ارائه یک روش معنایی مبتنی بر یادگیری ماشین برای شناسایی استلزامات متنی

زینب رحیمی[۱] و مهرنوش شمس‌فرد[۲]
[۱] دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی، تهران.
rahimi.zeinab@gmail.com

[۲]دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی، تهران.
m-shams@sbu.ac.ir

چکیده

بسیاری از کاربردهای پردازش زبان طبیعی مانند سیستم‌های پرسش و پاسخ، استخراج اطلاعات و خلاصه سازی متن با پدیده بیان معنای یکسان در قالب جملات متفاوت و بیان غیرمستقیم معنا در متون روبرو هستند. حوزه پژوهشی استلزامات متنی به شناسایی، استخراج یا تولید جملات قابل نتیجه گیری از متن اختصاص دارد. در این مقاله یک روش جدید برای شناسایی استلزامات متنی (RTE[۱]) معرفی می‌گردد. در این روش رابطه استلزام متنی به صورت مسئله دسته بندی دوکلاسه در نظر گرفته شده و یک سیستم یادگیری ماشین با استفاده از ویژگی‌های معنایی برای حل آن آموزش داده شده است. این ویژگی‌ها به صورت عمده از بازنمایی معنایی زوج متن و فرضیه استخراج شده اند. نتایج ارزیابی نمایانگر عملکرد قابل قبول سیستم معنایی ارائه شده است که در آن مقدار ٦٢ برای معیار f بر روی مجموعه ای ترکیبی از دادگان رقابت‌های RTE بدست آمده است. صرف نظر از این موضوع که سیستم پیشنهادی نسبت به سیستم‌های شرکت کننده در رقابت‌های RTE سال‌های مختلف عملکرد خوب و قابل قبولی ارائه نموده است، در مقایسه با سیستم پایه پیاده سازی شده که بر اساس یک معیار ترکیبی شباهت سنجی طراحی شده است، کارایی بهتری داشته است.

کلمات کلیدی

شناسایی استلزامات متنی، یادگیری ماشین، بازنمایی معنایی.

۱- مقدمه

وجود تفاوت در نحوه بیان معنا، یک مسئله متداول در پردازش زبان طبیعی است که در آن معنای یکسان می‌تواند در قالب روساخت متفاوت بیان شود و یا به عبارت دیگر از متون مختلف معنای یکسان دریافت گردد. همچنین گاهی در متون زبان طبیعی مواردی وجود دارد که به صراحت بیان نشده ولی با استفاده از دانش زمینه از متن قابل نتیجه گیری است. بنابراین برای فهم بهتر لازم است تا مواردی که مستقیما در متن بیان نشده نیز استخراج گردد. در ادبیات پردازش زبان طبیعی به پدیده نتیجه گیری یک متن از متن دیگر استلزام متنی گویند. استلزام متنی به صورت زوج متن و فرضیه (t,h)

به صورت t ⇨ h نمایش داده شده و در منابع علمی تعاریف زیر برای آن ارائه شده است:

– تعریف کلاسیک [1]: متن t فرضیه h را استلزام می‌کند اگر h به ازای تمام حالات ممکن که t درست است، درست باشد.

– تعریف کاربردی [1]: متن t فرضیه h را استلزام می‌کند اگر انسانی t را بخواند h را نتیجه بگیرد.

–تعریف ریاضی [2]: متن t فرضیه h را استلزام می‌کند اگر

$$P(h \text{ is true}|t) > P(h \text{ is true})$$

با این تعابیر، به صورت کلی در زمینه تحلیل استلزامات متنی اقدامات مختلفی صورت می‌گیرد که این اقدامات عموما به ۳ دسته مجزا دسته بندی می‌شوند[3]:

– تشخیص استلزام‌های متنی (تشخیص درستی زوج‌های t و h)

– تولید استلزام‌های متنی (تولید همه نتایج ممکن h از روی متن t)

– استخراج استلزام‌های متنی (تولید زوج‌های (t,h))

