



دانشکده مهندسی کامپیوتر

استخراج زیرادعا به روش فرایادگیری

پروژه کارشناسی مهندسی کامپیوتر

زهرا حسینی

اساتید راهنما

دکتر سید صالح اعتمادی و دکتر احمد اکبری اذیرانی

تیر ۱۴۰۱



تأییدیه‌ی هیأت داوران جلسه‌ی دفاع از پروژه

نام دانشکده: دانشکده مهندسی کامپیوتر

نام دانشجو: زهرا حسینی

عنوان پروژه: استخراج زیرادعا به روش فرایادگیری

تاریخ دفاع: تیر ۱۴۰۱

رشته: مهندسی کامپیوتر

ردیف	سمت	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضا
۱	استاد راهنما	دکتر سید صالح اعتمادی	استادیار	مؤسسه دانشگاه علم و صنعت ایران	

تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب زهرا حسینی به شماره دانشجویی ۹۶۵۳۱۲۲۶ دانشجوی رشته مهندسی کامپیوتر مقطع تحصیلی کارشناسی تأیید می‌نمایم که کلیه‌ی نتایج این پروژه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه‌برداری شده 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 در خصوص 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: زهرا حسینی

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.

اساتید راهنما: دکتر سید صالح اعتمادی

دکتر احمد اکبری ازیرانی

تاریخ:

امضا:

تقدیم به:

این اثر، هرچند اندک، تقدیم می‌شود به پانیز سلطانی، دانش‌آموخته‌ی دانشکده مواد و متالورژی دانشگاه علم و صنعت، و دیگر جان باختگان پرواز PS752 اوکراین.
یاد و نامشان جاودانه، راهشان سبز باد.

قدردانی

از جناب آقای دکتر اعتمادی استاد راهنمای این پایان نامه نهایت تشکر را بابت تمام راهنمایی ها و زحماتشان در طول دوره ی کارشناسی اینجانب را دارم. از جناب آقای دکتر ناصر مزینی که زحمت داوری این پایان نامه را داشتند نیز تشکر می نمایم. لازم می دانم که از زحمات بی دریغ مهندس مجید زرهرن تشکر ویژه بنمایم. همچنین از همکاری های آقایان بابک بهکام کیا، آرمین غلامپور و خانم مهسا قادران سپاسگزارم. در پایان از خانواده عزیزم که در تمام مدت تحصیل بنده، مشوق و یاور بنده بودند تشکر میکنم.

زهره حسینی

تیر ۱۴۰۱

چکیده

امروزه افزایش منابع اینترنتی غیر اصولی منجر به انتشار گسترده اخبار جعلی شده است. عدم شناسایی سریع اطلاعات نادرست می تواند صدمات جبران ناپذیر قابل توجهی را ایجاد کند. در نتیجه، بسیاری از محققان بر روی شناسایی اخبار جعلی کار کرده اند. طبقه بندی خودکار یک ادعا به عنوان اطلاعات درست یا اطلاعات نادرست یک کار چالش برانگیز است. حتی یک متخصص در یک حوزه خاص باید جنبه های متعددی را قبل از صدور حکم در مورد صحت یک ادعا بررسی کند. در این پژوهش، ما در ابتدا به بررسی کارهای پژوهشی انجام شده در زمینه ی تشخیص اخبار جعلی و مقایسه فرآیندهای آن ها پرداختیم. در پژوهش های قبلی اگر یک ادعا، ترکیبی از ادعاهای غلط و درست بود، ادعای کلی را به عنوان ادعای غلط برچسب میزدند. در این پژوهش، برای بهبود کیفیت کار سامانه های پیشین، بر روی ایده ی استخراج ادعا، از یک ادعای اصلی تحقیقاتی انجام دادیم. با استخراج زیر ادعاهای یک ادعا، هر زیر ادعا به صورت مجزا میتواند صحت سنجی شود. برای انجام آزمایش ها و ارزیابی ها، مجموعه داده ایی جمع آوری کردیم و بر روی مدل T5 و GPT-NEO با روش یادگیری چند داده ¹ خروجی گرفتیم و بر روی خروجی ها ارزیابی انجام دادیم.

واژگان کلیدی: هوش مصنوعی، پردازش زبان های طبیعی، تشخیص اخبار جعلی، استخراج حقیقت، مدل های زبانی، بررسی حقیقت، تشخیص موضع، فرا یادگیری

¹Few-shot learning

فهرست مطالب

خ فهرست تصاویر

د فهرست جداول

۱ فصل ۱: مقدمه

۱-۱ پیشگفتار

۱-۲ بیان مسئله

۱-۲-۱ تشخیص اخبار جعلی

۱-۲-۲ استخراج حقیقت و بررسی حقیقت

۱-۲-۳ تشخیص موضع

۱-۲-۴ تشخیص اخبار جعلی قابل توضیح

۱-۲-۵ استخراج ادعا

۱-۳ ترتیب نگارش گزارش

۶ فصل ۲: مروری بر منابع

۲-۱ مقدمه

۲-۲ تشخیص اخبار جعلی

۲-۳ تشخیص موضع

۲-۴ تشخیص اخبار جعلی توضیحی

۲-۵ استخراج ادعا

۱۰	فصل ۳: پیاده‌سازی
۱۱	۱-۳ مقدمه
۱۱	۲-۳ جمع‌آوری داده
۱۴	۳-۲-۱ دستور العمل جمع‌آوری داده در زبان انگلیسی
۱۷	۳-۲-۲ دستور العمل جمع‌آوری داده در زبان فارسی
۲۱	۳-۳ نحوه پیاده‌سازی
۲۱	۳-۳-۱ مدل ABCD
۲۲	۳-۳-۲ مدل T5
۲۳	۳-۳-۳ مدل GPT-NEO
۲۴	فصل ۴: بازیابی عملکرد
۲۵	۴-۱ مقدمه
۲۵	۴-۲ نتایج و ارزیابی
۲۶	۴-۲-۱ خروجی‌های مدل ABCD
۲۷	۴-۲-۲ خروجی‌های مدل T5
۲۸	۴-۲-۳ خروجی‌های مدل GPT-Neo
۳۰	فصل ۵: جمع‌بندی و پیشنهادها
۳۱	۵-۱ مقدمه
۳۱	۵-۲ جمع‌بندی
۳۱	۵-۳ نوآوری
۳۲	۵-۴ پیشنهادها
۳۳	مراجع
۳۶	واژه‌نامه فارسی به انگلیسی
۳۷	واژه‌نامه انگلیسی به فارسی

فهرست تصاویر

۸	۲-۱ تصویری از مفهوم تشخیص موضع
۲۲	۳-۱ بررسی عملکرد سامانه ABCD
۲۲	۳-۲ نمونه هایی از ورودی ها و خروجی های مربوطه از مدل ۵T
۲۶	۴-۱ نحوه محاسبه امتیاز bleu

فهرست جداول

۱۲	۳-۱ توزیع ادعاهای بررسی شده در مجموعه داده های HOVER و FEVER
۱۳	۳-۲ درصد موافقت ادعاهای استخراجی بین دو نفر در زبان انگلیسی
۱۳	۳-۳ درصد موافقت ادعاهای استخراجی بین دو نفر در زبان فارسی
۲۶	۴-۱ بررسی خروجی مدل ABCD
۲۷	۴-۲ امتیاز های bleu برای خروجی مدل ABCD
۲۷	۴-۳ بررسی خروجی مدل T5
۲۸	۴-۴ امتیاز های bleu برای خروجی مدل T5
۲۸	۴-۵ بررسی خروجی مدل GPT-Neo
۲۸	۴-۶ امتیاز های bleu برای خروجی مدل GPT-Neo

فصل ۱

مقدمه

۱-۱ پیشگفتار

تشخیص خودکار اخبار جعلی در بیان دیگر ارزیابی صحت ادعاها در اخبار است. این یک مسئله جدید اما حیاتی پردازش زبان های طبیعی^۱ است زیرا هم رسانه های خبری سنتی و هم رسانه های اجتماعی تأثیرات اجتماعی-سیاسی عظیمی بر هر فرد در جامعه دارند. به عنوان مثال، قرار گرفتن در معرض اخبار جعلی می تواند باعث نگرش نادرست و بدبینی نسبت به نامزدهای سیاسی شود. اخبار جعلی حتی به رویدادهای خشونت آمیز در دنیای واقعی که امنیت عمومی را تهدید می کند، مربوط می شوند.

