، در سال ۲۰۱۹، این مجموعه داده را گسترش دادند و از پست های بستر شبکه اجتماعی Reddit نیز استفاده کردند. ما در این پژوهش، تصمیم به استفاده از مجموعه داده های سال ۲۰۱۷ گرفتیم به این علت که ابتدا به تحلیل داده های تنها یک بستر پردازیم و با جنبه های مختلف آن در یاددهی مدل آشنا شویم سپس در گام های بعدی این پروژه به بررسی تشخیص شایعه در بقیه بستر ها نیز پردازیم.

¹ Validation

۳-۲-۴- نحوه استفاده از داده‌ها در مدل

داده های اصلی این مجموعه داده که از رابط برنامه نویسی توییتر^۱ گرفته شده است شامل تمام اطلاعات یک توییت مانند لینک، تعداد دنبال کنندگان، تعداد پست ها، متن توییت و ... است. داده هایی که ما به آن نیاز داریم متن توییت، شناسه توییت، شناسه توییتی که به آن پاسخ داده است و برچسب موضوع کاربران است. این داده ها را از مجموعه داده اصلی جدا میکنیم و در نهایت قالب داده^۲ مورد استفاده ما همانند شکل (۴-۱) تهیه شده است.

tweet_id	text	source	label
544288681021145090	Up to 20 held hostage in Sydney Lindt Cafe sie...	None	support
544290348525449216	"@Independent: Up to 20 held hostage in Sydney...	544288681021145090	support
544290635013189632	@Independent @Hragy oh god !!!!	544288681021145090	comment
544288801926156289	"@Independent: Up to 20 held hostage in Sydney...	544288681021145090	support
544297324466282497	@Independent Can we take the gloves off and go...	544288681021145090	comment

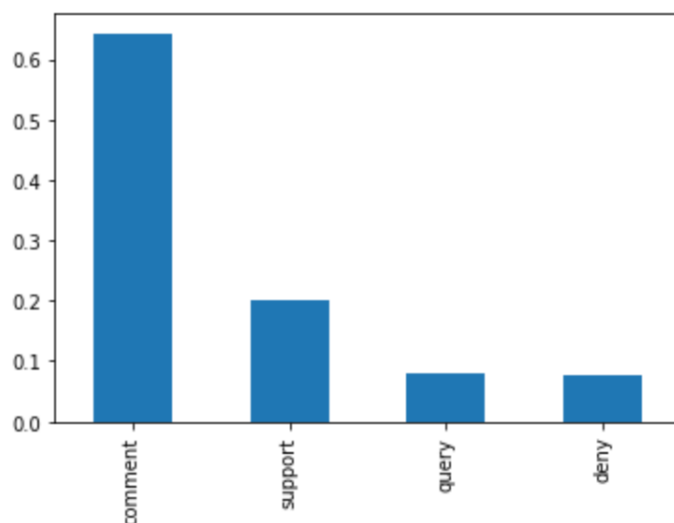
شکل (۴-۱) مجموعه داده مرتب شده جهت استفاده در مدل یادگیری

همانطور که در توضیحات مجموعه داده مطرح شد کلاس های مجموعه داده های مسئله موضوع کاربر بسیار نامیزان هستند و کلاس نظر دادن که کمترین اطلاعات را به ما میدهد دارای بیشترین تعداد داده و کلاس رد کردن که بیشترین اطلاعات را میدهد دارای داده های بسیار کمی است بطوریکه در یک سری مقالات مدل نهایی نتوانسته هیچ داده ای از این کلاس تشخیص دهد. این

¹ Twitter API

² Dataframe

عدم تعادل را در شکل (۲-۴) مشاهده میکنیم. همچنین، شکل (۳-۴) توزیع عددی داده ها را برای هر کلاس برچسب و دو دسته یاددهی و اعتبار سنجی ارائه داده است.