در این مقاله بر مسئله شناسایی استلزامات متنی (دسته اول) با استفاده از یک روش معنایی مبتنی بر یادگیری ماشین تمرکز شده است. همچنین یک روش پایه نیز برای مقایسه با روش پیشنهادی پیاده سازی و تست شده است. در ادامه مقاله، در بخش دوم به مرور روش‌های پیشین پرداخته شده است، در بخش سوم روش پیشنهادی و جزئیات آن آورده شده است و در نهایت در بخش چهارم و پنجم، ارزیابی و نتیجه گیری ارائه شده است.

۲- مروری بر کارهای پیشین

روش‌های مطرح شده در زمینه استلزام متنی به صورت کلی به ۴ دسته واژگانی، یادگیری ماشین، مبتنی بر گراف و معنایی قابل تقسیم بندی هستند که غالبا روی مسئله تشخیص استلزام‌ها (مورد اول از ۳ وظیفه مطرح شده در قسمت قبل) کار شده است:

– روش واژگانی: این روش‌ها تحلیل سطحی بر روی متن ورودی انجام می دهند. در این روش‌ها پس از انجام پیش پردازش روی متن ، بازنمایی انجام شده (کیسه کلمات و یا n-gram) و با استفاده از منابع دانشی مثل وردنت، امتیازدهی انجام می‌دهند. این امتیازدهی عموما با توجه به همپوشانی لغات، تطبیق توالی‌های کلمات و یا زیر رشته‌های مشترک [4] صورت می‌گیرد. در روش‌های واژگانی، روش‌های پیش پردازش مورد استفاده می‌تواند روش‌های ساده ای مثل توکن بندی، ریشه یابی، لم یابی و کشف ایستواژه‌ها [5] بوده و یا از روش‌های پیچیده تر مثل تشخیص افعال چندبخشی، تشخیص اصطلاحات و ضرب المثل ها، تشخیص موجودیت‌های اسمی و تشخیص آرگومان‌های زمان استفاده شود.

– روش یادگیری ماشین: در این روش‌ها مسئله RTE تبدیل به مسئله طبقه بندی دو کلاسه می‌شود و با استفاده از یک مجموعه ویژگی و یک روش یادگیری ماشین عملیات انجام می‌شود [6]. این مجموعه ویژگی‌ها می‌توانند در قالب‌های زیر باشند:

O فاصله [7]

- تعداد کلمات مشترک
- طول زیر رشته مشترک
- طولانی ترین زیردرخت نحوی مشترک [8]

O عوامل محرک[۲] [9]

- ویژگی قطبیت (وجود کلمات باردار و مثبت بودن یا منفی بودن محتوای t و h می‌تواند در تطبیق آنها تعیین کننده باشد)
- ویژگی تضاد (حضور یا عدم حضور واژگان متضاد در دو جمله، مانند open و close)
- ویژگی‌های الحاقی (هنگام مقایسه t و h، اگر تفاوت دو جمله تنها رخداد کلمه ای مثل regularly در انتهای t باشد، تفاوت در معنا و استلزام ایجاد می‌کند)

O ویژگی‌های زوج

مدلسازی ساختار و محتوای t و h با استفاده از معیارهایی دقیقتر از فاصله مثل روش‌های مقایسه درخت‌های تجزیه آنها [10]

– روش مبتنی بر تطبیق گراف: نمایش t و h در قالب گراف و استفاده از روش‌های تطبیق گراف [11]. این بازنمایی می‌تواند با استفاده از درخت تجزیه [10] و استخراج وابستگی‌ها در قالب یکسری قانون صورت بگیرد و سپس با استفاده از مدلسازی تابع‌های هزینه امتیازدهی صورت بگیرد.

