

شبکه‌های توجه گراف و درخت تجزیه وابستگی نحوی در تحلیل احساسات

مبتنی بر جنبه: مرور نظام‌مند

شیماروحانی

دانشجوی کارشناسی ارشد، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان، ایران

صبا صارمی‌نیا

استادیار، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان، ایران

چکیده

تحلیل احساسات مبتنی بر جنبه، یکی از شاخه‌های مهم پردازش زبان طبیعی است که هدف آن استخراج و تحلیل احساسات مرتبط با جنبه‌های خاصی از متن می‌باشد. با توجه به پیچیدگی زبان انسانی و نیاز به درک بهتر ارتباطات بین کلمات، استفاده از روش‌های پیشرفته مانند شبکه توجه گراف و درخت تجزیه وابستگی، توجه زیادی را به خود جلب کرده است. شبکه توجه گراف با بهره‌گیری از ساختارهای گرافی و مکانیزم توجه، توانایی مدل‌سازی ارتباطات پیچیده بین کلمات را دارد، در حالی که درخت تجزیه وابستگی با تجزیه متن به ساختارهای وابسته، ارتباطات نحوی و معنایی را بهتر درک می‌کند. در این مقاله مروری، به بررسی استفاده از این دو روش در تحلیل احساسات مبتنی بر جنبه می‌پردازیم. نتایج این بررسی نشان می‌دهد که در زبان انگلیسی ترکیب شبکه توجه گراف و درخت تجزیه وابستگی توانسته بهبود قابل توجهی در دقت و کارایی سیستم‌های تحلیل احساسات مبتنی بر جنبه در حوزه‌های متنوع رویدادهای اجتماعی، کالاهای الکترونیکی، محصولات و خدمات رستوران‌ها و برندهای مختلف ایجاد کند. همچنین نتایج نشان‌دهنده روند افزایشی پژوهش‌های انجام شده در سال‌های اخیر است که اهمیت این موضوع را بیان می‌کند.

کلیدواژه‌ها: پردازش زبان طبیعی، کلان داده، تحلیل احساسات، درخت تجزیه وابستگی نحوی، شبکه توجه گراف

۱. مقدمه

به دلیل رواج یافتن استفاده از شبکه‌های اجتماعی و وبسایت‌های آنلاین، حجم زیادی از محتوای متنی توسط شخصیت‌های حقیقی و حقوقی تولید می‌شوند. استخراج دانش از این نظرات می‌تواند به درک عمیق رویکردهای اجتماعی، سیاسی، تجاری و ... کمک کند.

تحلیل احساسات در این زمینه نقش مهمی را ایفا می‌کند. تحلیل احساسات یکی از شاخه‌های مهم و کاربردی در زمینه پردازش زبان طبیعی است که به شناسایی و تفسیر نظرات، احساسات و نگرش‌های کاربران نسبت به محصولات، خدمات، رویدادها و سایر موجودیت‌ها می‌پردازد. این فناوری به کسب‌وکارها و سازمان‌ها امکان می‌دهد تا به شکلی دقیق‌تر و علمی‌تر از نظرات و بازخوردهای مشتریان بهره‌برداری کنند و در نتیجه، تصمیمات بهتری در زمینه بازاریابی، بهبود محصولات و خدمات، و مدیریت ارتباط با مشتریان اتخاذ نمایند (Liu., 2012).

تجزیه و تحلیل احساسات نقش مهمی در دنیای کنونی ایفا می‌کند و به کسب و کارها اجازه می‌دهد، نظرات و احساسات مشتری را نسبت به محصولات یا خدمات خود درک کنند. این بینش ارزشمند به شرکت‌ها کمک می‌کند تا تصمیمات آگاهانه بگیرند، رضایت مشتری را بهبود بخشند و شهرت کلی برند خود را افزایش دهند. براساس مطالعه‌ای که توسط گارتنر انجام شده است؛ تا سال ۲۰۲۲، بالغ بر ۸۵٪ از تعاملات با مشتری بدون دخالت انسان مدیریت می‌شود و همین موضوع اهمیت روبه رشد کیفیت و دقت در تحلیل احساسات در عصر دیجیتال را برجسته می‌کند (Gartner, 2021).

تحلیل احساسات می‌تواند در سطوح مختلفی مثل سطح سند، سطح جمله و سطح جنبه انجام شود. یکی از مهمترین سطح تحلیل‌ها، سطح جنبه می‌باشد؛ زیرا در خلاصه‌سازی متون و ارائه گزارشات ساختارمند از متن بسیار پرکاربرد است. هدف اصلی تحلیل احساسات مبتنی بر جنبه^۱ که در سال‌های اخیر توجه بسیاری از پژوهشگران را به خود جلب کرده است، شناسایی و استخراج احساسات مربوط به جنبه‌های خاصی از متن است که می‌تواند در زمینه‌های مختلفی از جمله تحلیل بازخورد مشتریان، بررسی نظرات کاربران در شبکه‌های اجتماعی و بهبود سیستم‌های پیشنهاددهنده مورد استفاده قرار گیرد (Liu, 2022). برای مثال در محیط کسب‌وکار، تحلیل احساسات در سطح جنبه می‌تواند به شرکت‌ها کمک کند تا بازخوردهای مشتریان را با دقت بیشتری تحلیل کنند. به جای اینکه تنها به مثبت یا منفی بودن کلی یک نظر توجه شود، تحلیل احساسات مبتنی بر جنبه اجازه می‌دهد که احساسات نسبت به جنبه‌های خاصی از یک محصول یا خدمات شناسایی شود. این رویکرد به مدیران اجازه می‌دهد تا نقاط قوت و ضعف دقیق‌تری را در محصولات خود تشخیص دهند و بر اساس آن تصمیم‌گیری کنند (Liu., 2012). در واقع به سازمان‌ها این امکان را می‌دهد که به طور مستقیم به نیازها و دغدغه‌های مشتریان پاسخ دهند. با شناسایی احساسات مرتبط با جنبه‌های خاصی از خدمات یا محصولات، شرکت‌ها می‌توانند اقدامات بهتری برای بهبود تجربه مشتری انجام دهند. این نوع تحلیل می‌تواند نشان دهد که کدام ویژگی‌ها بیشترین تأثیر مثبت یا منفی را بر روی تجربه مشتری دارند (Pontiki et al., 2016).