شکل (۲-۴) توزیع داده های موجود در هر کلاس برچسب

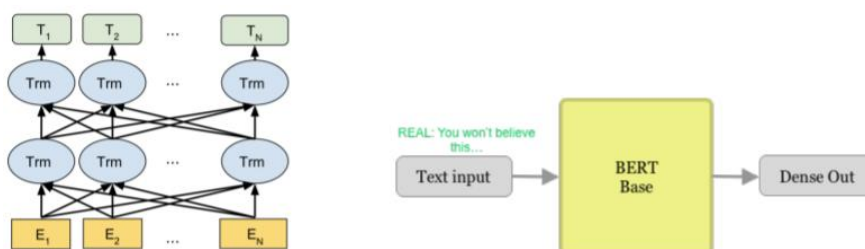
label_count		
label	data_type	
comment	train	2471
	val	436
deny	train	292
	val	52
query	train	304
	val	54
support	train	774
	val	136

شکل (۳-۴) تعداد داده های هر کلاس در دو دسته یاددهی و اعتبار سنجی

۴-۲-۴- معماری های استفاده شده

ما از پیش پردازش متن مانند special token, padding, attention mask استفاده می‌کنیم و یک مدل طبقه بندی با استفاده از BERT و کتابخانه مدل های انتقالی به نام Hugging Face می‌سازیم. دو نوع مدل BERT تعریف شده است که BERT Base و BERT Large هستند. نوع پایه شامل تعداد ۱۲ رمزگذار انتقالی، ۱۲ مدل توجه و ۳۴۰ میلیون پارامتر است.

دو قدم در ساختن یک مدل BERT وجود دارد: از پیش یادهی مدل که روی داده های بدون برچسب و چندین مسئله از پیش یادهی شده آموزش می‌بیند و سپس تنظیم دقیق مدل که ابتدا با استفاده از پارامتر های از پیش یادهی شده مقداردهی اولیه می‌شود و سپس تمام پارامتر ها با استفاده از داده های برچسب خورده تنظیم دقیق می‌شوند. مدل از پیش یادهی شده BERT می‌تواند تنها با یک لایه آخر تنظیم دقیق شود تا تعداد زیادی از مسائل را مانند ترجمه ماشین عصبی^۱، پاسخ دادن به سوالات، تحلیل احساسات و خلاصه کردن متن بدون تصحیح معماری قابل توجهی یاد بگیرد. عناصر یک مدل BERT شامل جهت گیری دو طرفه، رمزگذاری، مدل انتقالی و کلمات تعبیه شد از پیش یادهی شده^۲ است.



شکل (۴-۴) معماری مدل BERT

۴-۲-۵- مراحل و روش پیاده سازی

در ادامه به توضیح مفصلی از پیاده سازی کد مسئله می‌پردازیم. بعد از اینکه توزیع داده ها را مشاهده کردیم و جداسازی مجدد آن ها را انجام دادیم (همانطور که در بخش (۴-۲-۳) توضیح

¹ Neural machine translation

² Pre-trained word embedding

داده شد) برچسب ها را از حالت متن به عدد تبدیل میکنیم و سپس به پیش پردازش متن میپردازیم. بیشتر مدل های شبکه عصبی با شکستن ترتیبی از رشته متن به کلمات و عبارات تک شروع میکنند فرآیندی که به آن tokenization گفته میشود. یک tokenizer یک لغت نامه^۱ میسازد و ترتیبی از کلمات را به ترتیبی از اعداد تبدیل میکند. هر عددی مربوط به یک کلمه در این لغت نامه است که تمام مجموعه نوشته^۲ را کدگذاری میکند. بطور مشابهی، مدل BERT از طریق سه مرحله token های ورودی را پردازش میکند و به این ترتیب هر token حاصل سه نوع نهفته شدگی است: token embedding – segment embedding – position embedding

همانند تمامی شبکه های عصبی دیگر مدل BERT متن خام را به عنوان ورودی نمیگیرد بنابراین اولین قدم این است که متن را به اعداد تبدیل کنیم. به این منظور ما از مدل bert-base-uncased استفاده میکنیم که تفاوتی بین English و english قائل نمیشود. برای بکار بردن مدل از پیش یادهی شده BERT باید از یک tokenizer که توسط کتابخانه BertTokenizer ارائه میشود استفاده کنیم. این به این دلیل است که این مدل یک لغت نامه ثابت دارد و BERT tokenizer یک روش مخصوصی برای مدیریت کلماتی که خارج از این لغت نامه هستند دارد. این tokenizer روی تمامی مجموعه داده یادهی میشود که این بدین معنی است کلمات یکتا در داده ها را میابد و به هر کدام یک عدد یکتا اختصاص میدهد.