– روش مبتنی بر معنا: این روش‌ها غالبا شامل روش‌های مبتنی بر معیارهای شباهت سنجی معنایی و استفاده از منابع معنایی بوده و در موارد اندکی بازنمایی معنایی منطقی انجام شده [12,13] و سپس استدلال صورت گرفته است (روش‌های مبتنی بر منطق). همچنین مقالاتی نیز وجود دارند که حاشیه نویسی معنایی انجام داده اند [14,15].

روش مطرح شده در این مقاله در دسته روش‌های مبتنی بر یادگیری ماشین قرار می‌گیرد که برخلاف سایر روش‌ها در این دسته از ویژگی‌های معنایی در آن استفاده شده است. به همین دلیل روش پیشنهادی در دسته آخر (روش‌های مبتنی بر معنا) نیز دسته بندی می‌شود. همچنین روش پایه پیاده سازی شده نیز در دسته اول یعنی روش‌های واژگانی قرار می‌گیرد.

۳- معرفی سیستم پیشنهادی

همانطور که پیشتر ذکر شد، در این مقاله یک روش معنایی مبتنی بر یادگیری ماشین برای شناسایی استلزامات متنی معرفی شده است.همچنین یک سیستم پایه مبتنی بر شباهت سنجی برای نمایش تفاوت روش معنایی و واژگانی پیاده سازی شده است. این دو روش بر روی مجموعه ای از زوج‌های متن و فرضیه دارای برچسب وجود یا عدم وجود رابطه استلزامی ارزیابی شده اند. این دو سیستم در ادامه معرفی شده اند.

۳-۱- سیستم پایه

در سیستم پایه، یک روش واژگانی پیاده سازی شده است. در این سیستم برای تشخیص اینکه آیا فرضیه از متن استلزام می‌شود یا خیر از شباهت بین دو جمله استفاده می‌شود و با توجه به یک حد آستانه این تصمیم‌گیری انجام می‌شود. برای محاسبه امتیاز شباهت از یک معیار ترکیبی استفاده می‌شود که سعی می‌کند دو عامل ساختار و محتوای جملات را در نظر بگیرد. به صورت دقیق تر برای محاسبه میزان شباهت دو جمله از جمع وزندار ۶ عامل زیر استفاده می‌شود:

- شباهت بر اساس اشتراک کلمات دو جمله با حذف ایستواژه‌ها (شباهت کسینوسی)
- شباهت بر اساس n-gramهای مشترک
  - ۲-گرام‌های مشترک
  - ۳-گرام‌های مشترک
- شباهت بر اساس مقوله نحوی (pos) کلمات
  - شباهت افعال
  - شباهت اسامی
  - شباهت اسامی خاص

فلوچارت روش پایه در شکل ۱ آورده شده است.

۳-۲- سیستم مبتنی بر معنا

در روش مبتنی بر معنا که اساس سیستم پیشنهادی در این مقاله است، سعی می‌شود تا از اطلاعات معنایی به صورت مستقیم برای شناسایی استلزام متنی کمک گرفته شود. برای این هدف از روش یادگیری ماشین استفاده می‌شود تا با در نظر گرفتن ویژگی‌های معنایی استخراج شده از زوج جملات، تصمیم گیری انجام دهد. در این راستا ابتدا با استفاده از روش برچسب‌زنی نقش معنایی[۳] (با استفاده از ابزار سنا[۴]) بازنمایی معنایی جملات تولید می‌شود. سپس با توجه به این بازنمایی یک مجموعه ویژگی معنایی از هر زوج متن و فرضیه استخراج شده و سیستمی مبتنی بر یادگیری ماشین آموزش داده شده است. در این مقاله ۳ روش متفاوت یادگیری ماشین برای آموزش سیستم مورد استفاده قرار گرفته است که جزئیات آن در ادامه آورده شده است. همچنین فلوچارت روش در شکل ۲ آورده شده است.

۳-۲-۱- بازنمایی معنایی

گفتیم که برای بازنمایی معنایی متون از روش برچسب زنی نقش معنایی (SRL) استفاده شده است. این روش که از آن با عنوان تجزیه معنایی سطحی نیز نام برده می‌شود، در پردازش زبان طبیعی شامل فرآیند تشخیص آرگومان‌های معنایی متناظر با مسند (معمولا فعل) در یک جمله و تخصیص نقش‌های خاص آن هاست [16].