تشخیص اخبار جعلی یک برنامه مهم در جهان است که پردازش زبان های طبیعی می تواند به آن کمک کند، زیرا همچنین تأثیرات گسترده تری بر چگونگی تسهیل صحت ادعاها و آموزش و کمک به عموم مردم برای تشخیص صحت اخبار ایجاد می کند.

۲-۱ بیان مسئله

تشخیص خودکار اخبار جعلی، به کاربر در صرفه جویی زمان و تشخیص راحت تر صحت اخبار کمک میکند. این مسئله جزو مسائل جدید در حوزه پردازش زبان های طبیعی به شمار می رود. در این پژوهش، زمانی برای مطالعه مسائل موجود در حوزه تشخیص اخبار جعلی و روش های حل آن ها صرف شد. از جمله از این مسائل می توان به موارد زیر اشاره کرد:

- تشخیص اخبار جعلی^۲
- استخراج حقیقت^۳
- بررسی حقیقت^۴
- تشخیص موضع^۵
- تشخیص اخبار جعلی قابل توضیح^۶

¹ Natural Language Processing

² Fake News Detection

³ Fact Extraction

⁴ Fact Checking

⁵ Stance Detection

⁶ Explainable Fake News Detection

در ادامه به توضیح مختصری برای هریک در این فصل بنده میکنیم.

۱-۲-۱ تشخیص اخبار جعلی

تشخیص اخبار جعلی، طبقه بندی اخبار جعلی از متن و سایر اطلاعات است. یعنی متن کامل یک خبر به عنوان ورودی داده میشود، خروجی مربوطه حاصل میشود. البته نوع و ورودی و خروجی قابل تعریف هستند. در برخی از پژوهش ها، فقط عنوان خبر به مدل داده میشود برای تولید خروجی.

۲-۲-۱ استخراج حقیقت و بررسی حقیقت

در موارد ذکر شده در بالا، دو مورد اول در بیشتر پژوهش ها در کنار یکدیگر ظاهر میشوند، لذا بهتر است برای تعریف نیست در کنار هم تعریف شده تا کاربرد آن ها مشخص است. به مجموع این دو وظیفه در واقع، Fact Extraction and Verification و یا به اختصار FEVER می گویند. در تعریف کلی به استخراج حقیقت های مرتبط با ادعای ورودی از مجموعه مستندات و بررسی ارتباط ادعا با مستندات را FEVER میگویند. هدف استخراج آن شناسایی صحت یک ادعای معین است. این وظیفه شامل وظایف فرعی بازیابی اسناد (و جملات) مربوطه از ویکی پدیا و تأیید اینکه آیا اطلاعات موجود در اسناد یک ادعای معین را تأیید یا رد می کند، است. این کار ضروری است و می تواند بلوک سازنده برنامه هایی مانند تشخیص اخبار جعلی باشد. در این سامانه، برخلاف سامانه جامع تر تشخیص اخبار جعلی، یک ادعا وارد مدل میشود.

۳-۲-۱ تشخیص موضع

تشخیص موضع یک متن نسبت به یک ادعا، اولین و مهم ترین قدم در تشخیص اخبار جعلی است [۱۲]. شناسایی دیدگاه نسبی یک متن با توجه به یک ادعای معین تحت دسته بندی هایی مانند موافق، مخالف، بحث و گفتگو و غیر مرتبط و یا به فرآیند خودکار فهمیدن اینکه سازمان های خبری درباره ی یک ادعا چه میگویند، تشخیص موضع گفته میشود.

۱-۲-۴ تشخیص اخبار جعلی قابل توضیح

با وجود موفقیت روش‌های تشخیص اخبار جعلی مبتنی بر یادگیری عمیق، اکثر این روش‌ها بر روی تشخیص موثر اخبار جعلی با ویژگی‌های پنهان تمرکز می‌کنند، اما نمی‌توانند توضیح دهند که چرا یک خبر به عنوان خبر جعلی تشخیص داده شد. توانایی توضیح اینکه چرا اخبار جعلی تشخیص داده شده اند بسیار کارآمد است.

۱-۲-۵ استخراج ادعا

بعد از انجام مطالعات در حوزه ی، تشخیص اخبار جعلی قابل توضیح، به ایده ی استخراج ادعا رسیدیم. در سامانه ها تشخیص اخبار جعلی موجود، یک ادعا به عنوان ورودی گرفته میشود و یک خروجی قطعی دارد. به این معنی که ادعای وارد شده یا صحیح است، غلط است و یا اطلاعات کافی برای تشخیص نداریم. در این حالت یک مشکل وجود دارد، اینکه ادعا ممکن است دارای بخش های غلط و هم بخش های صحیح باشد. ما به دنبال توسعه سامانه ایی هستیم که، در متن ورودی یک ادعا مشخص کند که کدام بخش های ادعا صحیح و کدام بخش ها غلط هستند و از خروجی قطعی بپرهیزیم. اولین گام برای دست یابی به این هدف، جمع آوری مجموعه داده و بررسی داده ها بود. سپس با توجه به صورت مسئله، از مدل های زبانی استفاده کردیم تا زیر ادعاهایی از ادعای اصلی تولید کنیم. این مسئله به بهبود کیفیت و عملکرد سامانه های موجود کمک چشمگیری خواهد کرد و میتواند به یک ابزار روزمره تبدیل شود. به عنوان نمونه ایی از ورودی و خروجی میتوان به مثال زیر اشاره کرد:

ورودی:

Cristiano Ronaldo, who has scored over 800 senior career goals for club and country and has played for Real Madrid, has won two Ballon d'Or awards. In 2011, he lost his best friend.

استخراج ادعا:

1. Cristiano Ronaldo has scored over 800 senior career goals for club and country. (Support)
2. Cristiano Ronaldo has played for Real Madrid. (Support)
3. Cristiano Ronaldo has won two Ballon d'Or awards. (Refute)
4. In 2011, Cristiano Ronaldo lost his best friend. (Not Enough Info)

خروجی:

Cristiano Ronaldo, who has scored over 800 senior career goals for club and country and has played for Real Madrid, has won two Ballon d'Or awards. In 2011, he lost his best friend.

۱-۳ ترتیب نگارش گزارش

در این گزارش، در فصل دوم مروری بر ادبیات موضوع انجام میشود. گام ها و پژوهش هایی که ما را در انتخاب موضوع اصلیمان یعنی استخراج ادعا، هدایت کردند را بررسی میکنیم. اینکه در هرکدام مسئله چه بوده است و به اختصار به راه حل های موجود هرکدام اشاره خواهیم کرد.

در فصل سوم، توضیحاتی در رابطه با مسئله مورد بحث داده میشود، در کنار روند جمع آوری داده مورد نیاز و روش های برچسب زدن داده. همچنین به توضیح فرآیند خروجی گرفتن و پیاده سازی های انجام شده، خواهیم پرداخت.

در فصل چهارم، نتایج حاصل از خروجی ها را بررسی خواهیم کرد. اینکه در پیاده سازی ها و شرایط مختلف آزمایشی چه خروجی هایی دریافت کردیم. همچنین مدل های استفاده شده از لحاظ مختلف مقایسه و بررسی خواهند شد.

در فصل پنجم، که فصل انتهایی نیز است، به چالش ها و کارهای آینده این موضوع اشاراتی خواهیم کرد. این موضوع هم در زمینه جمع آوری داده و هم در زمینه ی ارزیابی مدل ها دارای چالش های بسیاری است که به جزئیات در این فصل بحث خواهند شد.