برای tokenize کردن داده هایمان از batch-encode-plus استفاده میکنیم که token های مخصوصی به ابتدا و انتهای هر جمله اضافه میکند. مدل BERT از این نشان های خاص برای دریافت صحیح ورودی استفاده میکند. نشان [SEP] به مدل کمک میکند تا متوجه پایان یک ورودی و شروع ورودی دیگر در همان ترتیب ورودی^۳ شود. نشان [CLS] یک نشان خاص طبقه بندی است که اجازه میدهد مدل BERT بداند که ما یک مسئله طبقه بندی داریم. ورودی این مدل باید به صورت زیر باشد:

[CLS] Sentence A [SEP] Sentence B [SEP]

¹ Dictionary

² Corpus

³ Sequence input

این باعث میشود مدل متوجه شود که جملات A و B پی در پی هستند یا خیر. اگر احتمال ۰.۵ داشته باشند پی در پی محسوب میشوند و در غیر اینصورت رندوم محسوب میشوند. حال داده های ما آماده برای مدل BERT است. حال ما مدل BERT را در چهار مرحله پیاده میکنیم: آماده سازی مدل از پیش یادهی شده BERT، ساخت DataLoader، آماده سازی Optimizer و آماده سازی Scheduler. هر کدام از این مرحله ها در ادامه مقاله توضیح داده میشود.

آماده سازی مدل از پیش یادهی شده BERT: از آنجایی که مدل BERT برای پیش یادهی شدن نمایش عمیق دو طرفه^۱ از متن بدون برچسب با استفاده از شرایط گذاشتن روی متون سمت چپ و راست در تمام لایه ها طراحی شده است، میتواند تنها با یک لایه اضافه دیگر تنظیم دقیق شود. ما از الگوریتم BertForSequenceClassification که برای مسئله های طبقه بندی طراحی شده است استفاده میکنیم.

ساخت DataLoader: ما از TensorDataset، که راهی برای بوجود آوردن یک مجموعه داده از داده هایی که در حافظه ذخیره شده میدهد، برای ساخت DataLoader استفاده میکنیم. این مدل داده را در شکل آرایه numpy و tensor میپذیرد. حال یک iterator برای مجموعه داده هایمان با استفاده از DataLoader از کتابخانه torch استفاده میکنیم. این موضوع به حفظ فضای حافظه و سرعت بخشیدن به فرآیند یادهی کمک میکند.

آماده سازی Optimizer و Scheduler: ما از بهینه ساز AdamW با نرخ یادگیری 10^{-5} ، مقدار اپسیلون 10^{-8} که مقادیر پیش فرض مدل هستند و همچنین عدد 10 برای epoch استفاده میکنیم. هر موقع که یک batch به مدل داده میشود scheduler صدا میشود. با صدا زدن optimizer.step() پارامتر های مدل تجدید میشوند و با صدا زدن scheduler.step() نرخ یادگیری تجدید میشود.

در نهایت با تعیین random seed یادهی را آغاز میکنیم و همچنین از کتابخانه tqdm برای دنبال کردن پیشرفت یادهی استفاده میکنیم. برای ارزیابی مدلمان از معیار های ارزیابی accuracy و f1-score استفاده میکنیم. به دلیل سنگین بودن کد و زمان بر بودن اجرای آن، از Google Colab برای اجرای برنامه استفاده کردیم.

