

تحول در تحلیل احساسات: یک ساختار نوین در معماری مدل زبانی

۱- حسین غلامعلی نژاد ۲- طه‌ورا رضانی مقدم

۱- استادیار، گروه مهندسی کامپیوتر، دانشگاه بزرگمهر قائنات

۲- کارشناس واحد فناوری اطلاعات، دانشگاه بزرگمهر قائنات

h_gholamalinejad@buqaen.ac.ir

چکیده

شناسایی احساسات از متن یک وظیفه حیاتی در پردازش زبان طبیعی است که می‌تواند منافع زیادی در حوزه‌های مختلفی مانند هوش مصنوعی، تعامل انسان و کامپیوتر و غیره داشته باشد. احساسات، افکار فیزیولوژیکی هستند که در پاسخ‌های انسانی به رویدادها بروز می‌کنند. تحلیل این احساسات بدون تغییرات صورت و صوت بسیار مهم است و نیازمند رویکرد نظارتی برای تفسیر دقیق احساسات است. علیرغم این چالش‌ها، در نظر گرفتن احساسات انسانی ضروری است، زیرا آنها به تدریج از طریق متن‌های نامناسب در پلتفرم‌های شبکه‌های اجتماعی مانند فیس‌بوک، توئیتر و دیگران منتقل می‌شوند. در این مقاله، سیستمی برای شناسایی احساسات مبتنی بر تکنیک‌های یادگیری عمیق پیشنهاد شده است. برای این هدف، سیستم طراحی شده بر روی دو مجموعه داده آزمایش شد: `Tweeter_en_db`: به زبان انگلیسی و `Snappfood` به زبان فارسی. این روش با استفاده از شبکه‌های عصبی بازگشتی و `LSTM` روی این دو مجموعه داده مختلف ارزیابی شد تا نشان دهد چگونه می‌توان دقت بالای در طبقه‌بندی احساسات به دست آورد. ارزیابی‌های دقیق نشان داد که سیستم پیش‌بینی احساسات با مدل `CNN` به دقت 90.70% و با مدل `LSTM` به دقت 88.47% روی مجموعه داده انگلیسی `Tweeter_en_db` رسید، و در مجموعه داده فارسی `Snappfood`، دقت 82.90% با مدل `CNN` و 85.08% با مدل `LSTM` حاصل شد. مقایسه نتایج روش‌های پیشنهادی با روش‌های قبلی روی مجموعه داده `Tweeter_en_db` نشان‌دهنده برتری 8% روش‌های پیشنهادی نسبت به روش‌های قبلی بود، در حالی که در مجموعه داده `Snappfood` برتری 2% مشاهده شد.

کلمات کلیدی: شبکه‌های عصبی کانولوشنی، مدل‌های زبانی بزرگ، تشخیص احساس، `LSTM`.

۱. مقدمه

مدل‌های زبانی بزرگ (LLMs) تحولات قابل توجهی در پردازش زبان طبیعی و هوش مصنوعی ایجاد کرده‌اند. این سیستم‌ها توانایی قابل توجهی در فهم و تولید زبان انسان دارند و در کاربردهایی مانند چت‌بات‌ها، تولید محتوا، ترجمه و تحلیل متن‌های متنی مورد استفاده قرار می‌گیرند. مدل‌های زبانی از سیستم‌های اولیه به سمت مدل‌های پیشرفته‌ای که امروزه مشاهده می‌کنیم، تکامل یافته‌اند. در دهه‌های ۱۹۵۰ و ۱۹۸۰، روش‌های مبتنی بر قواعد و آمار کارایی کافی نداشتند، اما با ظهور شبکه‌های عصبی و یادگیری عمیق در دهه ۲۰۰۰، پیشرفت‌های قابل توجهی حاصل شد.

مدل‌هایی مانند `Word2Vec` و به‌خصوص معماری ترنسفورمر، تحولات بزرگی در پردازش معنایی و مدیریت زبان ایجاد کرده‌اند. از سال ۲۰۲۰ میلادی، توسعه اندازه و مقیاس این مدل‌ها، باعث بهبود عملکرد و تاثیر چشمگیری در تعاملات انسان و ماشین شده است. مدل‌های زبانی بزرگ در کاربردهای گسترده‌ای مانند چت‌بات‌ها، تولید محتوا، ترجمه و تحلیل احساسات کاربرد دارند.