گفتیم که در روش پیشنهادی برای تعیین نقش‌های معنایی از ابزار سنا استفاده شده است. برچسب‌ها در این ابزار برپایه دستورالعمل PropBank [17] بوده و شامل ۶ آرگومان اصلی (Arg0-5 که نمایانگر نقش‌هایی از قبیل عامل و پذیرنده هستند) و ۱۴ توصیف کننده (مانندADV،CAU، TMP، EXT و ..) هستند. شکل ۳ بازنمایی معنایی با استفاده از SRL را برای یک جفت جمله نمونه نشان می‌دهد.

۳-۲-۲- بردار ویژگی

پس از بازنمایی معنایی جملات، یک بردار ویژگی بر اساس اطلاعات معنایی ساخته می‌شود. پس برای هر زوج متن و فرضیه یک بردار ویژگی ساخته شده که با توجه به برچسب آن زوج، در فرآیند آموزش سیستم مورد استفاده قرار می‌گیرد. مجموعه ویژگی‌های مورد استفاده عبارت‌اند از:

- درصد شباهت آرگومان‌ها

برای مقایسه محتوای زوج جملات، نیاز داریم که آرگومان‌های مسندهای متن و فرضیه با هم مقایسه شوند. برای محاسبه این ویژگی آرگومان‌های اصلی متن و فرضیه با هم مقایسه شده و امتیاز آنها نرمال می‌شود. چون ممکن است تعداد آرگومان‌ها در دو جمله متفاوت باشد، این فرآیند یکبار از سمت فرضیه به متن و بار دیگر به صورت برعکس انجام می‌شود.

- شباهت معنایی مسند با استفاده از ورب نت با در نظر گرفتن نفی با استفاده از برچسب AM-NEG

برای افزایش دقت در روش پیشنهادی، از منبع دانشی ورب نت برای مقایسه مسندها استفاده شده است. اگر مسندهای متن و فرضیه از نظر معنایی یکسان باشند ویژگی مقدار یک می‌گیرد. همچنین با توجه به اینکه در بازنمایی مبتنی بر SRL برچسب AM-NEG وجود دارد، هر گاه این برچسب در متن یا فرضیه موجود بود مسندها برای رابطه تضاد بررسی می شوند.

- شباهت مکان با بررسی برچسب AM-LOC

با استفاده از برچسب زن نقش‌های معنایی، بخش‌هایی از جمله که دارای معنای مکان هستند با برچسب AM-LOC مشخص می شوند. این ویژگی با بررسی میزان شباهت دو عبارت دارای برچسب مکان با استفاده از معیار کسینوسی محاسبه می گردد.

- شباهت زمان با توجه به برچسب‌های AM-TMP و AM-EXT

مشاهده و بررسی برخی نمونه‌های بازنمایی شده نشان داد که اطلاعات زمانی در قالب دو برچسب AM-TMP و AM-EXT بیان می‌شوند. پس این دو برچسب باهم مقایسه می‌شوند.