یکی از چالش‌های اصلی در تحلیل احساسات مبتنی بر جنبه، درک صحیح ارتباطات بین کلمات و ساختارهای نحوی و معنایی متن است. برای مقابله با این چالش، روش‌های مختلفی پیشنهاد شده است که از جمله مهم‌ترین آن‌ها می‌توان به شبکه توجه گراف^۲ و درخت تجزیه کننده وابستگی^۳ اشاره کرد. شبکه توجه گراف با استفاده از مکانیزم توجه، ارتباطات پیچیده بین کلمات را مدل‌سازی کرده و به درک بهتر ساختارهای گرافی کمک می‌کند (Veličković et al., 2017). درخت تجزیه وابستگی نیز با تجزیه متن به ساختارهای وابسته، ارتباطات نحوی و معنایی کلمات را استخراج می‌کند که می‌تواند در بهبود دقت تحلیل احساسات موثر باشد (Chen & Manning, 2014).

¹ Aspect Base Sentiment Analysis

² Graph Attention Network

³ dependency parser tree

ترکیب این دو روش می تواند بهبود قابل توجهی در عملکرد سیستم های تحلیل احساسات مبتنی بر جنبه ایجاد کند. در این مقاله مروری، به بررسی جامع استفاده از شبکه توجه گراف و درخت تجزیه وابستگی در تحلیل احساسات مبتنی بر جنبه می پردازیم. در ادامه این پژوهش در بخش ۲ به بررسی پیشینه پژوهش در زمینه تحلیل احساسات مبتنی بر جنبه می پردازیم؛ در بخش ۳ روش انجام پژوهش را بیان می کنیم؛ در این بخش سوالی که این پژوهش برای پاسخ به آن انجام شده را مطرح و تحلیل آماری از یافته ها انجام می دهیم. در بخش ۴ درخت وابستگی نحوی و پژوهش های انجام شده با استفاده از آن را بررسی می کنیم. در بخش ۵ به بررسی شبکه توجه گراف می پردازیم. در بخش ۶ پژوهش های انجام شده با استفاده از ترکیب درخت وابستگی نحوی و شبکه توجه گراف را مرور می کنیم و در نهایت به بیان جهت گیری های تحقیقاتی آینده و نتیجه گیری می پردازیم.

۲. پیشینه پژوهش

تحلیل احساسات مبتنی بر جنبه یک زمینه تحقیقاتی مهم در پردازش زبان طبیعی است که به شناسایی و استخراج احساسات مربوط به جنبه های خاصی از متن می پردازد. این زمینه تحقیقاتی به خصوص در تحلیل نظرات کاربران و بازخورد مشتریان در زمینه های مختلفی نظیر شبکه های اجتماعی و سیستم های پیشنهاددهنده کاربرد بسیار دارد.

روش های اولیه برای تحلیل احساسات مبتنی بر جنبه بیشتر بر مبنای استفاده از ویژگی های دستی و الگوریتم های یادگیری ماشین سنتی نظیر ماشین های بردار پشتیبان و جنگل تصادفی بودند. با ظهور تکنیک های یادگیری عمیق، مدل هایی مانند شبکه های عصبی کانولوشنی^۱، شبکه های عصبی بازگشتی^۲، واحدهای بازگشتی دروازه دار^۳ و حافظه بلند مدت کوتاه^۴ به کار گرفته شدند تا بازنمایی های پخش شده با ابعاد پایین را به طور خودکار یاد بگیرند و جایگزین مهندسی ویژگی های دستی شوند (Yuan et al., 2022).

یکی از چالش های اصلی در تحلیل احساسات مبتنی بر جنبه، هماهنگی صحیح بین اصطلاحات جنبه و کلمات زمینه مرتبط برای شناسایی صحیح قطبیت احساسی است. مکانیزم توجه^۵ یکی از رویکردهای موثر برای حل این چالش است که می تواند به طور تطبیقی، بر اساس میزان تأثیر آن در طبقه بندی قطبیت جنبه، وزنی به هر کلمه زمینه ای اختصاص دهد (Gan & Tang, 2021). مطالعات نشان داده اند که استفاده از شبکه های حافظه ای و همچنین توجه موقعیتی می تواند به بهبود عملکرد مکانیزم توجه کمک کند (Gu et al., 2018).

در سال های اخیر، مدل های توجه گراف و شبکه های کانولوشن گراف^۶ به دلیل توانایی آن ها در درک روابط پیچیده و وابستگی های نحوی بین کلمات در یک جمله، توجه زیادی را به خود جلب کرده اند. این مدل ها با استفاده از گراف های وابستگی نحوی، ارتباطات طولانی مدت بین کلمات را بهبود می بخشند و اطلاعات محلی و وابستگی های نحوی را بهتر درک می کنند (Yuan et al., 2022). برای مثال شبکه توجه گراف مبتنی بر وابستگی^۷ و شبکه های گراف توجه خاص جنبه^۸ از ترکیب اطلاعات محلی و وابستگی های

¹ Convolution Neural Networks

² Recurrent Neural Networks, RNN

³ Gated return units, GRU

⁴ Long short-term memory, LSTM

⁵ Attention mechanism

⁶ Graph Convolutional Networks, GCN

⁷ Dependency-based Graph Attention Network, DGAT

⁸ Aspect-Specific Graph Attention Networks, ASGAT

نحوی استفاده می کنند تا ویژگی های معنایی غنی تری استخراج کنند و عملکرد بهتری در مجموعه داده های مختلف نشان دهند (Gan & Tang, 2021).

۳. روش پژوهش

مرور نظام مند یک روش پژوهشی ساختارمند و دقیق است که برای جمع آوری و تحلیل مطالعات موجود در یک حوزه مشخص استفاده می شود. هدف این روش، شناسایی، ارزیابی و ترکیب تمامی شواهد مربوط به یک سؤال پژوهشی خاص است تا نتایجی جامع و معتبر به دست آید. مراحل انجام مرور نظام مند شامل تعریف سؤال پژوهش، تعیین معیارهای شمول و خروج، جستجوی سیستماتیک در پایگاه های داده، انتخاب مقالات بر اساس معیارها، ارزیابی کیفیت مقالات و تحلیل داده ها می باشد (Moher et al., 2009). در این مطالعه، از روش مرور نظام مند برای شناسایی و بررسی پژوهش ها در زمینه استفاد از روش های شبکه های توجه گراف و درخت تجزیه وابستگی در تحلیل احساسات مبتنی بر جنبه استفاده شده است. مراحل انجام مرور سیستماتیک به شرح زیر است:

۳/۱. تعریف سؤال پژوهش

آیا استفاده از شبکه های توجه گراف و درخت تجزیه وابستگی می تواند به بهبود تحلیل احساسات مبتنی بر جنبه کمک کند؟

۳/۲. معیارهای انتخاب مقالات

- مقالات انتخاب شده باید به تحلیل احساسات مبتنی بر جنبه با استفاده از شبکه های توجه گراف و درخت تجزیه وابستگی پرداخته باشند.
- در مجلات معتبر و کنفرانس های بین المللی منتشر شده باشند.
- به زبان انگلیسی نوشته شده باشند.
- بین سال های ۲۰۱۰ تا ۲۰۲۴ منتشر شده باشند.