¹ Deep bidirectional representation

شکل زیر خروجی مدل بر روی داده های تست در هر epoch را نشان میدهد. مشاهده میکنیم که از اولین epoch به دومی تغییر قابل توجهی در f1-score داریم و سپس تغییرات بسیار نامحسوس میشوند. دلیل آن میتواند این باشد که در epoch اول داده ها به طور مناسبی مدل نشده اند. همچنین با توجه به مقادیر معیار های ارزیابی متوجه میشویم که تا دهمین epoch مدل داده ها را overfit نمیکند.



شکل (۴-۵) ارزیابی مدل در هر epoch

جدول زیر مقایسه مدل مطرح شده در این پژوهش با چهار مدل به روز است. مشاهده میکنیم که مدل ما به نتایج قابل مقایسه ای با مقاله های [40] [8] [14] رسیده است. همچنین، لازم است ذکر کنیم که معیار accuracy برای مسئله طبقه بندی موضع کاربر معیار مناسبی نیست زیرا که این مسئله دارای چهار کلاس است و داده ها به شدت عدم توازن دارند. از آنجایی که accuracy متناسب با کلاس بزرگ تر تغییر میکند و در مسئله موضع کاربر، کلاس بزرگ تر کلاس کم اهمیت تر است نمیتوانیم عملکرد مدل بر کلاس مهم تر را که دارای تعداد نمونه های بسیار کمتری است متوجه شویم. به همین دلیل، از معیار f1-score استفاده میشود که در سال ۲۰۱۹ که مجموعه داده های RumourEval-2019 ارائه داده شد این موضوع بیان گردید.

Reference	Model	Macro F1-score	Accuracy
Pamungkas <i>et al.</i> [14]	SVM	47%	79.5%
Veyseh <i>et al.</i> [8]	CNN	48.2%	82%
Kumar <i>et al.</i> [27]	Capsule Network and Multi-layer Perceptron	55%	77.6%
Yu <i>et al.</i> [40]	Hierarchical Transformer with BERT	50.9%	76.3%
Our Model	Pre-trained BERT	49.12%	70.65%

۳-۴- چالش ها و فرصت ها

چالش های متعددی برای مسئله طبقه بندی موضع کاربران در مورد شایعه در شبکه های اجتماعی وجود دارد. با وجود پژوهش های فراوانی که در این زمینه صورت گرفته، همچنان کار های زیادی برا رسیدن به یک سیستم قابل اعتماد برای صحت سنجی شایعه لازم است.

به دلیل اینکه مجموعه داده های موجود برای طبقه بندی موضع کاربران در شبکه های اجتماعی بصورت عمده ای در زبان انگلیسی هستند، بستر های رسانه های اجتماعی غیر انگلیسی از داشتن یک مکانیزم توسعه یافته تشخیص خودکار شایعه محروم اند. البته مجموعه داده های غیر انگلیسی ای تا به حال جمع آوری شده است ولی پژوهشگران بیشتری میتواند از این مجموعه داده ها استفاده کنند و روش های جدیدی برای تشخیص شایعه و یا دسته بندی موضع کاربر پیشنهاد دهند. زمانی که یک شایعه منتشر میشود میتواند در چندین بستر شبکه های اجتماعی و به زبان

های مختلف شیوع یابد. بنابراین، یک مدل چندزبانه میتواند کمک زیادی به یک سیستم تشخیص شایعه بکند. ویژگی های مختلف شبکه های اجتماعی مانند متن، تصویر، ویدئو و صدا میتوانند جهت بهبود بخشیدن نتایج مدل های طبقه بندی موضع کاربر به کار گرفته شوند. علاوه بر این، مجموعه داده ها برای داده هایی غیر از متن بسیار محدود است که این موضوع یک فرصت برای فعالیت در این زمینه و جمع آوری داده ایجاد میکند. مشکلی دیگر با مجموعه داده ها این است که اکثر آن ها از بستر توییتر جمع آوری شده اند. ویژگی های بستر های مختلف با یکدیگر فرق میکنند و کار کردن با داده های بستر های دیگر میتواند به ما مفاهیم جدید در رابطه با تشخیص شایعه دهد.