مدل‌های زبانی بزرگ در حوزه‌های متعددی کاربرد دارند:

- چت‌بات‌ها و دستیارهای مجازی: مدل‌های زبانی بزرگ در ایجاد چت‌بات‌هایی که قادر به برقراری ارتباط طبیعی با کاربران، پاسخ

به سوالات و انجام وظایف مختلف هستند، کمک می‌کنند. [۱-۳]

- تولید محتوا: این مدل‌ها می‌توانند مقالات، داستان‌ها و دیگر اشکال محتوا را به طور خودکار تولید کنند. [۴]
- ترجمه زبان: LLMs در بهبود کیفیت ترجمه‌های ماشینی و افزایش دقت ترجمه‌ها مؤثر بوده‌اند. [۵]
- تحلیل احساسات: این مدل‌ها قادر به شناسایی و تحلیل احساسات و دیدگاه‌های موجود در داده‌های متنی هستند. [۶]

در بخش بعدی، در بخش دوم، چندین پژوهش قبلی مرور می‌شود. سپس در بخش سوم، مدل پیشنهادی معرفی می‌گردد. پس از آشنایی با ساختارها، در بخش چهارم، آزمایش‌های اجرا شده ارائه می‌شود. این بخش، به معرفی مجموعه‌های داده، سخت‌افزار مورد استفاده در انجام آزمایش‌ها و نتایج حاصله اختصاص یافته است. در نهایت، در بخش پنجم، به جمع‌بندی نهایی می‌رسیم.

۲. کارهای مرتبط پیشین

در [۷]، مدل perc-DETECTLLM پیشنهاد شده است که ترکیبی از مدل‌های باینری و بیزی برای تخمین درصد متن تولیدشده توسط LLM‌ها است. این مدل با استفاده از ۵۰۰,۰۰۰ نمونه از دامنه‌های مختلف برای آموزش و ارزیابی، به دقت ۷۲.۱٪ و نرخ یادآوری ۸۱.۰٪ در شناسایی متون تولیدشده توسط LLM‌ها دست یافت.

مقاله [۸] روش جدیدی برای شناسایی متن و تحلیل احساسات در داده‌های چندرسانه‌ای مانند تصاویر و ویدئوها بررسی می‌کند. این روش، MTR-SAM، شامل شبکه LK-PAN برای بهبود شناسایی متن و یک مدل تحلیل احساسات بهینه‌شده است. با استفاده از رمزگشای CTC با GTC، سرعت شناسایی ۳۱.۷۷٪ افزایش یافت و دقت در مجموعه داده‌های ویدئویی ۱۰.۷۸٪ بهبود یافت. این روش می‌تواند در تحلیل نظرات عمومی و شناسایی اطلاعات مضر در وبسایت‌ها کمک کند.

مقاله [۹] استفاده از مدل‌های شبکه عصبی بازگشتی (RNN)، به‌ویژه حافظه کوتاه‌مدت بلند (LSTM)، برای تحلیل احساسات در متون را بحث می‌کند. نویسندگان بر تأثیر LSTM در حفظ اطلاعات بلندمدت و کوتاه‌مدت در مقایسه با مدل‌های RNN سنتی تأکید دارند. نتایج نشان می‌دهد که LSTM دقت و نرخ یادآوری بالاتری دارد. همچنین، این پژوهش چالش‌های تحلیل متن چینی را بررسی کرده و بر اهمیت توسعه مدل‌های زبانی مؤثر تأکید دارد.

در مقاله [۱۰]، مدل جدیدی به نام TISRI برای تحلیل احساسات مبتنی بر جنبه در داده‌های چندرسانه‌ای (ABMSA) معرفی شده است. این مدل احساسات مرتبط با جنبه‌های مختلف در نمونه‌های چندرسانه‌ای را شناسایی کرده و مشکلات مدل‌های قبلی، از جمله نادیده گرفتن روابط معنایی بین متن و تصاویر، را برطرف می‌کند. مدل TISRI از یک ماژول برای شناسایی همبستگی‌های معنایی استفاده می‌کند و نتایج نشان می‌دهد که عملکرد رقابتی در دو مجموعه داده توپیتر دارد.