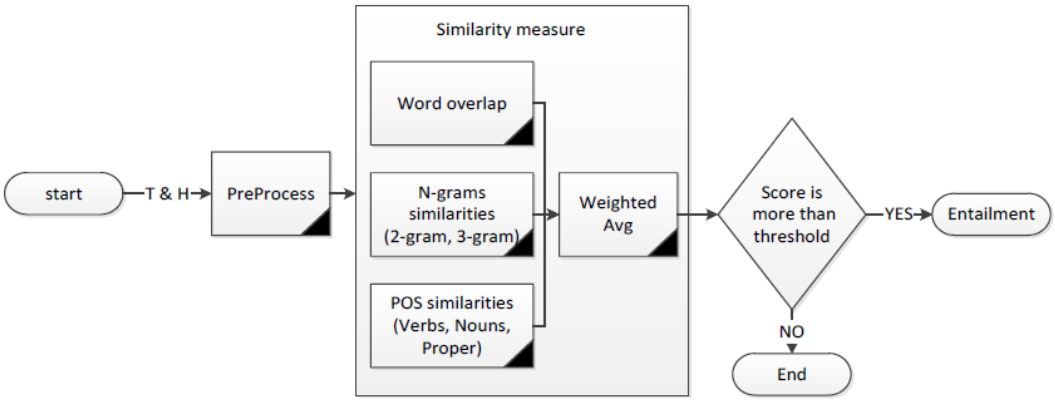
- شباهت علّی با استفاده از برچسب‌های AM-CAU و AM-PNC

با توجه به اینکه علیت و هدف با دو برچسب AM-CAU و AM-PNC بیان می‌شوند برای این بخش این نقش‌ها مقایسه شدند.

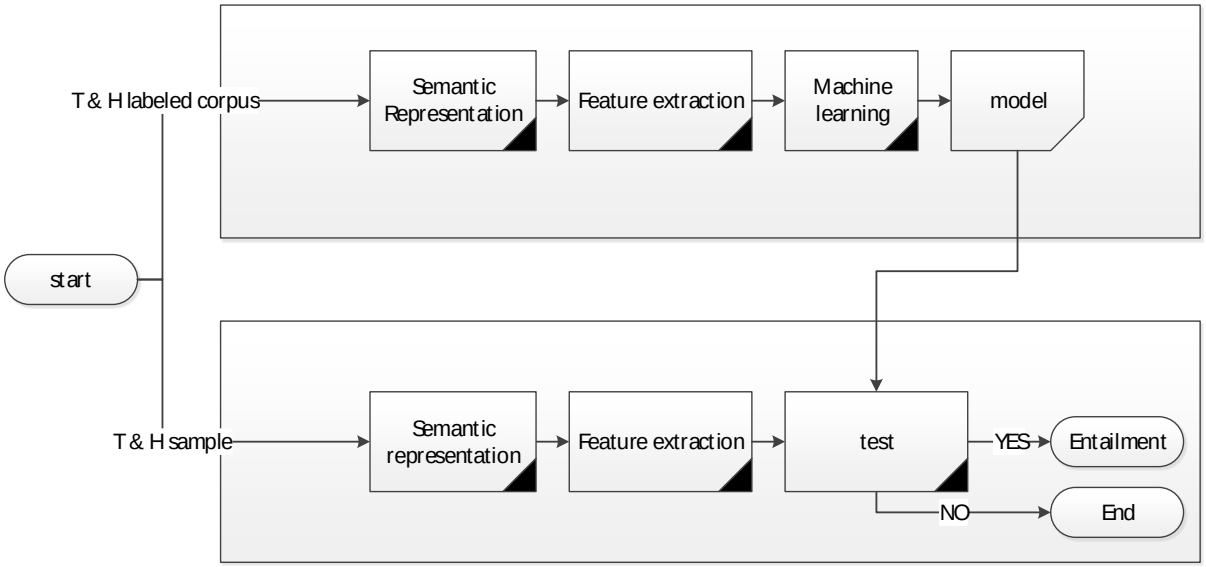
- جهت ب

در برخی جه
بازنمایی برچ
دو جمله متر
این مورد در

- شباهت



شکل ۱: فلوچارت روش پایه



شکل ۲: فلوچارت روش

```
Text: a plane carrying 200 French Jewish immigrants has landed in Israel,
at a time of strained diplomatic relations between Israel and France.
A0  a plane
V   carrying
A1  200 French Jewish immigrants
A1  a plane carrying 200 French Jewish immigrants
V   landed
AM-LOC  in Israel ,
AM-TMP  at a time of strained diplomatic relations between Israel and France


hyp: Nearly 2,000 Jews emigrate from France to Israel every year.
A0  Nearly 2,000 Jews
V   emigrate
A3  from France
A4  to Israel
AM-TMP  every year.
```

شکل ۳: بازنمایی معنایی برای

در این ویژگی اشتراک دسته بندی افعال (در دسته بندی وردنت) بررسی شده است.