۳/۳. منابع و پایگاه های داده

جستجو در پایگاه های داده ScienceDirect و Scopus و Goole Scholar انجام شد.

۳/۴. فرآیند جستجو

برای جستجو از کلمات کلیدی زیر استفاده شد:

- "Syntactic Attention Networks"
- "Graph Attention Networks"
- "Dependency Parsing"
- "Aspect-based Sentiment Analysis"

۳/۵. فرآیند بررسی و انتخاب مقالات

در ابتدا با استفاده از کلیدواژه های مذکور، حدود ۱۶۰ مقاله در پایگاه های داده مذکور یافت شد. پس از بررسی عنوان و چکیده، ۳۵ مقاله مرتبط انتخاب شدند. متن کامل این مقالات مورد بررسی دقیق تر قرار گرفت و در نهایت ۱۲ مقاله که تمرکز اصلی آن ها بر توسعه روش های مبتنی بر استفاده از شبکه های توجه گراف و درخت وابستگی نحوی بود، برای تحلیل نهایی انتخاب شدند.

جدول ۱. پژوهش‌های انتخاب شده

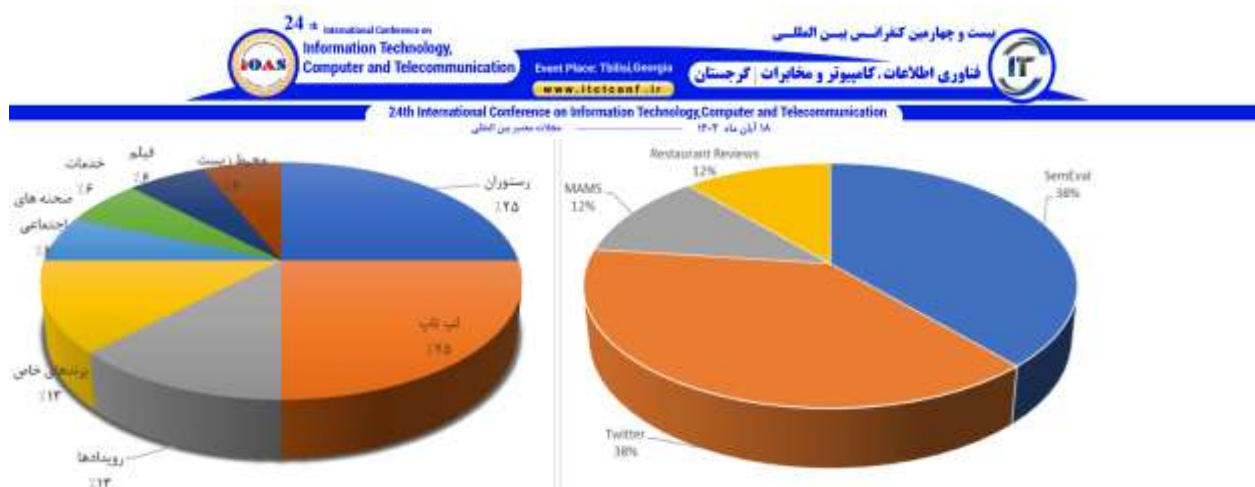
منبع	مجموعه داده	حوزه						زبان	روش‌های استفاده شده
		رستوران	لپ تاپ	رویادها	برندهای خاص	صفحه‌های	سایر		
(Gu et al., 2018)	SemEval 2014	☑	☑					انگلیسی	شبکه توجه دو جهته آگاه از موقعیت مبتنی بر واحد برگشتی دروازه ای ^۱ دو طرفه
(Sun et al., 2019)	Twitter, SemEval 2014, 2016	☑	☑	☑	☑			انگلیسی	درخت وابستگی نحوی و شبکه کانولوشن گراف به همراه مکانیزم توجه
(Wang et al., 2020)	Twitter, SemEval 2014	☑	☑	☑	☑			انگلیسی	درخت وابستگی نحوی و شبکه توجه گراف رابطه‌ای
(Gan & Tang, 2021).	Twitter, SemEval 2014, 2015, 2017	☑	☑	☑	☑			انگلیسی	شبکه توجه گراف، درخت تجزیه وابستگی نحوی، شبکه حافظه بلندمدت-کوتاه دوجهته، ماتریس موقعیت محلی
(Li et al., 2021)	Twitter, SemEval	☑	☑	☑	☑			انگلیسی	درخت وابستگی نحوی و شبکه کانولوشن گراف به همراه مکانیزم توجه و حافظه بلندمدت-کوتاه دوجهته
(Yuan et al., 2022)	Twitter, SemEval 2014, 2015, 2016	☑	☑	☑	☑			انگلیسی	تعبیه موقعیت، تعبیه یال، شبکه توجه گراف، درخت تجزیه وابستگی نحوی، شبکه حافظه بلندمدت-کوتاه دوجهته
(Wu et al., 2022)	Twitter, SemEval 2014	☑	☑					انگلیسی	شبکه توجه گراف رابطه‌ای، درخت وابستگی نحوی و شبکه حافظه بلندمدت-کوتاه دوجهته
(Wang et al., 2023)	SemEval2014, restaurant reviews, ACL14 Twitter, MAMS	☑	☑				خدمات، محیط زیست، فیلم	انگلیسی	درخت تجزیه وابستگی، درخت تجزیه اجزا، شبکه توجه گراف و شبکه کانولوشن گراف