زمانی که یک مدل بر روی شایعه خاصی یاددهی میشود و سپس برای تست از شایعه ای از یک موضوع متفاوت استفاده میشود، مانند شایعه مربوط به سلامت و نامرتبط به سلامت، عملکرد مدل کاهش میابد. در نتیجه، توسعه ویژگی میتواند به عنوان یک استراتژی محکم برای دریافت بینش بیشتر از ویژگی های موضع کاربران که ارائه دقیق تری از شایعه هستند استفاده شود. تشخیص زودهنگام شایعه نیز نقشی حیاتی در جلوگیری از شیوع شایعه و آسیب احتمالی آن به افراد و جامعه دارد. با وجود اطلاعات بسیار کمی که در ساعاتی نزدیک به پست شدن شایعه در دسترس است، این مسئله با سختی های زیادی روبرو است. پژوهشگرانی که تشخیص زودهنگام شایعه را بررسی میکنند به یافته های جالبی از نشانه های شایعه در ساعات اولیه آن رسیده اند. به عنوان مثال، ساختار شیوع شایعه در زمانی نزدیک به پخش شدن آن پیچیده تر از پست هایی است که شایعه نیستند.

تمام پژوهشی که بر روی طبقه بندی موضع کاربران صورت گرفته است برای بهتر کردن دقت یک سیستم تشخیص شایعه بوده است که بعد از آن از این سیستم در یک سناریو واقعی استفاده شود تا مانع شیوع وسیع شایعه شود. در نتیجه، مدل های پیشنهاد شده و یافته ها باید در یک سیستم واقعی تست شوند. یک بحث موجود در یک سیستم واقعی این است که به کاربران گزارش دهیم که چرا یک خبر به عنوان شایعه نادرست تشخیص داده شده است. هدف از این کار، آموزش دادن کاربران شبکه اجتماعی در جهت چگونگی قضاوت صحت یک شایعه است. این یک برنامه آموزش عموم است که به مردم یاد میدهد چگونه شایعه های درست و نادرست را شناسایی کنند. علاوه بر یک الگوریتم یادگیری ماشین که ویژگی های یک شایعه را یاد میگیرد، انسان ها نیز میتوانند چگونگی ارزیابی یک شایعه را یاد بگیرند.

۴-۴- خلاصه و جمع بندی

در این فصل ابتدا به معرفی مدل انتخابی و ساختار آن و داده هایی که با آن کار کردیم پرداختیم. سپس، نتایج را بیان کردیم و کار خود را با مدل های دیگر مقایسه کردیم. در آخر به بررسی یک سری چالش ها و فرصت های پژوهش در این زمینه پرداختیم.

فصل 5

فصل ۵: جمع‌بندی، نتیجه‌گیری و پیشنهادها

آنچه در فصل‌های گذشته آموختیم و انجام دادیم را در این فصل جمع‌بندی کرده و سپس چشم‌انداز کار و پیشنهادهای موجود روی آن را بیان می‌کنیم.

۱-۵- جمع‌بندی

در این تحقیق در گام نخست ما به تعاریف اولیه مورد نیاز برای داشتن فهم صحیحی از مدل‌های پردازش زبان طبیعی پرداختیم. سپس با ورود به بررسی انواع الگوریتم‌ها و مقایسه‌ی آن‌ها، مدلی بر اساس یک مقاله مرجع پیشنهاد کردیم و در مورد جزئیات آن صحبت نمودیم. در گام آخر آن را پیاده‌سازی کرده و پس از بیان خروجی‌های گرفته شده و مقایسه آن با مقاله مرجع، چالش‌ها و ناکامی‌های موجود در این پژوهش را ذکر کردیم.

۲-۵- نتیجه‌گیری

۱-۲-۵- دستاوردها

در این پژوهش ما توانستیم با مطالعاتی که انجام دادیم یک مدل شبکه عصبی را انتخاب کنیم و علاوه بر بدست آوردن دانش پیاده‌سازی آن، معایب و مزایای آن را نیز بدانیم تا برای ادامه پژوهش بتوانیم مدل‌های با دقت بالاتر طراحی کنیم. ما در این پیاده‌سازی توانستیم با بهبود‌هایی که روی مدل و ورودی‌ها دادیم و همچنین از روش‌های خلاقانه‌ای که استفاده کردیم خروجی خوبی با وجود عدم توازن در داده‌ها بگیریم و ارزیابی قابل مقایسه‌ای با مقالات به روز داشته باشیم.