مقاله [۱۱] تکنیک‌های یادگیری ماشین برای شناسایی و تحلیل احساسات را بررسی می‌کند. نویسندگان با مرور ۱۲۳ مقاله، روندهای پژوهشی، الگوریتم‌های محبوب و چالش‌ها را تحلیل کرده و نیاز به مرور جامع در این زمینه را برجسته می‌کنند.

در [۶]، چالش‌های تحلیل احساسات در متون غیرانگلیسی بحث شده است. نویسندگان مطالعه‌ای ارائه می‌دهند که از یک مدل ترکیبی شامل ترنسفورمرها و مدل‌های زبانی بزرگ (LLMs) برای تحلیل احساسات متون به زبان‌های مختلف استفاده می‌کند و دقتی بیش از ۸۶٪ گزارش می‌دهد.

مقاله دیگری [۱۲] تأثیر بازخورد مشتریان بر رفتار مصرف‌کنندگان در خریدهای آنلاین را بررسی می‌کند. با استفاده از مدل ARIMA، روندهای آینده در امتیازدهی ستاره‌ای پیش‌بینی شده و تأثیر بررسی‌های منفی بر رفتار مصرف‌کنندگان تحلیل شده است.

مطالعه دیگری [۹] استفاده از مدل‌های زبانی بزرگ (LLMs) برای استخراج داده‌های خاص از توپیت‌ها و شناسایی مسائل مطرح‌شده توسط شهروندان را بررسی می‌کند. مدل‌های BERT و BERTweet برای طبقه‌بندی توپیت‌ها استفاده شده‌اند و مدل BERT بالاترین دقت (۹۰.۹٪) را در شناسایی مسائل به دست آورده است.

۳. مدل پیشنهادی

برای این پژوهش، دو نوع شبکه طراحی شد: یکی مبتنی بر کانولوشن و دیگری مبتنی بر LSTM. برای هر یک از این انواع، ساختارهای مختلفی پیشنهاد و بررسی شد. در ادامه، این ساختارها تشریح می‌شوند.

۳.۱. مدل مبتنی بر کانولوشن

مدل پیشنهادی شامل یک معماری شبکه عصبی کانولوشنی (CNN) با استفاده از لایه‌های مختلف است. اجزای این مدل به شرح زیر است:

۱. لایه ورودی: این لایه مسئول دریافت داده‌های ورودی است و به‌عنوان نقطه شروع پردازش داده‌ها عمل می‌کند. این لایه می‌تواند داده‌های متنی، بصری یا صوتی را بپذیرد.
۲. لایه تعبیه: این لایه برای تبدیل داده‌های ورودی (به‌ویژه متن) به بردارهای با ابعاد پایین‌تر استفاده می‌شود. این بردارها ویژگی‌های معنایی کلمات را نشان داده و به مدل کمک می‌کنند تا روابط بین کلمات را بهتر درک کند.

۳. لایه LSTM: LSTM نوعی شبکه عصبی بازگشتی (RNN) است که برای پردازش توالی‌های زمانی طراحی شده است. این لایه می‌تواند اطلاعات بلندمدت را حفظ کرده و به یادگیری روابط پیچیده در داده‌های ترتیبی کمک کند.
۴. لایه‌های Conv1D: این لایه‌ها برای استخراج ویژگی‌ها از داده‌های ترتیبی استفاده می‌شوند. هر لایه کانولوشنی به‌طور مؤثری ویژگی‌های محلی را شناسایی کرده و مدل را قادر به تشخیص الگوهای خاص می‌کند.
۵. لایه‌های Max Pooling: این لایه‌ها به کاهش ابعاد ویژگی‌ها و فشردن داده‌سازی اطلاعات کمک می‌کنند. با انتخاب حداکثر مقدار از مناطق خاص، این لایه‌ها به کاهش بار محاسباتی و جلوگیری از بیش‌برازش کمک می‌کنند.
۶. لایه‌های Dropout: این لایه‌ها برای جلوگیری از بیش‌برازش در طول آموزش استفاده می‌شوند. با غیرفعال کردن تصادفی تعدادی از نورون‌ها در هر دوره آموزشی، این لایه‌ها به کاهش وابستگی مدل به ویژگی‌های خاص کمک می‌کنند.
۷. لایه Flatten: این لایه خروجی چندبعدی از لایه‌های قبلی را به یک بردار یک‌بعدی تبدیل می‌کند. این تبدیل برای ورود به لایه‌های کاملاً متصل بعدی ضروری است.
۸. لایه‌های Dense: این لایه‌ها به پردازش نهایی داده‌ها و پیش‌بینی کمک می‌کنند. هر نورون در این لایه به تمام نورون‌های لایه قبلی متصل است و یادگیری روابط پیچیده‌تر را تسهیل می‌کند.
۹. لایه خروجی: این لایه نهایی مدل است که خروجی نهایی (مانند طبقه‌بندی یا پیش‌بینی) را تولید می‌کند. این لایه معمولاً شامل تعداد نورون‌هایی برابر با تعداد کلاس‌ها برای وظایف طبقه‌بندی است.