فصل ۲

مروری بر منابع

۲-۱ مقدمه

ایده مطرح شده در این پژوهش بر پایه پژوهش های انجام شده در تشخیص اخبار جعلی، استخراج حقیقت و بررسی آن است. پس از گسترش مدل های یادگیری عمیق، وظایف متنوعی برای حوزه های مختلف هوش مصنوعی تعریف شد. یکی از آنها تشخیص اخبار جعلی بود. در ادامه به مرور مختصری میپردازیم [۳].

۲-۲ تشخیص اخبار جعلی

تشخیص اخبار جعلی به دلیل انتشار گسترده اخبار جعلی از طریق بستر های رسانه ای اهمیت فزاینده ای در میان جامعه پژوهشی پیدا کرده است [۵]. روش های متنوعی برای تشخیص اخبار جعلی وجود دارد که به دو دسته کلی زیر تقسیم میشوند.

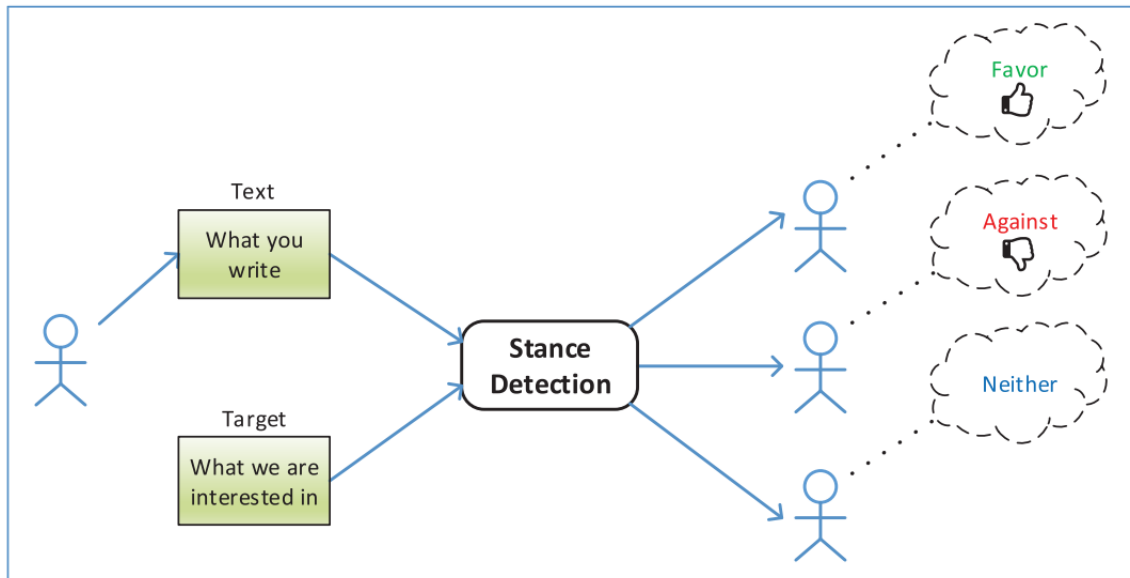
- بر اساس محتوا: هدف روش تشخیص اخبار جعلی مبتنی بر محتوا، تشخیص اخبار جعلی با تجزیه و تحلیل محتوای مقاله، متن یا تصویر یا هر دو در داخل مقاله خبری است [۲۰].

- بر اساس بافت اجتماعی: سه جنبه عمده از بافت اجتماعی وجود دارد [۴]: نمایه های کاربر، پست ها و پاسخ های کاربر، و ساختار شبکه تشخیص اخبار جعلی مبتنی بر شبکه، شبکه های اجتماعی مختلف مانند دوستی، توییت-ریتوییت، شبکه های را برای شناسایی اخبار جعلی مطالعه می کند [۱۶].

چندین مجموعه داده برای تشخیص اخبار جعلی وجود دارد که پیش بینی می کند آیا کل مقاله درست یا جعلی است.

۲-۳ تشخیص موضع

تشخیص موضع وظیفه ارزیابی اینکه نویسنده در چه سمتی از بحث از متن قرار دارد. این با تشخیص اخبار جعلی متفاوت است زیرا برای صحت نیست بلکه برای ثبات است [۷، ۲]. تشخیص موضع می تواند یکی از وظایف فرعی تشخیص اخبار جعلی باشد، زیرا می تواند برای جستجوی اسناد برای شواهد اعمال شود. PHEME [۲۱]، یکی از مجموعه داده های اخبار جعلی، توییت هایی مرتبط با اخبار دارد که رفتار کاربرانی را که به آنها اعتماد یا غیر قابل اعتماد هستند، نشان می دهد.



شکل ۲-۱: تصویری از مفهوم تشخیص موضع

[۱۰]

۴-۲ تشخیص اخبار جعلی توضیحی

هدف این پژوهش ها این است که به این سوال پاسخ دهند، چرا یک خبر خاص جعلی تشخیص داده می شود؟ [۱۵] برای اعتماد به یک سامانه تشخیص، ابتدا باید آن را درک کنیم. سناریوی ایده آل نمایش عوامل واقعی است که در نتیجه نهایی نقش داشته اند. با این حال، زمانی که از مدل های یادگیری عمیق استفاده می شود، بسیار چالش برانگیز است. این مدل ها از میلیون ها محاسبات ساده تشکیل شده اند، بنابراین مگر اینکه بتوان آنها را به روشی قابل تفسیر توسط انسان توصیف کرد، به عنوان جعبه سیاه در نظر گرفته می شوند. همچنین در حوزه ی پزشکی و سلامت پژوهشی انجام شده بود که به عنوان توضیح خلاصه ای از مستند مربوط به ادعا را خروجی میداد [۹].

۲-۵ استخراج ادعا

در زمینه استخراج ادعا کمتر پژوهشی انجام شده است. در پژوهشی، بررسی شده بود که کدام یک از جملات مقالات علمی چاپ شده ادعا هستند و کدام یک نیستند. برای این عمل، مجموعه داده ایی جمع آوری شده بود و آزمایش هایی انجام داده بودند [۱]. در پژوهش دیگری نیز به کمک انتقال یادگیری هدف مشترکی را دنبال کرده بوند تا سامانه های تشخیص اخبار جعلی را بهبود بدهند. به اینصورت که جملات مقاله ورودی را یک به یک بررسی میکردند که ادعا هستند یا خیر [۱۳]. خروجی های هر دوی این مقاله ها به صورت دودویی است، یعنی صرفاً مشخص میشود که یک جمله، آیا ادعا هست یا نیست؟ در حالی که در پژوهش ما، هدف تولید کردن زیرادعا از یک ادعای ورودی است و به بیان دیگر، تبدیل یک ادعای پیچیده به مجموعه ایی از زیرادعاهایش.

فصل ۳

پیاده‌سازی

۳-۱ مقدمه

برای استخراج ادعا، به دنبال روش های پیاده سازی بودیم. این وظیفه، از آنجایی که هم به نوعی تولید زبان طبیعی است و هم استخراج اطلاعات از یک جمله، وظیفه پیچیده ایی به شمار می‌رود. برای دست یابی به این هدف ما ابتدا به بررسی مجموعه داده ی موجود پرداختیم. دو مجموعه داده ی FEVER [۱۷] و HOVER [۸] را نهایتاً انتخاب کردیم. در این دو مجموعه، سه دسته بندی وجود دارد.

در مجموعه داده FEVER سه دسته ادعا با برچسب های، SUPPORTED، REFUTED و NOTENOUGHINFO مشخص شده اند. همچنین در مجموعه داده ی HOVER برچسب ادعاها به صورت، SUPPORTED و NOT_SUPPORTED است.

با بررسی این مورد، به این نتیجه رسیدیم که عموماً ادعاهایی با برچسب NOTENOUGHINFO در مجموعه داده ی Fever و NOT_SUPPORTED در مجموعه داده ی HOVER ادعاهایی هستند که دارای چندین بخش هستند و سامانه های موجود قادر به تخصیص دادن یک برچسب واحد به آنها نیستند و آنها را با این عناوین برچسب زده اند.