۲-۲-۵- محدودیت‌ها

در این تحقیق ما با محدودیت‌های زیادی مواجه بودیم. از عدم دسترسی به کارت‌های گرافیک با مدل مورد نیاز برای یاددهی مدل‌ها تا محدودیت‌هایی که روی سرویس‌های آنلاین از لحاظ زمانی و حجمی داشتیم.

۳-۲-۵- پیشنهادها

پیشنهاد می‌شود که برای کارهای حوزه‌ی یادگیری عمیق و یا حتی یادگیری ماشین، از طرف دانشگاه سخت‌افزارها و ابزارهای لازم جهت پیاده‌سازی آن‌ها فراهم شود تا پژوهشگران بتوانند از آن‌ها استفاده کنند. این بستر می‌تواند به‌صورت دستی و در هر پروژه جداگانه باشد و یا به‌صورت یک پروژه‌ی بزرگ در قالب یک سرویس ابری^۱ باشد که دسترسی‌ها و سخت‌افزارها را مدیریت کند و واسطه‌ی انسانی از میان برداشته شود. در ادامه پیشنهاد می‌شود که یک سیر مطالعاتی و پیاده‌سازی برای چنین موضوعاتی تعریف شود و پروژه‌هایی که در این زمینه انجام می‌شوند به همدیگر زنجیر^۲ شوند تا بتوان علاوه بر خروجی تحقیقاتی مختص یک موضوع، بتوان خروجی کاربردی از این مدل‌ها گرفت و آن‌ها را در مسائل واقعی پیاده‌سازی کرد و به کار برد. به عنوان مثال، یک فرد می‌تواند کار تحقیقاتی و بررسی مدل‌های مختلف را انجام دهد و نتایج خود را ارائه دهد، فرد بعدی با استفاده از نتایج این فرد بیايد و مدل پیشنهادی خود را ارائه و سپس پیاده‌سازی کند، نفر بعدی بازار هدف را بررسی کرده و برای این مدل کاربرد^۳ پیدا کند، یک نفر دیگر این مدل را متناسب با آن کاربرد تغییر داده و آن را به مرحله‌ی محصول برساند^۴، و در نهایت این محصول را تحویل یک سازمان مانند دانشگاه یا شرکت‌های مختلف برای استفاده بدهد.

¹ Cloud Service

² Chain

³ Usecase

⁴ Productionize

فصل 6

فصل ٦ : مراجع

مراجع

- [1] A. Zubiaga, H. Ji, and K. Knight, "Curating and contextualizing Twitter stories to assist with social newsgathering," in Proc. ACM Int. Conf. on Intell. User Interfaces (IUI), Mar. 2013, pp. 213–224.
- [2] A. Hermida, F. Fletcher, D. Korell, and D. Logan, "Share, like, recommend," *Journalism Stud.*, vol. 13, no. 5-6, pp. 815–824, Mar. 2012.
- [3] A. Zubiaga and H. Ji, "Tweet, but verify: epistemic study of information verification on Twitter," *Soc. Netw. Anal. Min.*, vol. 4, no. 1, pp. 1–12, Mar. 2014.
- [4] M. Wendling, "The saga of 'pizzagate': The fake story that shows how conspiracy theories spread," 2016, [Accessed: 14.03.2022]. [Online]. Available: <https://www.bbc.com/news/blogs-trending-38156985>
- [5] S. Murphy, "AP Twitter hack falsely claims explosions at White House," 2013, [Accessed: 14.03.2022]. [Online]. Available: <https://mashable.com/archive/ap-hacked-white-house>
- [6] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Comput. Surv.*, vol. 51, no. 2, pp. 1–36, Feb. 2018.
- [7] A. Zubiaga, M. Liakata, R. Procter, G. W. S. Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PLOS ONE*, vol. 11, no. 3, p. e0150989, Mar. 2016.
- [8] A. P. B. Veyseh, J. Ebrahimi, D. Dou, and D. Lowd, "A temporal attentional model for rumor stance classification," in Proc. ACM Conf. on Info. and Knowl. Manage. (CIKM), Nov. 2017, pp. 2335–2338.
- [9] R. Watson, "On the size of a rumour," *Stochastic Process. Appl.*, vol. 27, pp. 141–149, 1987.
- [10] "A contribution to the psychology of rumour," in *Collected Works of C.G. Jung, Volume 4: Freud & Psychoanalysis*. Princeton University Press, Dec. 2014, pp. 35–47.
- [11] L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. Wong Sak Hoi, and A. Zubiaga, "SemEval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours," in Proc. Int. Workshop on Semantic Eval. (SemEval), Aug. 2017, pp. 69–76.