۳.۲. مدل مبتنی بر LSTM

مدل پیشنهادی مبتنی بر LSTM ترکیبی از لایه‌های LSTM و Dense است.

۱. نام مدل: "LSTM" نشان‌دهنده استفاده از لایه‌های LSTM برای پردازش توالی‌های زمانی است.
۲. لایه ورودی: برای دریافت داده‌های ورودی طراحی شده است.
۳. Conv1D: ویژگی‌ها را استخراج می‌کند.
۴. Dense^۱ و Dense^۲: داده‌ها را پردازش کرده و پیش‌بینی می‌کنند؛ این‌ها لایه‌های کاملاً متصل را نشان می‌دهند.
۵. واحدهای LSTM: تعداد واحدهای LSTM به حفظ اطلاعات بلندمدت کمک می‌کند، با نمونه‌هایی از ۶۴ و ۲۵۶ واحد.
۶. قابلیت آموزش لایه تعبیه: نشان می‌دهد که آیا لایه تعبیه قابل آموزش است یا خیر.
۷. لایه خروجی: خروجی نهایی را تولید می‌کند.

۴. آزمایش‌ها و نتایج

برای ارزیابی اثربخشی الگوریتم پیشنهادی، از دو مجموعه داده استفاده شد. مجموعه داده اول، Tweeter_en_db، شامل توییت‌ها و احساسات مرتبط با آن‌ها است که شامل ۱۷۵,۴۳۵ نمونه آموزشی و ۴۳,۸۵۹ نمونه آزمایشی، در مجموع ۲۱۹,۲۹۴ نمونه است. در این مجموعه داده، داده‌ها به سه کلاس تقسیم شده‌اند: مثبت، منفی و طبیعی. مجموعه داده دوم، Snappfood، شامل نظرات مشتریان درباره غذا، کیفیت غذا و کیفیت خدمات اسنپ‌فود است که شامل ۵۶,۷۰۰ نمونه آموزشی و ۷,۰۰۰ نمونه آزمایشی است. در این مجموعه داده، داده‌ها به دو کلاس تقسیم شده‌اند: بد و خوب.

آزمایش‌ها با استفاده از سیستمی با پردازنده Core i9-۱۲۹۰۰، پردازنده گرافیکی Nvidia GeForce RTX ۳۰۶۰ و ۳۲ گیگابایت رم انجام شد. زبان اجرا پایتون و کتابخانه مورد استفاده Keras بود.

نتایج به‌دست‌آمده از اعمال داده‌های Tweeter_en_db به شبکه‌های پیشنهادی و شبکه‌های استاندارد معرفی‌شده قبلی در جدول زیر نشان داده شده است.

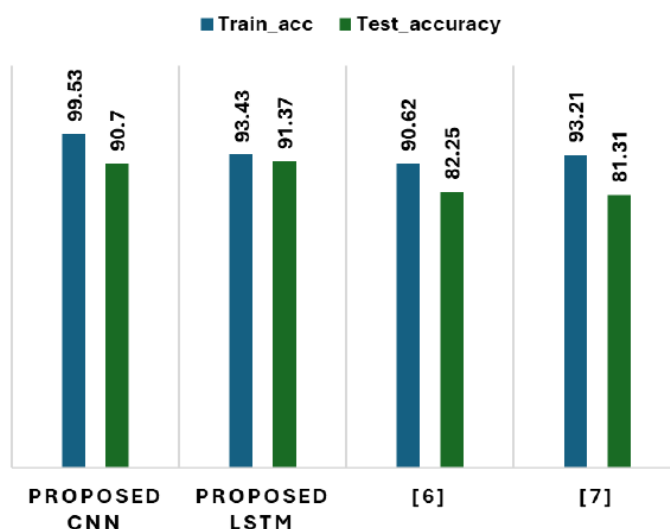
جدول ۱. نتایج روی مجموعه داده Tweeter_en_db