به این ترتیب گام بعدی ما در این پژوهش، جمع آوری داده مناسب به وسیله ی این دو منبع بود. در ادامه به توضیح نحوه جمع آوری داده، نحوه استخراج ادعاها و دستورالعمل های تهیه شده برای استخراج ادعاها می‌پردازیم.

۳-۲ جمع آوری داده

در این مرحله ^۱، مجموعه داده FEVEROUS^۲ را که نسخه ی ۲۰۲۱ مجموعه داده FEVER است را دریافت کردیم که دارای ۷۱۲۹۱ ادعا است همچنین مجموعه داده HOVER^۳ را نیز دریافت کردیم. به صورت رندوم ۷۵ عدد از ادعاهای NOTENOUGHINFO در FEVER و ۷۵ ادعای دیگر از NOT_SUPPORTED های مجموعه داده HOVER را انتخاب کردیم.

^۱ از زحمات آقای بابک بهکام کیا به وزه در این بخش بینهایت سپاسگزارم، بدون همکاری ایشان در مدت کم این بخش قابل انجام نبود.

^۲<https://fever.ai/dataset/feverous.html>

^۳<https://hover-nlp.github.io/>

با پایین بودن درصد موافقت بین دو نفر که ادعا ها را بررسی کردند، متوجه شدیم که چالش های بسیاری دارد استخراج ادعا توسط نیرو انسانی. چند مورد از چالش ها به موارد زیر میتوان اشاره کرد:

۱. زبان انگلیسی، زبان مادری هیچ یک از افراد نبود، لذا باعث ایجاد خطاهایی میشد که دو نفر توافق نداشتند.

۲. در برخی موارد ادعاهایی وجود داشت که نیازمند اطلاعات عمومی خاص در آن زمینه بود و استخراج ادعا بدون آن اطلاعات میسر نبود.

Label \ DataSet	Total	NOTENOUGHINFO	NOT_SUPPORTED
FEVEROUS	71291	7920	-
HOVER	18171	-	7148

جدول ۳-۱: توزیع ادعاهای بررسی شده در مجموعه داده های HOVER و FEVER

هر کدام از این ۷۵ ادعا توسط یک نفر بررسی شده اند و زیرادعاهای آنها استخراج شد. همچنین ۵۰ ادعا توسط هر دو نفر بررسی شد. در بررسی بین دو نفر موارد متناقض زیادی پیدا میشد که دو نفر نظر یکسانی نداشتند. به عنوان مثال در ادعای زیر:

Claim: Mauricio Rosales had no career goals across any season from 2012-2021, including with his current club Olimpo.

Extracted Claims by First Annotator:

1. Mauricio Rosales had no career goals across any season from 2012-2021.
2. Olimpo is an Argentine football club sited in Carlos Casares.

3. Mauricio Rosales was tthe owner of the Olimpo club.

Extracted Claims by Second Annotator:

1. Mauricio Rosales had no career goals across any season from 2012-2021.
2. Olimpo is an Argentine football club sited in Carlos Casares.
3. Mauricio Rosales had no career goals with his current club Olimpo.
4. Mauricio Rosales is a player in Olimpo.

همانطور که مشخص شده، دو نفر دارای اشتراکی هستند ولی نفر دوم معتقد بود که ادعای سومی استخراج شده توسط نفر اول در متن اولیه وجود ندارد که با رنگ قرمز مشخص شده است و به علاوه اینکه دو ادعای بیشتر نیز توسط نفر دوم از متن استخراج شده که نفر اول معتقد بوده است متن اولیه دارای این ادعاها نیست.

agreement%	-	+guideline
First Annotator	60%	88%
Second Annotator	68%	92%

جدول ۳-۲: درصد موافقت ادعاهای استخراجی بین دو نفر در زبان انگلیسی

به بررسی مجموعه داده فارسی مشابه یعنی، ParsFEVER [۱۹] پرداختیم تا تاثیر زبان مادری را در استخراج ادعاها بررسی کنیم و همچنین نقطه شروعی برای انجام پژوهشی مشابه در زبان فارسی باشد.

agreement%	-	+guideline
First Annotator	70%	90%
Second Annotator	60%	90%

جدول ۳-۳: درصد موافقت ادعاهای استخراجی بین دو نفر در زبان فارسی

۳-۲-۱ دستورالعمل جمع‌آوری داده در زبان انگلیسی

دستورالعمل تهیه شده به طور خلاصه، پس از بررسی‌های بسیار و گفت و گو بر سر هر کدام به شرح زیر است:

۱. اگر در جمله and یا or دیده شود می‌توانیم برای یک طرف ادعا ها را در آورده و همان ادعا ها را برای آن یکی طرف نیز نسبت دهیم. (باید دقت کرد در بعضی جملات and بین دو کلمه نماید، ممکن است یک طرف and شامل ۴ کلمه باشد و طرف دیگر شامل ۱ کلمه)

۲. هر وقت or ببینیم، می‌توانیم بگوییم که x is y یا x and y are same ولی این نکته همیشه درست نیست. ممکن است که کلمه or بین دو کلمه‌ای که دو چیز متفاوت را تعریف می‌کنند بیاید که این دو چیز برخی صفات مشترک دارند.

۳. اگر شغل، صفت، لقب، ملیت یا ... به کسی یا چیزی نسبت داده شود، می‌توانیم با جایگشت‌های مختلف از هر کدام از آنها ادعا‌های متفاوتی بدست بیاوریم. (هر کدام از ویژگی‌های ذکر شده اگر فقط برای کلمه مورد نظر نسبت داده شده نباشد، نمی‌تواند به تنهایی یک ادعا تشکیل دهد و باید به همراه ویژگی یا ویژگی‌هایی که به آن‌ها نسبت داده شده است، باشد).
مثال:

باراک اوباما اولین رئیس‌جمهور سیاه پوست ایرانی تبار آمریکا است.
باراک اوباما رئیس‌جمهور است / باراک اوباما سیاه پوست است / ... اما "باراک اوباما اولین است"
معنی نمی‌دهد پس کلمه "اولین" نمی‌تواند به تنهایی ادعا بسازد. پس به این شکل می‌توان ادعا ساخت:

باراک اوباما اولین رئیس‌جمهور است.

۴. اگر ویژگی‌های ذکر شده در قانون بالا برای یک شی مجهول باشد، نمی‌توانیم آن ویژگی‌ها را حذف کنیم و ادعا جدید درست کنیم. مثال:

The tallest man in the America has two houses.

در این مثال نمی‌توانیم the tallest را حذف کنیم.

۵. در اکثر اوقات ”،” نقش and یا or را دارد.

۶. اگر دو جمله با یک کلمه مبهم به هم دیگر متصل شوند، نمی‌توان آن را به جملات ساده ترتجیزه کرد (منظور ما کل دو جمله است، امکان دارد بتوانیم از هر کدام از آن جمله‌ها ادعا استخراج کنیم. مثال:

claim: The movie starring Lightning McQueen was released by the same company which used songs written by Lin-Manuel Miranda in another of its movies.

۷. اگر در جمله از حروف مالکیت استفاده شده باشد، می‌توانیم با استفاده از has/have/had یک ادعا جدید بسازیم. مثال:

Dana Fox wrote Tom Vaughan’s unpopular 2008 film.

Tom Vaughan had an unpopular 2008 film.

توجه داشته باشید که unpopular و ۲۰۰۸ صفت هستند و با توجه به قوانین بالا می‌توانیم ادعاهای بیشتری بسازیم.

۸. اگر در جمله بدل یا نکته اضافه‌تری در پرانتز داشتیم، آن‌ها را نیز به عنوان ادعا بررسی کنیم. (اگر مترادف کلمه در داخل پرانتز بود نیازی نیست که ادعاهای مجزا در نظر بگیریم.)

claim: The Iris latistyla has two bract (protective layer)

The Iris latistyla has two bract.

The Iris latistyla has two protective layer.

۹. اگر در جمله with به معنی "داشتن" بود، ادعا‌های جدید را میتوان با have-has-had با توجه به فاعل و زمان جمله اصلی، نوشت.

claim: The Iris latistyla has two bract (protective layer) and the flowers are 5 cm in diameter with two pairs of petals.