¹ Gated Recurrent Unit

منبع	مجموعه داده	حوزه						زبان	روش‌های استفاده شده
		رستوران	لب تاب	رویدادها	برندهای خاص	صفحه‌های	سایر		
(Shi et al., 2023)	SemEval	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					انگلیسی	درخت وابستگی نحوی و شبکه کانولوشن گراف به همراه مکانیزم توجه
(Wu et al., 2023)	restaurant reviews, Twitter, MAMS	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	خدمات، محیط زیست، فیلم	انگلیسی	درخت وابستگی نحوی و شبکه توجه گراف آگاه از دامنه
(Zhang et al., 2023)	Restaurant14, Laptop14, Twitter	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			انگلیسی	شبکه واحد برگشتی دروازه ای دو طرفه، شبکه توجه گراف با در نظر گرفتن ترکیب نحوی، معنایی و دانشی
(Wang et al., 2024)	Twitter, SemEval 2014, MAMS	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	خدمات، محیط زیست، فیلم	انگلیسی	شبکه توجه گراف رابطه‌ای، درخت وابستگی نحوی و شبکه حافظه بلندمدت-کوتاه دوجهته، شبکه کانولوشن گراف معنایی، ماژول ترکیب نحوی و ماژول ترکیب معنایی براساس مکانیزم دروازه

۴. تحلیل آماری

پژوهش‌های یافت شده در زمینه مورد نظر از سال ۲۰۱۸ تا سال ۲۰۲۴ منتشر شده‌اند. مشاهده می‌شود که تعداد مقالات از سال ۲۰۱۸ تا ۲۰۲۳ روند افزایشی داشته و در سال ۲۰۲۳ به اوج خود رسیده است. همانطور که در جدول ۱ مشاهده می‌شود، تقریباً تمام پژوهش‌های انجام شده در این زمینه روی مجموعه داده‌های معروف Twitter, SemEval, MAMS صورت گرفته‌اند؛ بنابراین حوزه موضوعی آن‌ها نیز مشابه و در زمینه‌های رستوران و لب تاب، محصولات و خدمات برندهای معروف، رویدادها، صفحه‌های اجتماعی و محیط زیست می‌باشد. تنوع موضوعی در پژوهش‌ها به خصوص آن‌هایی که از مجموعه داده Twitter استفاده کرده‌اند، نشان‌دهنده گستردگی کاربرد روش ارائه شده در زمینه‌های مختلف است (شکل ۱).



در تمامی مقالات از روش های شبکه توجه گراف و درخت وابستگی نحوی برای بهبود تشخیص روابط نحوی بین کلمات جنبه و زمینه و بهبود نتیجه تحلیل احساسات مبتنی بر جنبه استفاده شده است. همچنین از شبکه حافظه بلندمدت-کوتاه دوجهته یا واحد برگشتی دروازه ای دو جهته برای ایجاد بازنمایی کلمات پس از مرحله تعبیه کلمات، استفاده شده است؛ که استفاده از مورد اول در مقالات بیشتر بوده است. حدود نیمی از مقالات از شبکه کانولوشن گراف برای استخراج اطلاعات معنایی و در نتیجه بهبود نتایج تحلیل احساسات مبتنی بر جنبه استفاده کرده اند. در سال 2023 نیز دو پژوهش از اطلاعات دانش دامنه برای بهبود استفاده کرده اند. اکثر پژوهش ها در مجموعه داده های رستوران و لب تاب (SemEval, MAMS) به دقت های بالای ۸۰٪ و در مجموعه داده توئیتر که حوزه موضوعی آن گسترده تر است به دقت های بالای ۷۰٪ رسیده اند.

نکته دیگری که در این بررسی حائز اهمیت است، این است که همه این پژوهش ها در زبان انگلیسی انجام شده اند و در زبان های دیگر به خصوص زبان فارسی شکاف عمیقی در زمینه تشخیص روابط نحوی بین کلمات جنبه و زمینه مشاهده می شود. در ادامه پژوهش به بررسی دقیق تر پژوهش های یافت شده پرداخته خواهد شد. خلاصه شده ی جدول ۱ در جدول ۲ قابل مشاهده است.

جدول ۲. خلاصه پژوهش های انجام شده

سال	زبان	حوزه								داده				روش تحلیل			
		رستوران	لب تاب	رویدادها	برندهای خاص	صفحه های اجتماعی	خدمات	فیلم	محیط زیست	SemEval	Twitter	MAMS	Restaurant Reviews	درخت وابستگی	شبکه توجه گراف	BiLSTM	CNN
2018	انگلیسی	☑	☑							☑				☑	☑	☑	
2019	انگلیسی	☑	☑	☑	☑					☑	☑			☑	☑		☑
2020	انگلیسی	☑	☑	☑	☑					☑	☑			☑	☑		

سال	زبان	حوزه								داده				روش تحلیل			
		رستوران	لپ تاپ	رویدادها	برندهای خاص	صفحه های اجتماعی	خدمات	فیلم	محیط زیست	SemEval	Twitter	MAMS	Restaurant Reviews	درخت وابستگی	شبکه توجه گراف	BiLSTM	CNN
2021	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2022	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>							<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
2023	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>							<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
2024	انگلیسی	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

همانطوری که در جدول ۲ مشهود است توجه به تحلیل محتوای متنی در حوزه های موضوعی متنوع مانند فیلم، محیط زیست و ارائه خدمات در سال های جدید با استفاده از روش های شبکه توجه گراف و درخت وابستگی نحوی افزایش یافته است. این موضوع نشان دهنده این است که استفاده از این روش ها در تشخیص محتوای متن از قدرت قابل قبولی برخوردار است. لذا در ادامه به توصیف این روش ها پرداخته میشود.

۵. درخت وابستگی نحوی

تجزیه وابستگی نحوی یکی از تکنیک های مهم در پردازش زبان طبیعی است که به تحلیل ساختار دستوری جملات با تعیین روابط بین کلمات "سر" و وابسته های آن ها می پردازد. این روش مشخص می کند که چگونه کلمات در یک جمله به هم متصل شده اند و بر روابط نحوی که ساختار جمله را تعیین می کنند، تمرکز دارد. هدف اصلی این است که این وابستگی ها را ترسیم کرده و یک درخت وابستگی ایجاد کنیم که فهم بهتری از ترکیب و معنای جملات ارائه می دهد (Dozat & Manning, 2016). تجزیه وابستگی

نحوی به خصوص در تحلیل احساسات مبتنی بر جنبه بسیار مفید است زیرا به شناسایی نقش‌های نحوی و معنایی کلمات در جمله کمک می‌کند و ارتباطات با فاصله طولانی بین کلمات را مشخص می‌سازد (Kiperwasser & Goldberg, 2016).