- میزان ارتباط بین آرگومان‌ها با استفاده از وردنت

وردنت به عنوان یک منبع دانشی مناسب در زبان انگلیسی ارتباطات واژگانی را مدل کرده است. در این ویژگی بررسی می‌شود که آرگومان‌ها چقدر بهم مرتبط هستند. به عبارت دیگر هم وقوعی آرگومانها در وردنت بررسی شده است و چنانچه کلمات موجود در آرگومانها در یک سینست وردنت رخ داده باشند امتیاز مثبت محسوب می‌شود.

- تطبیق زمان و شخص افعال و POS آن

میزان تطابق افعال از نظر شخص و زمان با توجه به برچسب اجزای کلام مد نظر بوده است.

۳-۲-۳- یادگیری ماشین

همانطور که قبل تر گفته شد در روش‌های شناسایی استلزام متنی با استفاده از یادگیری ماشین، مسئله به صورت دسته بندی دوکلاسه مدل می شود. در روش پیشنهادی نیز به همین صورت پس از تولید بردار ویژگی برای هر زوج، سیستم با الگوریتم‌های متفاوتی ارزیابی شد و سه الگوریتم متفاوت Naïve Bayes، Multi-Layer Perceptron و Meta Multiclass classifier انتخاب و با استفاده از ابزار وکا آموزش داده شد.

٤- ارزیابی

همانطور که توضیح داده شد، در این مقاله دو روش پایه و مبتنی بر معنا برای شناسایی استلزامات متنی توسعه یافتند. در این بخش ارزیابی این دو روش ا رائه شده است.

٤-١- دادگان

مجموعه دادگان مورد استفاده در این مقاله به صورت ترکیبی از دادگان رقابت‌های RTE1-3 تهیه شده و بخش تست و توسعه شامل ۴۵۰۰ زوج جمله برچسب خورده بوده است. پس از اعمال ابزار سنا برای بازنمایی معنایی حدود ۱۰۰۰ زوج به دلیل عدم تولید خروجی حذف شدند. پس ۳۵۲۴ زوج دارای برچسب وجود یا عدم وجود استلزام متنی برای آموزش سیستم مورد استفاده قرار گرفت. تست سیستم با روش ۱۰-fold برای سیستم مبتنی بر معنا و استفاده از داده مشابه برای سیستم پایه انجام گرفت. پراکندگی کلاس‌های استلزام در دادگان به صورت زیر است:

جدول۱: پراکندگی کلاس‌های استلزام

کلاس مثبت (وجود استلزام)	۱۷۶۴
کلاس منفی (و نامشخص)	۱۷۶۰

٤-٢- معیار ارزیابی

معیارهای ارزیابی برای شناسایی استلزامات متنی مشابه اغلب کاربردها در پردازش زبان طبیعی دقت (P)، فراخوانی (R) و معیار F هستند که به صورت زیر تعریف می‌شوند.

P= تصمیم‌های YES درست سیستم/ تمامی تصمیم‌های YES سیستم

R= تصمیم‌های YES درست سیستم/ تمامی نمونه‌های YES

معیار F= 2*P*R/ (P+R)

٤-٣- آزمایشات و بحث

ارزیابی سیستم مبتنی بر معنای پیشنهادی با استفاده از روش 10-fold و ارزیابی روش پایه به سادگی با مقایسه برچسب سیستم با برچسب مرجع انجام شد. نتایج ارزیابی در جدول ۱ آورده شده است:

جدول۲: ارزیابی سیستم ها

System/Measure	Precision	Recall	F-Measure
Base system	0.564	0.37.1	0.443
Naïve Bayes	0.528	0.704	0.604
Multi-layer Perceptron	0.511	0.751	0.608
Meta Multiclass Classifier	0.521	0.774	0.622