The Iris latistyla's flowers have two pairs of petals.

۱۰. اگر در جمله "Neither/nor" داشتیم، میتوان ادعا‌های منفی تولید کرد.

claim: The courtside reporter in the 2015–16 Fox Sports North was neither Marney Gellner nor Anthony LaPanta.

Marney Gellner was not The courtside reporter in the 2015–16 Fox Sports North.

Anthony LaPanta was not The courtside reporter in the 2015–16 Fox Sports North.

۱۱. اگر در جمله "either/or" داشتیم، میتوان ادعا‌های مثبت تولید کرد.

۱۲. وقتی جمله پیچیده^۴ هست میتوان آن را ساده تر کرد و ادعا‌های جدیدی استخراج کرد:

claim: the Ammon family whose story was sensationalized by news outlets.

The Ammon family's story was sensationalized by news outlets.

۱۳. وقتی از ضمائر مالکیت استفاده شده میشه جمله جدید رو با استفاده از فاعل و s' نوشت:

⁴Complex sentences <https://www.grammarly.com/blog/complex-sentence/>

claim: Pantophaea has grey forewings with a wide dark grey postmedial band and a dark diffuse spot at the costa, near the apex, and its Hindwings are white with a small dark spot at tornus.

Pantophaea's Hindwings are white with a small dark spot at tornus.

۱۴. آگه تاریخ تولد و فوت رو داخل پرانتز آورد میشه سن رو حساب کرد به عنوان ادعا ارایه داد، تاریخ تولد، تاریخ فوت را به عنوان ادعا ارائه کرد.

claim: Max Stewart (14 March 1935 – 19 March 1977) Max Stewart died when he was 42.

۳-۲-۲ دستور العمل جمع‌آوری داده در زبان فارسی

یکی از موارد بحث برانگیز زبان فارسی، افعال متعدی و لازم بودند. در ابتدا به اختصار به این مورد اشاره کردیم:

انواع فعل در زبان فارسی: فعل ناگذرا یا لازم فعلی است که معنایش به خودی خود کامل است و جز فاعل (نهاد) به کلمه دیگری احتیاج ندارد. به فعلی که برای کامل شدن معنایش علاوه بر فاعل به مفعول نیز نیاز دارد، فعل متعدی (گذرا) می‌گویند.

پس به طور کلی اگر یک ادعا را بشکنیم و اگر فعل این ادعا های فرعی متعدی باشد و مفعول نداشته باشد و یا برعکس یعنی فعل لازم باشد و دارای مفعول نیز باشد، آن ادعای فرعی به دست آمده اشتباه است.

۱. اگر در ادعا حالتی برای نهاد دیده شود، با استفاده از یک فعل اسنادی (است، بود، شد) یک ادعای فرعی می‌توان ساخت. در این حالت نهاد تبدیل به مسند الیه شده و حالت توصیفی نیز یک مسند میشود.

مثال:

غلامرضا موسوی متولد ۲۸ فروردین ۱۳۳۰ نورپرداز فیلم گل یخ است. -> غلامرضا موسوی نورپرداز است. ممکن است از یک ادعا چندین ادعای فرعی با این ساختار استخراج شود:

مثال:

جیمز ادوارد فرانکو کشاورز آمریکایی است.

جیمز ادوارد فرانکو کشاورز است.

جیمز ادوارد فرانکو آمریکایی است.

۲. اگر در ادعا تاریخ تولد فردی آمده باشد می‌توان چندین ادعای فرعی استخراج کرد، به این صورت که تاریخ تولد دارای دو بخش است یعنی سال و ماه تولد. می‌توانیم ۳ ادعای فرعی با استفاده از این اطلاعات به دست آورد.

مثال:

غلامرضا موسوی متولد فروردین ۱۳۳۰ نورپرداز فیلم گل یخ است.

غلامرضا موسوی متولد فروردین است.

غلامرضا موسوی متولد ۱۳۳۰ است.

بعضی اوقات ممکن است روز تولد نیز همراه ماه بیاید، در این صورت علاوه بر ادعای بالا می‌توان ادعای زیر را نیز استخراج کرد.

مثال:

غلامرضا موسوی متولد ۲۸ فروردین ۱۳۳۰ نورپرداز فیلم گل یخ است.

غلامرضا موسوی متولد ۲۸ فروردین است.

۳. اگر در ادعا به "و" ببینیم، به راحتی می‌توانیم دو ادعای فرعی با استفاده از طرفین این "و" استخراج کرد.

مثال:

از ویژگی‌های فیلم‌های عباس کیارستمی پس از انقلاب می‌توان به شاعرانگی و حذف کارگردان اشاره کرد.

از ویژگی‌های فیلم‌های عباس کیارستمی پس از انقلاب می‌توان به حذف کارگردان اشاره کرد.

از ویژگی‌های فیلم‌های عباس کیارستمی پس از انقلاب می‌توان به شاعرانگی اشاره کرد.

نکته: باید دقت کرد که حرف ”و“ دو اسم یا دو جمله مستقل را بهم وصل کند، در غیر اینصورت نباید جدا کنیم.

مثال: سرکه سبزیجات را حل و هضم می کند.

در این مثال حل و هضم در واقع یک کلمه مرکب است و جدا کردن آنها ادعایی با معنی متفاوت اضافه تولید نمیکند.

۴. بهترین روش ساخت یک ادعای فرعی استفاده از فعل و نهاد آن است. یعنی ابتدا نهاد و فعل را در ادعا تشخیص می دهیم و سپس اگر فعل ادعا یک فعل متعدی باشد، سریعاً مفعولش را از ادعای اصلی استخراج می کنیم و بعد اضافه کردن مفعول، به یک ادعای فرعی میرسیم.

مثال:

عبدالحمید اسماعیل زهی با درخواست یکی از چهره‌های مشهور اهل سنت ایران، معاونت مدرسه عالی دارالعلوم زاهدان را به عهده گرفت.

نهاد: عبدالحمید اسماعیل زهی ، فعل : عهده گرفت

فعل متعدی است پس باید مفعول را نیز پیدا کنیم -> مفعول : معاونت مدرسه عالی دارالعلوم زاهدان
پس در نهایت به ادعای فرعی زیر می رسیم:

عبدالحمید اسماعیل زهی معاونت مدرسه عالی دارالعلوم زاهدان را به عهده گرفت.

در ادامه اگر فعل جمله گذرا به متمم نیز باشد، می توان با استفاده از متمم آن نیز یک ادعای فرعی دیگر ساخت.

مثال بالا:

عبدالحمید اسماعیل زهی با درخواست یکی از چهره‌های مشهور اهل سنت ایران، معاونت مدرسه‌ای را به عهده گرفت.

نکته: در این مثال برای این که ادعا پیچیده تر نشود، از آوردن کل مفعول خودداری شده است.

۵. برای قانون بالا بعضی اوقات معنی فعل فقط با مفعول کامل نمیشود و به متمم نیز نیاز است.

مثال:

سرکشی جرج ششم به مناطق آسیب‌دیده تحسین مردم آلمان را برانگیخت.

بدون متمم: جرج ششم سرکشی کرد.

با متمم: جرج ششم به مناطق آسیب‌دیده سرکشی کرد.

۶. یک سری فعل‌ها را می‌تواند نشانی از یک ادعا در نظر گرفت. به معانی فعل‌ها باید دقت کنیم.

مثال:

ادعای اصلی: کوین مایکل کاستنر در فیلم یک دنیای بی نقص بازی کرده است.

ادعای فرعی: کوین مایکل بازیگر است.

۷. در صورت نام بردن تعدادی از عناوین می‌توان بر روی تعداد و شمارش ادعای جدید ساخت.

مثال:

ادعای اصلی: کلینت ایستوود جایزه‌ی اسکار سال ۲۰۰۷ بهترین کارگردان و بهترین تهیه‌کننده را برای

خود کسب کرد.