روش‌های مختلفی برای تجزیه وابستگی نحوی وجود دارد که به دو دسته اصلی تقسیم می‌شوند: روش‌های مبتنی بر قواعد و روش‌های مبتنی بر یادگیری ماشین. روش‌های مبتنی بر قواعد از مجموعه‌ای از قوانین نحوی استفاده می‌کنند تا ساختار جملات را تحلیل کنند. این روش‌ها به دلیل پیچیدگی قوانین نحوی و نیاز به تخصص زبانی، معمولاً کمتر استفاده می‌شوند (Chen & Manning, 2014). در مقابل، روش‌های مبتنی بر یادگیری ماشین و یادگیری عمیق، از مدل‌های آماری و شبکه‌های عصبی برای یادگیری خودکار ساختارهای نحوی از داده‌های برچسب‌گذاری شده استفاده می‌کنند. از جمله این مدل‌ها می‌توان به شبکه‌های عصبی بازگشتی، حافظه بلند-کوتاه مدت و شبکه‌های عصبی کانولوشنی اشاره کرد (Dozat & Manning, 2016).

تجزیه وابستگی نحوی می‌تواند به طور قابل توجهی دقت و عملکرد ABSA را بهبود بخشد. این روش با تجزیه جملات به ساختارهای وابسته، می‌تواند ارتباطات نحوی بین کلمات را استخراج کرده و به مدل‌ها کمک کند تا احساسات مرتبط با هر جنبه را دقیق‌تر شناسایی کنند. به عنوان مثال، درخت‌های وابستگی نحوی می‌توانند به تعیین کلمات کلیدی مرتبط با هر جنبه کمک کنند و احساسات مثبت یا منفی مرتبط با آن‌ها را بهتر مشخص کنند (Kiperwasser & Goldberg, 2016). بنابراین کارهای متعددی در این زمینه انجام شده برای مثال مدل‌های گراف توجه مبتنی بر وابستگی که از ترکیب اطلاعات نحوی و معنایی برای استخراج ویژگی‌های غنی‌تر و دقیق‌تر استفاده می‌کنند (Kiperwasser & Goldberg, 2016)؛ یا استفاده از شبکه‌های عصبی کانولوشنی مبتنی بر درخت که با استفاده از ساختارهای وابستگی نحوی، ویژگی‌های محلی و جهانی متن را استخراج کرده و دقت طبقه‌بندی احساسات را افزایش می‌دهند (Chen & Manning, 2014). همچنین مدل‌های حافظه تکراری با مکانیزم توجه که از توجه موقعیتی و اطلاعات نحوی برای بهبود تطبیق بین جنبه‌ها و کلمات زمینه استفاده می‌کنند و به شناسایی دقیق‌تر احساسات کمک می‌کنند (Ma & Hovy, 2016).

۶. شبکه‌های توجه گراف

شبکه‌های توجه گراف به‌عنوان یک رویکرد دگرگون‌کننده در قلمرو شبکه‌های عصبی ظهور کرده‌اند که به‌طور خاص برای مدیریت داده‌های ساختار یافته به‌صورت گراف طراحی شده‌اند. در حالی که شبکه‌های عصبی سنتی در داده‌های شبکه‌ای یا متوالی برتری دارند، این شبکه‌ها به‌طور منحصربه‌فردی برای گرفتن روابط پیچیده و وابستگی‌های ذاتی در نمایش داده‌های مبتنی بر گراف مناسب هستند. در قلب شبکه‌های توجه گراف، کاربرد نوآورانه مکانیسم توجه نهفته است. مشابه نقش آنها در مدل‌های متوالی مانند ترانسفورمرها^۱ و مکانیسم‌های توجه در شبکه‌های توجه گراف، گره‌های درون یک گراف را قادر می‌سازند تا به‌طور انتخابی بر روی اطلاعات نودهای همسایه تمرکز کنند. این سازگاری به هر گره اجازه می‌دهد تا اهمیت ویژگی‌های مختلف را از گره‌های همسایه خود بسنجد و توانایی مدل برای تشخیص و استفاده از اطلاعات مرتبط بر اساس ساختار گراف را افزایش دهد. در عمل، شبکه توجه گراف تجمع اطلاعات را از طریق فرآیندی شبیه به ارسال پیام در میان گره‌ها تسهیل می‌کند. برخلاف روش‌های مرسوم که ممکن است به میانگین‌گیری یکنواخت یا وزن‌های ثابت برای تجمیع تکیه کنند، شبکه‌های توجه گراف وزن توجه را بر اساس محتوای گره‌ها و روابط آنها که توسط یال‌های گراف تعریف شده‌اند، به‌صورت پویا محاسبه می‌کنند. این انطباق پویا تضمین می‌کند که هر گره می‌تواند به‌طور موثر اطلاعات همسایگان خود را با در نظر گرفتن اهمیت هر رابطه یکپارچه کند (Velickovic et al., 2017).

¹ Transformers

شبکه‌های توجه گراف یکی از روش‌های پیشرفته در پردازش زبان طبیعی هستند که برای تحلیل احساسات مبتنی بر جنبه استفاده می‌شوند. در واقع این شبکه‌ها ترکیبی از شبکه‌های عصبی گرافی^۱ و مکانیزم‌های توجه است که به منظور مدل‌سازی وابستگی‌های نحوی و معنایی بین کلمات در یک جمله به کار می‌رود. چندین مطالعه و پژوهش بر روی استفاده از شبکه توجه گراف در تحلیل احساسات مبتنی بر جنبه انجام شده است؛ به عنوان مثال در مطالعه‌ای از ترکیب شبکه‌های توجه گراف و لایه‌های کانولوشن برای استخراج اطلاعات با فاصله طولانی و اطلاعات محلی استفاده کردند تا هماهنگی بهتری بین جنبه هدف و کلمات زمینه‌ای ایجاد کنند (Wang et al., 2020).