- [12] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, “Rumor has it: Identifying misinformation in microblogs,” in Proc. Conf. on Empirical Methods in Natural Lang. Process. (EMNLP), Jul. 2011, pp. 1589–1599.
- [13] G. Cai, H. Wu, and R. Lv, “Rumors detection in Chinese via crowd responses,” in Proc. IEEE/ACM Int. Conf. Advances in Social Netw. Anal. and Mining (ASONAM), Aug. 2014, pp. 912–917.
- [14] E. W. Pamungkas, V. Basile, and V. Patti, “Stance classification for rumour analysis in Twitter: Exploiting affective information and conversation structure,” ArXiv, vol. abs/1901.01911, 2019.
- [15] A. Aker, L. Derczynski, and K. Bontcheva, “Simple open stance classification for rumour analysis,” in Proc. Recent Advances in Natural Lang. Process. Meet Deep Learning (RANLP), Nov. 2017, pp. 31–39.
- [16] A. Aker, A. Zubiaga, K. Bontcheva, A. Kolliakou, R. Procter, and M. Liakata, “Stance classification in out-of-domain rumours: A case study around mental health disorders,” in Proc. Int. Conf. on Social Informatics (SocInfo), 2017, pp. 53–64.
- [17] M. Lukasik, K. Bontcheva, T. Cohn, A. Zubiaga, M. Liakata, and R. Procter, “Gaussian processes for rumour stance classification in social media,” ACM Trans. Inf. Syst., vol. 37, no. 2, pp. 1–24, Apr. 2019.
- [18] M. Lukasik, T. Cohn, and K. Bontcheva, “Classifying tweet level judgements of rumors in social media,” in Proc. Conf. on Empirical Methods in Natural Lang. Process. (EMNLP), Sep. 2015, pp. 2590–2595.
- [19] M. Lukasik, P. K. Srijith, D. Vu, K. Bontcheva, A. Zubiaga, and T. Cohn, “Hawkes processes for continuous time sequence classification: an application to rumour stance classification in Twitter,” in Proc. Annual Meeting of the Assoc. for Comput. Linguistics (Volume 2: Short Papers), Aug. 2016, pp. 393–398.
- [20] S. Dungs, A. Aker, N. Fuhr, and K. Bontcheva, “Can rumour stance alone predict veracity?” in Proc. Int. Conf. on Comput. Linguistics (COLING), Aug. 2018, pp. 3360–3370.
- [21] G. Giasemidis, N. Kaplis, I. Agraftotis, and J. R. C. Nurse, “A semisupervised approach to message stance classification,” IEEE Trans. Knowl. Data Eng., vol. 32, no. 1, pp. 1–11, Jan. 2020.
- [22] A. Zubiaga, E. Kochkina, M. Liakata, R. Procter, and M. Lukasik, “Stance classification in rumours as a sequential task exploiting the tree structure of social media conversations,” in Proc. Int. Conf. on Comput. Linguistics: Technical Papers (COLING), Dec. 2016, pp. 2438–2448.