ادعای فرعی: کلینت ایستوود دو جایزه‌ی اسکار سال در ۲۰۰۷ کسب کرد.

۸. به مکان‌های ذکر شده در ادعا دقت کنیم زیرا منبع مناسبی برای ادعای جدید هستند:

مثال:

ادعای اصلی: جمشید در یکی از محله‌های قدیمی بیست و یکمین شهر پرجمعیت ایران به دنیا آمد.

ادعای فرعی: جمشید ایرانی است.

۹. صفت‌هایی که معنی جمله را از بین نمی‌برند را حذف کنیم برای تولید ادعای جدید.

مثال:

ادعای اصلی: محمدحسین مهدویان در بخش غربی شهر بهار نارنج متولد شد.

ادعای فرعی: محمدحسین مهدویان در شهر بهار نارنج متولد شد.

۱۰. هیچ ادعای فرعی نباید زیر مجموعه یک ادعای فرعی دیگری باشد.

مثال:

فیلم آدم برفی بعد از سه سال توقیف در مهر ماه سال ۷۶ اکران شد.

– فیلم آدم برفی در مهر ماه سال ۷۶ اکران شد.

– فیلم آدم برفی در سال ۷۶ اکران شد.

در این حالت ادعای فرعی اول را به عنوان یک ادعای فرعی قبول نمی‌کنیم.

۳-۳ نحوه پیاده‌سازی

با توجه به اینکه تولید مجموعه داده‌ای بزرگ به صورت خودکار امکان‌پذیر نبود در مورد پروژه ما، و همچنین استخراج ادعاها توسط نیروهای انسانی کار بسیار زمانبری بود از روش Meta-Learning [۱۸] برای آموزش مدل و گرفتن خروجی استفاده کردیم. این نوع آموزش‌ها مناسب زمان‌هایی که تعداد داده کمی در دسترس داریم.

در این پژوهش به طور خاص از Few-Shot Learning استفاده کردیم. یک مسئله فرا یادگیری، آموزش یک مدل یادگیری عمیق با استفاده از چند مثال آموزشی است، با توجه به تجربه قبلی با وظایف بسیار مشابه که مجموعه‌های آموزشی بزرگی برای آنها در دسترس داریم. برای این مرحله از دو مدل استفاده کردیم که به شرح زیر هستند:

۱-۳-۳ مدل ABCD

وظیفه در نظر گرفته شده، یعنی استخراج ادعا در بیان دیگر یک وظیفه‌ی استخراج جملات و تبدیل جمله ساده به جملات کوچکتر است. به همین منظور از پیاده‌سازی یک مقاله که هدف آن تبدیل جملات پیچیده به مجموعه‌ای از جملات کوچک تشکیل دهنده جمله اولیه بودند، استفاده کردیم. [۶]

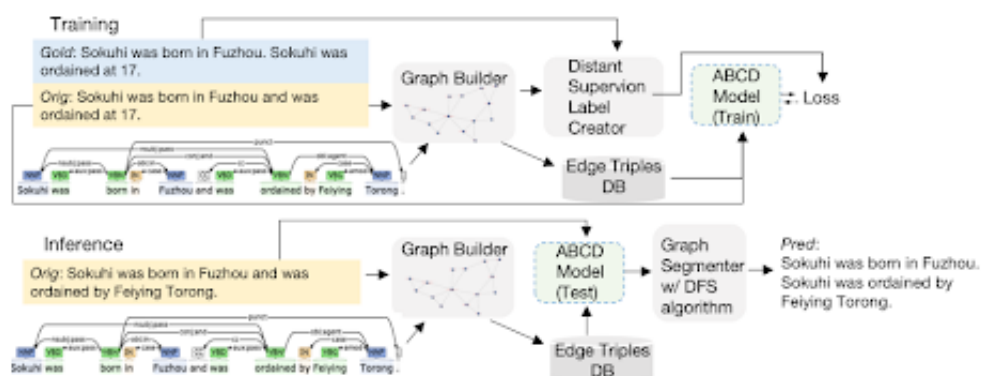


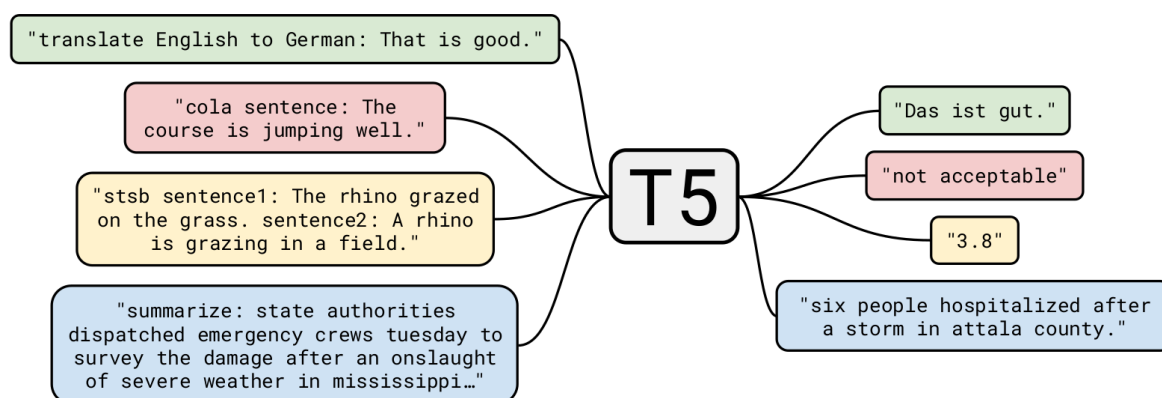
Figure 4: ABCD system overview during training (top) and inference (bottom).

شکل ۳-۱: بررسی عملکرد سامانه ABCD

[۶]

۳-۲-۳ مدل T5

برخلاف BERT که فقط بلوک‌های رمزگذار داشت و GPT-2 که فقط بلوک‌های رمزگشا داشت، T5 از هر دو استفاده می‌کند.



شکل ۳-۲: نمونه‌هایی از ورودی‌ها و خروجی‌های مربوطه از مدل T5

[۱۴]

T5 یک مدل شبکه عصبی جدید بسیار بزرگ است که بر روی ترکیبی از متن بدون برچسب و داده های برچسب گذاری شده از وظایف پردازش زبان طبیعی محبوب آموزش داده شده است، سپس به صورت جداگانه برای هر یک از آنها تنظیم^۵ شده است. وظایفی که هدف نویسندگان حل کردن آنها است. این مدل بسیار خوب کار می کند، و بهترین نتیجه را در بسیاری از برجسته ترین وظایف طبقه بندی متن برای زبان انگلیسی، پاسخگویی به سؤال و خلاصه سازی نیز نتیجه عالی دارد.

۳-۳-۳ مدل GPT-NEO

از مدل های محبوب تولید زبان طبیعی میتوان به GPT-3 اشاره کرد که تعداد پارامتر زیادی نیز دارد، ولی این مدل به صورت رایگان قابل دسترس نبود. لذا به عنوان جایگزین از مدل GPT-Neo پیاده سازی شبیه سازی مدل GPT-3 است استفاده کردم. البته این شبیه سازی نیز به طور کامل در دسترس نیست و فقط میتوان از رابط کاربری آن استفاده کرد.

GPT-Neo پیاده سازی مدل و مدل های زبان اتورگرسو موازی داده است که از Mesh Tensorflow برای محاسبات توزیع شده روی TPU ها استفاده می کند.

⁵Fine-tune

فصل ۴

بازبینی عملکرد

۴-۱ مقدمه

در این فصل به بررسی نتایج و خروجی های به دست آمده میپردازیم. از هر سه روش نامبرده،، یعنی ABCD، T5 و GPT-Neo خروجی هایی حاصل شد که به ترتیب به ذکر و بررسی چند مورد خواهیم پرداخت. همچنین روش های ارزیابی را ذکر خواهیم کرد.