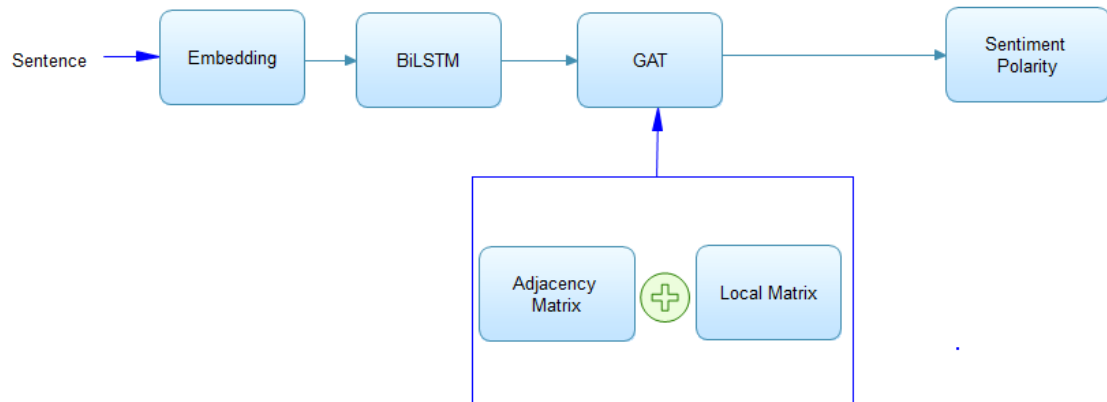
۷. شبکه توجه گراف همراه با درخت تجزیه وابستگی

یکی از موضوعات مهم در تحلیل احساسات مبتنی بر جنبه، مدل کردن روابط نحوی و معنایی بین کلمات جنبه و سایر کلمات متن می‌باشد. پژوهش‌های متعددی در این زمینه انجام شده است. در سال 2019 «سان و همکاران» پژوهشی ارائه دادند؛ در این پژوهش یک مدل چندسطحی مبتنی بر شبکه‌های توجه گراف برای تحلیل احساسات مبتنی بر جنبه پیشنهاد شد که از گراف‌های نحوی به عنوان بخشی از ورودی استفاده می‌کند تا وابستگی‌های بین کلمات را بهتر مدل کند. نتایج نشان داد که مدل پیشنهادی عملکرد بهتری نسبت به مدل‌های پایه دارد (Sun et al., 2019). همچنین در سال ۲۰۲۰، یک مدل جدید ترکیبی از شبکه توجه گراف و حافظه بلند-کوتاه مدت برای تحلیل احساسات مبتنی بر جنبه پیشنهاد شد که از اطلاعات نحوی و توالی زمانی کلمات استفاده کرده بود. این مدل توانست در مجموعه داده‌های مختلف نتایج برتری کسب کند (Yuan et al., 2020). در یک مطالعه برای مدل کردن وابستگی‌ها با فاصله طولانی بین کلمات جنبه و متن از شبکه‌های توجه گراف و شبکه‌های کانولوشن گراف استفاده شده است (Li et al., 2021). در مطالعه‌ای دیگر به علت چالش برانگیز بودن شناسایی و مدل کردن روابط پیچیده بین کلمات مثل شروط و تضاد، مدل تطبیقی تقسیم‌بندی نشانگر شبکه کانولوشن گراف برای تحلیل احساسات مبتنی بر جنبه ارائه شده است؛ در این مدل از درخت تجزیه وابستگی، درخت تجزیه اجزاء، شبکه توجه گراف و شبکه کانولوشن گراف برای مدل کردن روابط نحوی و معنایی بین کلمات در دو ماژول مجزا استفاده شده است (Wang et al., 2023). در پژوهشی دیگر از تعبیه کلمه برت^۲ و شبکه حافظه بلند-کوتاه مدت دوجهته^۳ برای ایجاد بازنمایی کلمات متن و جنبه و از ماتریس مجاورت و ماتریس وزن محلی به ترتیب برای مدل‌سازی روابط نحوی بین کلمات و توجه به اطلاعات محلی استفاده شده است (Gan & Tang, 2021). این بازنمایی‌ها به عنوان ورودی شبکه توجه گراف استفاده شده‌اند. ساختار کلی مدل ارائه شده در این پژوهش در شکل ۲ قابل مشاهده است.

¹ Graph Neural Networks

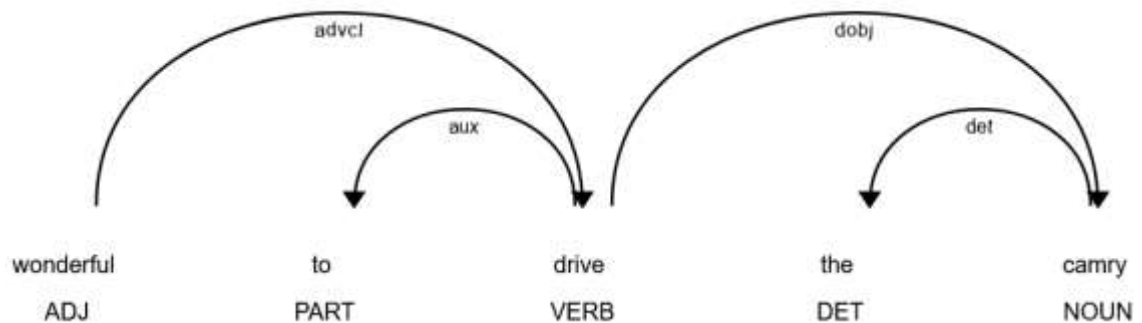
² Bert Embedding

³ Bi-LSTM



شکل ۲. ساختار مدل ASGAT (Gan & Tang, 2021)

طبق روابط نحوی موجود در جمله قابل مشاهده در شکل ۳ در ماتریس مجاورت کلماتی که باهم ارتباط نحوی دارند عدد ۱ و آن‌هایی که رابطه‌ای ندارند عدد صفر می‌گیرند (جدول ۳).



شکل ۳. درخت وابستگی نحوی (Gan & Tang, 2021)

جدول ۳. ماتریس مجاورت (Gan & Tang, 2021)

	wonderful	to	drive	the	Camery
wonderful	0	0	1	0	0
to	0	0	1	0	0
drive	1	1	0	0	1
the	0	0	0	0	1
Camery	0	0	1	1	0

در ماتریس محلی یک عدد به عنوان محدوده تعیین می‌شود و کلماتی که در این محدوده نسبت به کلمه جنبه قرار داشته باشند عدد ۱ و بقیه ۰ می‌گیرند. در جدول ۴ عدد ۳ به عنوان محدوده مجاز تعیین شده و کلماتی که با کلمه جنبه Camery کمتر از

عدد ۳ فاصله دارند در ماتریس در ستون کلمه جنبه مقدار ۱ را دریافت می کنند. این پژوهش با استفاده از روش های ذکر شده توانست نتایج بهتری روی 5 مجموعه داده نسبت به مدل های قبلی بدست آورد (Gan & Tang, 2021).