همانطور که جدول ۲ نشان می‌دهد، تمامی سه روش مبتنی بر یادگیری ماشین به نتایج بهتری نسبت به سیستم پایه دست یافتند. اگرچه در روش پایه از ویژگی‌های واژگانی، چندگرام‌ها و ویژگی اجزای کلام استفاده شده است، اما نتایج نشان می‌دهد که ویژگی‌های مبتنی بر معنا به تصمیم انسان نزدیکتر هستند و ظاهرا مجموعه مناسبی هستند. همچنین به نظر می‌رسد توصیف کننده‌های استخراج شده از بازنمایی معنایی SRL جملات را به خوبی توصیف نموده اند. در میان سیستم‌های مبتنی بر یادگیری ماشین سیستم فرا طبقه بند چندکلاسه بهترین نتیجه یعنی معیار F ۶۲ را ارائه نموده است. (لازم به کر است که ارزیابی برای کلاس منفی و نامشخص نیز انجام شد که نتیجه برای بهترین سیستم معیار F ۵۶,۷ بود، اما به دلیل تمرکز فعلی کاربرد بر تشخیص استلزام و نه رد استلزام، ارزیابی کلاس مثبت مورد توجه قرار گرفت.)

در بخش دیگری از ارزیابی میزان تاثیر هر یک از ویژگی‌ها با استفاده از ابزار وکا (بخش انتخاب ویژگی‌ها و روش SVM attribute evaluator) محاسبه شده و ویژگی‌ها بر این اساس رتبه بندی شدند. این رتبه‌بندی به ترتیب اشتراک آرگومان‌ها، زمان، جهت، ویژگی‌های نحوی فعل، مکان، ارتباط بر اساس وردنت، معنی فعل، علیت و نوع فعل است.

با توجه به اینکه دادگان آموزشی به اقتضای حجم به صورت ترکیبی برای آموزش و تست سیستم استفاده شده بودند امکان مقایسه دقیق نتایج

سیستم با سایر سیستم‌ها در رقابت‌های RTE وجود نداشت. اما به صورت کلی با مشاهده نتایج سیستم‌ها در مقالات ارائه شده عملکرد سیستم پیشنهادی نسبت به بهترین سیستم‌ها عملکرد خوب و مشابهی است. این ادعا بر این اساس است که در [18]عنوان شده که بهترین نتیجه در رقابت‌های RTE1-3، صحت ۸۰ درصد (در نظر گرفتن نمونه‌های درست هر دو کلاس بخش بر کل نمونه ها) بوده است. این در حالی است که صحت بهترین سیستم پیشنهادی در این مقاله بر روی دادگان تجمیع شده RTE1-3 برابر ۷۷,۴ محاسبه شده است که نتیجه نزدیک و قابل قبولی است.

۵- نتیجه گیری

در این مقاله یک روش معنایی مبتنی بر یادگیری ماشین برای شناسایی استلزامات متنی پیشنهاد شده است. همچنین یک سیستم ساده مبتنی بر شباهت سنجی نیز برای مقایسه با سیستم اصلی توسعه یافته است. در روش مبتنی بر معنا مجموعه ای از ویژگی‌های معنایی برای آموزش سیستم یادگیری ماشینی مورد استفاده قرار گرفتند که با استفاده از آن سیستم بتواند تصمیم گیری کند که آیا بین متن و فرضیه رابطه استلزامی برقرار است یا خیر. ویژگی‌های مورد استفاده نیز به صورت عمده از روی بازنمایی معنایی زوج متن و فرضیه استخراج می‌گردند.

همانطور که پیش بینی می‌شد با توجه به نتایج ارزیابی ها، روش مبتنی بر معنا نسبت به روش پایه عملکرد بهتری داشت که این موضوع به دلیل نمایش بهتر محتوا با استفاده از ویژگی‌های معنایی است. به عنوان کار آینده در این زمینه می‌توان به تعریف دقیق تر ویژگی‌ها و استفاده از مدل‌های دیگر بازنمایی معنایی اشاره نمود.