۴-۲ نتایج و ارزیابی

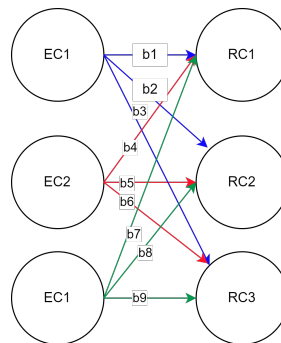
ارزیابی خروجی های این وظیفه جزو چالش های اصلی است. به عنوان مثال ما از معیار bleu [۱۱] استفاده کردیم. برای محاسبه ی این معیار از رابطه ی زیر باید تبعیت کرد:

$$\text{score bleu} = \exp \left(\frac{1}{n} \sum_{k=1}^n p_k \right) \quad (۴-۱)$$

که p_n امتیاز bleu تنها بر اساس n-gram است و به صورت زیر تعریف میشود.

$$p_n = \frac{\sum_{\text{n-gram} \in \hat{y}} \text{count}_{\text{clip}}(\text{n-gram})}{\sum_{\text{n-gram} \in \hat{y}} \text{count}(\text{n-gram})} \quad (۴-۲)$$

با توجه به اینکه در این وظیفه خروجی، چندین جمله است و نمی توان برای هریک از آنها یک جمله منبع معین در نظر گرفت، در زمینه ی محاسبه ی امتیاز bleu با چالش رو به رو میشویم. برای ارائه راه حلی جایگزین، به جای اینکه در bleu هر جمله خروجی با یک جمله منبع مقایسه شود، با تمام جملات مقایسه میشود. این عمل را برای تمام جملات خروجی انجام میدهم. سپس به ترتیب جفت جملاتی که بیشتر امتیاز را دارند، حذف میکنیم و باهم در نظر میگیریم.



شکل ۴-۱: نحوه محاسبه امتیاز bleu

با توجه به شکل ۴-۱ اگر EC را نماد ادعای استخراج شده^۱ و RC^۲ را ادعای منبع در نظر بگیریم و b_1 تا b_9 را امتیازهای بین جفت جمله‌ها در نظر میگیریم. به عنوان مثال اگر EC1 بیشترین امتیاز را با RC2 داشته باشد یعنی b_2 از تمام امتیازهای دیگر بزرگتر باشد، EC1 و RC2 حذف شده و به بررسی سایر ادعاها میپردازیم.

۴-۲-۱ خروجی‌های مدل ABCD

Input Claim: A Distributed Bragg Reflector laser has diffraction grating on one or both end mirrors, which reflect a narrow band of light back in the cavity which help them be more spectrally stable.	
Annotator's Extracted Claim	ABCD output
RC1: A Distributed Bragg Reflector laser has diffraction grating on one or both end mirrors.	EC1: A Distributed Bragg Reflector laser has diffraction grating on one or both end mirrors .
RC2: Distributed Bragg Reflector laser reflect a narrow band of light back in the cavity	EC2: one mirrors reflect a narrow band of light back in the cavity .
RC3: Distributed Bragg laser's reflection helps them be more spectrally stable.	EC3: cavity help them be more spectrally stable .

جدول ۴-۱: بررسی خروجی مدل ABCD

¹Extracted Claim²Reference Claim

	RC1	RC2	RC3	Corpus bleu
EC1	b1: 0.86	b2: 0.16	b3: 4.14e-124	cb1: 1
EC2	b4: 8.12e-13	b5: 0.65	b6: 0	cb2: 0.72
EC3	b7: 0	b8: 4.54e-122	b9: 0.26	cb3: 0.28

جدول ۴-۲: امتیازهای bleu برای خروجی مدل ABCD

خروجی بررسی شده، از بهترین خروجی ها بود. از مشکلات این سامانه میتوان به این اشاره کرد که در شکستن جملات فقط به حروف ربط توجه میکند و همچنین در بازسازی جملات استخراجی دچار اشکالاتی نظیر تشخیص نادرست فاعل و موضوع جمله است.

۴-۲-۲ خروجی های مدل T5

Input Claim: Walter studied at Gottingen where he received a doctorate and he worked at the University of Cologne.	
Annotator's Extracted Claim	T5 output
RC1: Walter studied at Gottingen.	EC1: Walter studied at Gottingen.
RC2: Walter received a doctorate from Gottingen.	EC2: Walter was a doctor.
RC3: Walter worked at the University of Cologne	EC3: Walter worked at the University of Cologne.

جدول ۴-۳: بررسی خروجی مدل T5

	RC1	RC2	RC3	Corpus bleu
EC1	b1: 1	b2: 9.09e-232	b3: 7.23e-231	cb1: 1
EC2	b4: 1.28e-231	b5: 9.29e-232	b6: 6.08e-232	cb2: 1.53e-231
EC3	b7: 1.12e-231	b6: 1.33e-231	b9: 1	cb3: 1

جدول ۴-۴: امتیازهای bleu برای خروجی مدل T5

از مشکلات ارزیابی bleu میتوان به این اشاره کرد که با اینکه ادعای استخراج شده دوم یعنی، Walter was a doctor. ادعای درستی است ولی کمی جمله بندی متفاوتی دارد دارای امتیاز به شدت پایانی است.

۴-۲-۳ خروجی های مدل GPT-Neo

Input Claim: Walter studied at Gottingen where he received a doctorate and he worked at the University of Cologne.	
Annotator's Extracted Claim	GPT-Neo output
RC1: Walter studied at Gottingen.	EC1: Walter studied at Gottingen.
RC2: Walter received a doctorate from Gottingen.	EC2: Walter received a doctorate from Gottingen and worked at the University of Cologne.
RC3: Walter worked at the University of Cologne	EC3: Gottingen is a university in Germany.

جدول ۴-۵: بررسی خروجی مدل GPT-Neo

	RC1	RC2	RC3	Corpus bleu
EC1	b1: 1	b2: 9.09e-232	b3: 7.23e-231	cb1: 1
EC2	b4: 1.14e-231	b5: 0.28	b6: 0.39	cb2: 0.67
EC3	b7: 0	b6: 1.16e-231	b9: 0	cb3: 1.17e-231

جدول ۴-۶: امتیازهای bleu برای خروجی مدل GPT-Neo

از مشکلات این حالت، تولید ادعایی کاملاً خارج از ادعای اصلی است. در این حالت بهتر است از روش های ارزیابی دیگر نیز در کنار روش ذکر شده استفاده کرد که در فصل بعدی اشاره خواهیم کرد.

فصل ۵

جمع‌بندی و پیشنهادها

۵-۱ مقدمه

در این پژوهش تلاش شده است، تا در کنار مروری کلی بر روی مباحث تشخیص اخبار جعلی، به بهبود کارایی سامانه‌های مربوطه پرداخت. ایده کلی، بهبود خروجی‌ها با نگاهی جزئی‌تر به ادعاهای ورودی است. همچنین به جمع‌آوری داده و روش متناسب داده‌ها برای گرفتن خروجی مناسب پرداختیم.

۵-۲ جمع‌بندی

یکی از مسائل نو و چالش برانگیز در حوزه پردازش زبان‌های طبیعی، تشخیص اخبار جعلی است. این مسئله دارای زیرمسئله‌های متعددی است که هرکدام در حال انجام پژوهش‌های نو و جدید برای بهبود عملکرد تشخیص اخبار جعلی در انواع حوزه‌ها هستند.

از موارد جدید میتوان به تشخیص اخبار جعلی توضیحی اشاره کرد. که تلاش دارید نشان دهد بر چه اساسی یک خبر جعلی یا واقعی است و تنها به خروجی قطعی بسنده نکند. در پژوهشی که انجام دادیم، هدفی مشابه داشتیم تا کاربر تجربه استفاده بهتری از سامانه‌های تشخیص اخبار جعلی داشته باشد و با نگاهی جزئی‌تر ادعاها بررسی شوند.