جدول 4. ماتریس وزن محلی (Gan & Tang, 2021)

Column1	wonderful	to	drive	the	Camery
wonderful	0	0	0	0	0
to	0	0	0	0	1
drive	0	0	0	0	1
the	0	0	0	0	1
Camery	0	1	1	1	1

«یوان و همکاران» در پژوهشی مدل ASGAN را ارائه دادند که یک پله فراتر قدم برداشت و علاوه بر بهبود مدلسازی روابط نحوی برای تحلیل احساسات مبتنی بر جنبه، ارتباطات نحوی را براساس نوع آن ها وزن دهی کرد. در واقع علاوه بر استفاده از درخت تجزیه وابستگی به عنوان ورودی گراف، از تعبیه موقعیت^۱ به عنوان یک ویژگی اضافی در بازنمایی کلمات برای افزایش قدرت یادگیری روابط نحوی مدل و از تعبیه یال^۲ برای تمایز قائل شدن بین انواع ارتباطات نحوی استفاده کرد (Yuan et al., 2022). در مدل ارائه شده در این مطالعه، از سه ماژول رمزگذار زمینه آگاه از فاصله^۳ برای ایجاد بازنمایی کلمات، شبکه توجه گراف نحوی^۴ برای یادگیری روابط بین کلمات و پالایش بازنمایی^۵ برای بهبود نتایج تحلیل احساسات استفاده شده است. همچنین همانطور که گفته که این مقاله ویژگی تعبیه موقعیت را ارائه کرده است؛ این ویژگی، برداری است که حداکثر فاصله نحوی مسیر بین هر کلمه تا کلمات جنبه را می شمارد. در شکل ۴ کلمه sitting area جنبه دو عبارتی است. برای محاسبه فاصله بین کلمه the تا این کلمه، دورترین فاصله نحوی یعنی فاصله مسیر نحوی بین کلمه the تا sitting که برابر با ۴ است را در نظر می گیرد. مقدار این بردار برای کل جمله برابر است با $P = \{4, 3, 2, 2, 2, 0, 0\}$ ؛ طول بردار برابر تعداد کلمات در جمله است و مقدار هر عضو فاصله نحوی آن کلمه تا دورترین کلمه عبارت جنبه را نشان می دهد.

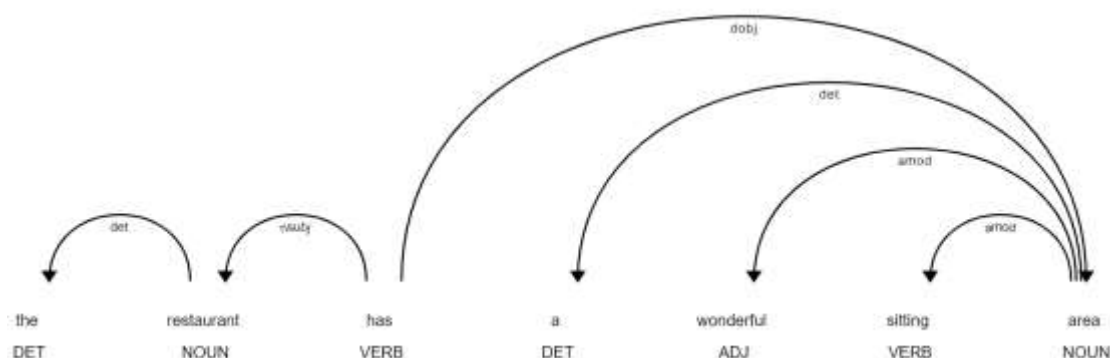
¹ Position Embedding

² Edge Embedding

³ Distance Aware Context Encoder

⁴ Syntactic Graph Attention Network

⁵ Representation Refinement



شکل ۴. وابستگی نحوی جمله نمونه (Yuan et al., 2022)

برای متمایز کردن روابط نحوی با نام یکسان نیز از وزن‌های توجه و ادات سخن گره منبع رابطه استفاده شده است (Yuan et al., 2022). در پژوهشی دیگر یک چهارچوب یکپارچه برای بهبود تحلیل احساسات مبتنی بر جنبه در وظایف مختلف استخراج جنبه، تعیین قطبیت و ... تعریف شده است که شامل یک مدل معنایی چندلایه کانولوشن گراف برای درک روابط معنایی و یک مدل نحوی چند لایه برای یادگیری وابستگی‌های نحوی می‌باشد. این پژوهش از دو مکانیزم توجه که یکی نوع رابطه و وابستگی را مدل می‌کند و دیگری برجسب‌های ادات سخن^۱ را برای تشخیص حدود جنبه و نظر کدگذاری می‌کند (Shi et al., 2023). در یک پژوهش یک شبکه توجه گراف وابستگی عبارت برای تحلیل احساسات مبتنی بر جنبه پیشنهاد شده که یک شبکه گراف توجه رابطه‌ای براساس گراف وابستگی عبارت، تجمیع یال‌های وابستگی جهت‌دار و اطلاعات عبارت، ساخته شده است. این پژوهش توانسته اطلاعات نحوی و وابستگی‌های کوتاه‌برد بلند واژه را به تصویر بکشد و تشخیص قطبیت جنبه را بهبود دهد (Wu et al., 2022). در پژوهشی دیگر برای در نظر گرفتن دانش حوزه هر متن، یک شبکه گراف وابستگی آگاه از دانش را براساس نمودار وابستگی شامل دانش دامنه، برجسب‌های وابستگی و مسیر نحوی پیشنهاد کرده و بهبود قابل توجهی را در تحلیل احساسات مبتنی بر جنبه نشان داده است (Wu et al., 2023). همچنین در پژوهشی از شبکه توجه گراف رابطه‌ای برای کدگذاری درخت وابستگی نحوی و از برجسب‌های وابستگی برای بهبود بازنمایی جنبه استفاده شده است (Wang et al., 2024). در مطالعه‌ای دیگر برای رفع مشکلات عدم حساسیت به اطلاعات ساختار نحوی، پیچیدگی تجزیه وابستگی نحوی و عدم بهره‌برداری از دانش احساسات خارجی، یک مدل مبتنی بر شبکه توجه گراف پیشنهاد شده که ترکیب نحو، معنا و دانش را در برمی‌گیرد (Zhang et al., 2023).