- [23] A. Zubiaga et al., “Discourse-aware rumour stance classification in social media using sequential classifiers,” *Inform. Process. Manag.*, vol. 54, no. 2, pp. 273–290, Mar. 2018.
- [24] L. Poddar, W. Hsu, M. L. Lee, and S. Subramaniam, “Predicting stances in Twitter conversations for detecting veracity of rumors: A neural approach,” in *Proc. IEEE Int. Conf. on Tools with Artif. Intell. (ICTAI)*, Dec. 2018, pp. 65–72.
- [25] J. Ma, W. Gao, and K.-F. Wong, “Rumor detection on Twitter with treestructured recursive neural networks,” in *Proc. Annual Meeting of the Assoc. for Comput. Linguistics (Volume 1: Long Papers)*, Jul. 2018, pp. 1980–1989.
- [26] N. Bai, Z. Wang, and F. Meng, “A stochastic attention CNN model for rumor stance classification,” *IEEE Access*, vol. 8, pp. 80 771–80 778, 2020.
- [27] A. Kumar and M. Upadhyay, “Rumor stance classification using a hybrid of capsule network and multi-layer perceptron,” *Turkish J. of Comput. and Math. Educ. (TURCOMAT)*, vol. 12, no. 13, pp. 4110–4120, 2021.v
- [28] C. Scarton and Y. Li, “Cross-lingual rumour stance classification: a first study with BERT and machine translation,” in *Proc. Truth and Trust Online Conf. (TTO)*, Oct. 2021, pp. 50–59.
- [29] L. Tian, X. Zhang, Y. Wang, and H. Liu, “Early detection of rumours on Twitter via stance transfer learning,” in *Proc. Eur. Conf. on Information Retrieval (ECIR)*, vol. 12035, Apr. 2020, pp. 575–588.
- [30] J. Ma, W. Gao, and K.-F. Wong, “Detect rumors in microblog posts using propagation structure via kernel learning,” in *Proc. Annual Meeting of the Assoc. for Comput. Linguistics (Volume 1: Long Papers)*, Jul. 2017, pp. 708–717.
- [31] P. Wei, N. Xu, and W. Mao, “Modeling conversation structure and temporal dynamics for jointly predicting rumor stance and veracity,” in *Proc. Conf. on Empirical Methods in Natural Lang. Process. and the Int. Joint Conf. on Natural Lang. Process. (EMNLP-IJCNLP)*, Nov. 2019, pp. 4787–4798.
- [32] E. Kochkina, M. Liakata, and A. Zubiaga, “All-in-one: Multi-task learning for rumour verification,” in *Proc. Int. Conf. on Comput. Linguistics (COLING)*, Aug. 2018, pp. 3402–3413.
- [33] A. Khandelwal, “Fine-tune longformer for jointly predicting rumor stance and veracity,” in *Proc. ACM IKDD CODS and COMAD (CODS COMAD)*, Dec. 2021, pp. 10–19.
- [34] L. Derczynski and K. Bontcheva, “Pheme: Veracity in digital social networks,” in *UMAP Workshops*, 2014.

-
- [35] L. Alekya, L. Lakshmi, G. Susmitha, and S.Hemanth. 2020. "A Survey On Fake News Detection in Social Media Using Deep Neural Networks". *International Journal of Scientific & Technology Research*.
- [36] Alvaro Ibrain Rodriguez and Lara Lioret Iglesias. 2019. "Fake News Detection Using Deep Learning". *University of Cantabria*.
- [37] Md Zahangir Alom, Tarek M. Taha, Chris Yakopcic, et al. 2019. "A State-of-the-Art Survey on Deep Learning Theory and Architectures". *MDPI*.
- [38] Xu Han, Zhengyan Zhang, Ning Ding, et al. 2021. "Pre-Trained Models: Past, Present and Future". *Science Direct*.
- [39] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". *Semantic Scholar*.
- [40] J. Yu, J. Jiang, L. M. S. Khoo, H. L. Chieu, and R. Xia, "Coupled hierarchical transformer for stance-aware rumor verification in social media conversations," in Proc. Conf. on Empirical Methods in Natural Lang. Process. (EMNLP), Nov. 2020, pp. 1392–1401.