۵-۳ نوآوری

نوآوری اصلی ایده مطرح شده است، یعنی استخراج ادعا. در پژوهش‌های پیشین به طور خاص به استخراج ادعا و مشخص کردن جملات صحیح و غلط یک ادعای اصلی پرداخته نشده است. در این پژوهش سعی بر این بود که دستورالعملی مناسب جهت تهیه مجموعه داده مناسب برای این وظیفه، چه در زبان فارسی و چه در انگلیسی جمع‌آوری شود. همچنین برای گرفتن خروجی‌هایی غیر از ادعاهای استخراج شده توسط نیروهای انسانی از سه روش استفاده کردیم.

مدل اول ما، ABCD بود که وظیفه آن تبدیل جملات پیچیده به مجموعه‌ای از جملات ساده بود. روش دوم، استفاده از یادگیری با تعداد کمی داده بود، از دو مدل استفاده شد. مدل اول T5 بود که توسط تعدادی داده که توسط دو نیروی انسانی بررسی شده بودند، آموزش دید. همین داده‌ها را بر روی مدل GPT-Neo نیز آزمایش

کردیم. چالش این مسئله بخش ارزیابی خروجی‌ها بود که از دو روش امتیاز bleu و bleu-corpora استفاده کردیم. در روش اول، تصمیم بر آن شد که تمام جایگشت‌ها را برای امتیاز bleu حساب کرده و جفت جمله‌هایی که بیشترین امتیاز را دارند با یکدیگر در نظر بگیریم.

۴-۵ پیشنهادها

با توجه به اینکه مسئله جدید است، موضوعات زیادی برای تحقیق و توسعه در آن وجود دارد. به طور خلاصه به موارد زیر میتوان اشاره کرد:

- جمع‌آوری و استخراج داده‌های بیشتر توسط نیروی انسانی
- جمع‌آوری داده‌های بیشتر و تکرار آزمایش‌ها برای زبان فارسی
- محاسبه‌ی شباهت کسینوس بین جملات استخراجی و جملات منبع به عنوان روش ارزیابی
- استفاده از تشخیص موضع در مرحله‌ی ارزیابی یعنی بررسی کنیم که ادعای اولیه چه موضعی نسبت به ادعای استخراجی دارد. زیرا در برخی از موارد مشاهده شد که مدل استفاده شده، ادعایی خارج ادعای اصلی داده شده تولید کرده است.
- در نظر گرفتن تعداد ادعاهای استخراج شده به عنوان یکی از معیارهای ارزیابی

مراجع

- [1] Achakulvisut, T., Bhagavatula, C., Acuna, D., and Kording, K. Claim extraction in biomedical publications using deep discourse model and transfer learning. *arXiv preprint arXiv:1907.00962* (2019).
- [2] AlDayel, A., and Magdy, W. Stance detection on social media: State of the art and trends. *Information Processing & Management* 58, 4 (2021), 102597.
- [3] Canziani, A., Paszke, A., and Culurciello, E. An analysis of deep neural network models for practical applications. *arXiv preprint arXiv:1605.07678* (2016).
- [4] Dixit, A., and Kalbhor, I. Fake news detection using machine learning: Survey paper. *International Research Journal of Innovations in Engineering and Technology* 6, 4 (2022), 104.
- [5] D’Ulizia, A., Caschera, M. C., Ferri, F., and Grifoni, P. Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science* 7 (2021), e518.
- [6] Gao, Y., Huang, T.-H., and Passonneau, R. J. Abcd: A graph framework to convert complex sentences to a covering set of simple sentences. *arXiv preprint arXiv:2106.12027* (2021).
- [7] Ghosh, S., Singhanian, P., Singh, S., Rudra, K., and Ghosh, S. Stance detection in web and social media: a comparative study. in *International Conference of the Cross-Language Evaluation Forum for European Languages* (2019), Springer, pp. 75–87.
- [8] Jiang, Y., Bordia, S., Zhong, Z., Dognin, C., Singh, M., and Bansal, M. HoVer: A dataset for many-hop fact extraction and claim verification. in *Findings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2020).
- [9] Kotonya, N., and Toni, F. Explainable automated fact-checking for public health claims. *arXiv preprint arXiv:2010.09926* (2020).

- [10] Küçük, D., and Can, F. Stance detection: A survey. *ACM Computing Surveys (CSUR)* 53, 1 (2020), 1–37.
- [11] Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. Bleu: a method for automatic evaluation of machine translation. in *Proceedings of the 40th annual meeting of the Association for Computational Linguistics* (2002), pp. 311–318.
- [12] Pomerleau, D., and Rao, D. Fake news challenge. *Exploring how artificial intelligence technologies could be leveraged to combat fake news*. url: <https://www.fakenewschallenge.org/>(visited on 03/13/2020) (2017).
- [13] Prabhakar, A. A., Mohtaj, S., and Möller, S. Claim extraction from text using transfer learning. in *Proceedings of the 17th International Conference on Natural Language Processing (ICON)* (2020), pp. 297–302.
- [14] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P. J., et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.* 21, 140 (2020), 1–67.
- [15] Shu, K., Cui, L., Wang, S., Lee, D., and Liu, H. defend: Explainable fake news detection. in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (2019), pp. 395–405.
- [16] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., and Liu, H. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big data* 8, 3 (2020), 171–188.
- [17] Thorne, J., Vlachos, A., Christodoulopoulos, C., and Mittal, A. FEVER: a large-scale dataset for fact extraction and VERification. in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)* (New Orleans, Louisiana, June 2018), Association for Computational Linguistics, pp. 809–819.
- [18] Vanschoren, J. Meta-learning: A survey. *arXiv preprint arXiv:1810.03548* (2018).
- [19] Zarharan, M., Ghaderan, M., Pourdabiri, A., Sayedi, Z., Minaei-Bidgoli, B., Eetemadi, S., and Pilehvar, M. T. ParsFEVER: a dataset for Farsi fact extraction and verification. in *Proceedings of *SEM 2021: The Tenth Joint Conference on Lexical and Computational Semantics* (Online, Aug. 2021), Association for Computational Linguistics, pp. 99–104.

-
- [20] Zhou, X., Jain, A., Phoha, V. V., and Zafarani, R. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice 1*, 2 (2020), 1–25.
- [21] Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G., and Tolmie, P. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one 11*, 3 (2016), e0150989.

واژه‌نامه فارسی به انگلیسی

Measure	اندازه
Few-shot Learning	یادگیری چنده داده ایی
Meta Learning	فرا یادگیری
Natural Language Generation	تولید زبان های طبیعی
Language model	مدل زبانی
Ordered	مرتب
System	سامانه
Probabilistic	احتمالی
Valuation	ارزیابی
Project	پژوهش
Score	امتیاز
Data Set	مجموعه داده
User	کاربر
Profile	نمایه

واژه‌نامه انگلیسی به فارسی

Stance Detection	تشخیص موضع
Fact Extraction	استخراج حقیقت
Evaluation	ارزیابی
Experiment	آزمایش
Conclusion	نتیجه
Claim	ادعا
Train	آموزش
User	کاربر
Profile	نمایه
Document-retrieval	استخراج مستندات

Abstract:

Nowadays, the increase of unprincipled internet sources has led to the widespread dissemination of fake news. Failure to quickly identify misinformation can cause significant irreparable damage. As a result, many researchers have worked on identifying fake news. Automatically classifying a claim as true or false is a challenging task. Even an expert in a particular field must consider multiple aspects before passing judgment on the veracity of a claim. In previous research, if a claim was a combination of false and valid claims, the whole claim was labeled as false. In this research, to improve the work quality of previous systems, we researched the idea of claim extraction from the main claim. By extracting the sub-claims of a claim, each sub-claim can be verified separately. In this research, we first investigated the research work done in the field of fake news detection and compared their processes. To conduct tests and evaluations, we collected a data set, outputted some data on the T5 and GPT-NEO model with the few-shot learning method, and evaluated the outputs.

Keywords: Artificial intelligence, Natural Language Processing, Fake News Detection, Fact Extraction, Language models, meta-learning



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

Sub-Claim Extraction Using Meta Learning

Bachelor of Science Thesis in Computer Engineering

By:

Zahra Hosseini

Supervisors:

Prof. Sauleh Eetemadi and Prof. Ahmad Akbari Azirani

July 2022