۸. جهت‌گیری‌های تحقیقاتی آینده

تحلیل احساسات مبتنی بر جنبه در زبان انگلیسی به نتایج خوبی دست پیدا کرده و تا حدودی به بلوغ رسیده است؛ اما در زبان فارسی شکاف‌های زیادی در این زمینه وجود دارد؛ برای مثال مجموعه داده برجسب گذاری شده در زبان فارسی کم است. همچنین مدل‌های زبانی یادگیری عمیق به اندازه کافی برای زبان فارسی بهینه نشده‌اند و اغلب مدل‌های موجود به مجموعه داده‌های محدود و با کیفیت پایین تکیه دارند. بنابراین ایجاد و برجسب‌گذاری مجموعه داده‌های بزرگتر و متنوع‌تر برای این زبان از اهمیت بالایی برخوردار است. همچنین چند معنایی کلمات فارسی پیچیدگی در تحلیل احساسات را افزایش می‌دهد. علاوه بر آن ساختار نحوی و معنایی جملات در زبان فارسی می‌تواند بسیار پیچیده باشد و نیاز به مدل‌های پیشرفته‌تری برای تحلیل دقیق احساسات دارد. پژوهش‌های محدودی در مورد استفاده از درخت‌های وابستگی نحوی و شبکه‌های توجه گراف برای زبان فارسی وجود دارد. بنابراین باید از تکنیک‌های

¹ Part Of Speech

بررسی شده در این پژوهش مثل تعبیه موقعیت، تعبیه یال و ساختار گراف برای بهبود مدل سازی روابط نحوی و معنایی در زبان فارسی استفاده کرد.

۹. نتیجه گیری و مباحثه

در این مقاله مروری، ما به بررسی کاربرد درخت وابستگی نحوی و شبکه های توجه گراف در تحلیل احساسات مبتنی بر جنبه پرداختیم. نتایج مطالعات مختلف نشان می دهد که ترکیب این دو روش می تواند بهبود قابل توجهی در دقت و کارایی مدل های تحلیل احساسات مبتنی بر جنبه به همراه داشته باشد. استفاده از درخت وابستگی نحوی، اطلاعات ساختاری جملات را به مدل ها ارائه می دهد و به درک بهتر روابط بین کلمات کمک می کند. از سوی دیگر، شبکه های توجه گراف با قابلیت مدل سازی روابط پیچیده بین کلمات، توانایی مدل را در تشخیص دقیق تر احساسات مربوط به هر جنبه افزایش می دهند. همچنین نوآوری های پژوهش های گذشته در اضافه کردن ویژگی هایی به مدل ها برای تشخیص بهتر وابستگی ها با فاصله طولانی بررسی شد که می تواند در زبان هایی مثل فارسی که هنوز نیاز به توسعه های بسیاری در این زمینه دارد استفاده شود.

۱۰. منابع

- Chen, D., & Manning, C. D. (2014). A fast and accurate dependency parser using neural networks. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* ,
- Dozat, T., & Manning, C. D. (2016). Deep biaffine attention for neural dependency parsing. *arXiv preprint arXiv:1611.01734* .
- Gan, L., & Tang, Q. (2021). Graph attention network with dependency parsing for aspect-level sentiment classification. *Proceedings of the 4th International Conference on Computer Science and Software Engineering* ,
- Gartner. (2021). *STAMFORD, Conn* .
- Gu, S., Zhang, L., Hou, Y., & Song, Y. (2018). A position-aware bidirectional attention network for aspect-level sentiment analysis. *Proceedings of the 27th international conference on computational linguistics* ,
- Kiperwasser, E., & Goldberg, Y. (2016). *Transactions of the Association for Computational Linguistics*, 4, 313-327 .
- Li, R., Chen, H., Feng, F., Ma, Z., Wang, X., & Hovy, E. (2021). *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* ,
- Liu, B. (2022). Springer Nature .
- Liu., B. (2012). *Morgan & Claypool Publishers* .
- Ma, X., & Hovy, E. (2016). *arXiv preprint arXiv:1603.01354* .
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). *PLoS Med*, 6(7), e1000097 .
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., de clerq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Zafra, S. M., & Eryigit, G. (2016). *SemEval-2016 Task 5: Aspect Based Sentiment Analysis* .
- Shi, J., Li, W., Bai, Q., Yang, Y., & Jiang, J. (2023). *Neurocomputing*, 557, 126730 .
- Sun, K., Zhang, R., Mensah, S., Mao, Y., & Liu, X. (2019, November). In K. Inui, J. Jiang, V. Ng, & X. Wan, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* Hong Kong, China.
- Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). *stat*, 1050(20), 10-48550 .
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). *arXiv preprint arXiv:1710.10903* .
- Wang, K., Shen, W., Yang, Y., Quan, X., & Wang, R. (2020). *arXiv preprint arXiv:2004.12362* .

- Wang, P., Tao, L., Tang, M., Wang, L., Xu, Y., & Zhao, M. (2024). *Engineering Applications of Artificial Intelligence*, 133, 108101 .
- Wang, P., Tao, L., Tang, M., Zhao, M., Wang, L., Xu, Y., Tian, J., & Meng, K. (2023). *Knowledge-Based Systems*, 270, 110559 .
- Wu, H., Huang, C., & Deng, S. (2023). *Information Fusion*, 92, 289-299 .
- Wu, H., Zhang, Z., Shi, S., Wu, Q., & Song, H. (2022). *Knowledge-Based Systems*, 236, 107736 .
- Yuan, L., Wang, J., Yu, L.-C., & Zhang, X. (2020). Proceedings of the 1st conference of the asia-pacific chapter of the association for computational linguistics and the 10th international joint conference on natural language processing ,
- Yuan, L., Wang, J., Yu, L.-C., & Zhang, X. (2022). *IEEE Transactions on Artificial Intelligence* .
- Zhang, S., Gong, H., & She, L. (2023). *Knowledge-Based Systems*, 275, 110662 .

Abstract

Aspect-based sentiment analysis is one of the important branches of natural language processing, which aims to extract and analyze sentiments related to certain aspects of text. Considering the complexity of human language and the need to better understand the connections between words, the use of advanced methods such as graph attention network and dependency parser tree has attracted a lot of attention. Using graph structures and attention mechanism, the attention graph network has the ability to model complex relationships between words, while the dependency parsing tree better understands syntactic and semantic relationships by parsing the text into dependent structures. In this review, we examine the use of these two methods in aspect-based sentiment analysis. The results of this study show that in the English language, the combination of graph attention network and dependency parser tree has been able to significantly improve the accuracy and efficiency of aspect-based sentiment analysis systems in the various fields of social events, electronic goods, products and services of restaurants and different brands. Also, the results show the increasing trend of research conducted in recent years, which shows the importance of this topic.

Keywords: Natural language processing, big data, sentiment analysis, syntactic dependency parsing tree, attention graph network, systematic review