مراجع

[1] S. Ghuge and A. Bhattacharya, “Survey in Textual Entailment,” Cfilt.Iitb.Ac.in, no. 1, pp. 1–28, 2013.

[2] O. Glickman and I. Dagan, “A probabilistic setting and lexical cooccurrence model for textual entailment,” in *Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment - EMSEE ’05*, 2005, no. June, pp. 43–48, 2005.

[3] I. Androutsopoulos and P. Malakasiotis, “A survey of paraphrasing and textual entailment methods,” J. Artif. Intell. Res., vol. 38, pp. 135–187, 2010.

[4] Y. Tan, “Applying CRFs and SVMs to Textual Entailment Recognizing”, Recognizing Textual Entailment (RTE),” vol. 1, no. January, pp. 1–8, 2012.

[5] P. Pakray, S. Neogi, S. Poria, and A. Gelbukh, “A Textual Entailment System using Anaphora Resolution,” 2011.

[6] E. Agichtein, W. Askew, and Y. Liu, “Combining Lexical , Syntactic , and Semantic Evidence For Textual Entailment Classification,” Proc. TAC, pp. 1–6, 2008.

[7] Y. Tan, “Recognizing Textual Entailment Using Multiple Features and Filters,” pp. 240–243, 2014.

[8] Y. Mehdad, A. Moschitti, and F. M. Zanzotto, “Syntactic/Semantic Structures for Textual Entailment Recognition,” in The 2010 Annual Conference of the North American Chapter of the ACL, 2010, no. June, pp. 1020–1028.

[9] H. Ren, D. Ji, and J. Wan, “WHU at TAC 2009: A Tri-categorization Approach to Textual Entailment Recognition,” in Proceedings of Text Analysis Conference, 2009.

[10] R. Bar-haim, “Semantic Inference at the Lexical-Syntactic Level,” Proteus, no. January, pp. 871–876, 2010.

[11] A. D. Haghighi, A. Y. Ng, and C. D. Manning, “Robust textual inference via graph matching,” in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT ’05*, 2005, pp. 387–394.

[12] B. MacCartney and C. D. Manning, “Natural Logic for Textual Inference,” Proc. ACL-PASCAL Work. Textual Entailment Paraphrasing - RTE ’07, no. June, pp. 193–200, 2007.

[13] J. Bos and K. Markert, “Recognising textual entailment with robust logical inference,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 3944 LNAI, pp. 404–426, 2006.

[14] A. Toledo, S. Katrenko, S. Alexandropoulou, H. Klockmann, A. Stern, I. Dagan, and Y. Winter, “Semantic annotation for textual entailment recognition,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7630 LNAI, no. PART 2, pp. 12–25, 2013.

[15] A. Burchardt and M. Pennacchiotti, “Fate: a framenet-annotated corpus for textual entailment,” Proc. Lr., vol. 2008, no. iii, pp. 539–546, 2008.

[16] D. Gildea and D. Jurafsky, Automatic Labeling of Semantic Roles,. In Proceedings of the 38th Annual Conference of the Association for Computational Linguistics (ACL-00), pp. 512–520, Hong Kong, October 2000.

[17] Martha Palmer, Dan Gildea, Paul Kingsbury, The Proposition Bank: A Corpus Annotated with Semantic Roles Computational Linguistics Journal, 31:1, 2005.

[18] I. DAGAN, B. DOLAN, B. MAGNINI, and D. ROTH, “Recognizing textual entailment: Rational, evaluation and approaches,” Nat. Lang. Eng., vol. 15, no. 4, pp. i–xvii, 2009.

زیرنویس‌ها

[1] Recognizing Textual Entailment

[2] trigger

[3] Semantic Role Labeling

[4] ronan.collobert.com/senna/