زیان آزمایشی دقت آزمایشی زیان آموزشی دقت آموزشی نام شبکه

CNN پیشنهادی	۹۹.۵۳	۰.۰۱۸۶	۹۰.۷۰	۰.۵۷۲۷
LSTM پیشنهادی	۹۳.۴۳	۰.۱۹۵۸	۹۱.۳۷	۰.۲۵۲۴
[۶]	۹۰.۶۲	۰.۱۹۶۲	۸۲.۲۵	۰.۷۶۵۲
[۷]	۹۳.۲۱	۰.۲۹۷۳	۸۱.۳۱	۰.۸۶۲۱

همان طور که نشان داده شده است، CNN پیشنهادی توانایی های آموزشی قابل توجهی را نشان می دهد و به دقت آموزشی نزدیک به کامل ۹۹.۵۳٪ و زبان آموزشی پایین دست یافته است که نشان دهنده برازش عالی به داده های آموزشی است. با این حال، LSTM پیشنهادی، اگرچه در معیارهای آموزشی عملکرد کمی پایین تر دارد، در دقت آزمایشی (۹۱.۳۷٪) و زبان آزمایشی پایین تر (۰.۲۵۲۴) برتری دارد و توانایی آن در تعمیم به داده های جدید و نادیده را نشان می دهد.

هر دو مدل پیشنهادی به طور قابل توجهی از شبکه های استاندارد قبلی که دقت آزمایشی زیر ۸۳٪ داشتند، عملکرد بهتری نشان دادند و کارایی آنها را در وظایف تحلیل احساسات در زمینه داده های توپیتر برجسته کردند. کارهای آینده می توانند بهینه سازی های بیشتری را بررسی کرده و نقاط قوت هر دو معماری را برای عملکرد بهتر ترکیب کنند.



شکل ۱. نتایج روی مجموعه داده Tweeter_en_db

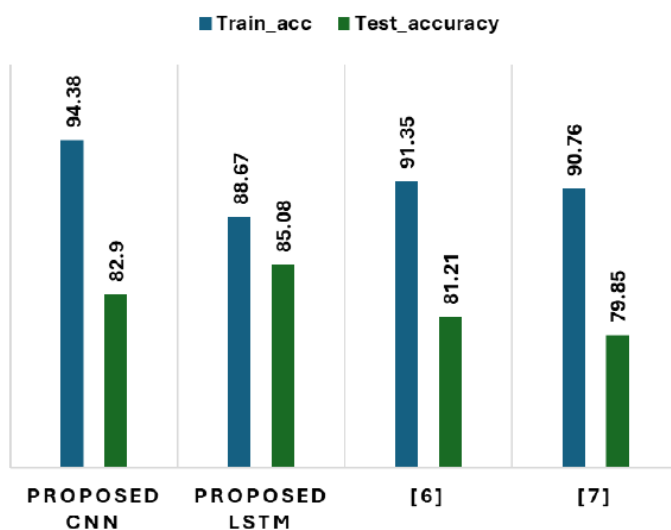
نتایج به دست آمده از اعمال داده های Snappfood به شبکه های پیشنهادی و شبکه های استاندارد معرفی شده قبلی در جدول زیر نشان داده شده است.

جدول ۲. نتایج روی مجموعه داده Snappfood

زبان آزمایشی	دقت آموزشی	زبان آموزشی	دقت آموزشی	نام شبکه
۰.۴۷۱۴	۸۲.۹۰	۰.۱۶۹۰	۹۴.۳۸	CNN پیشنهادی
۰.۳۸۲۷	۸۵.۰۸	۰.۲۸۹۵	۸۸.۶۷	LSTM پیشنهادی
۰.۶۵۷۲	۸۱.۲۱	۰.۳۴۶۷	۹۱.۳۵	[۶]
۰.۸۵۴۷	۷۹.۸۵	۰.۴۶۹۸	۹۰.۷۶	[۷]

همان طور که نشان داده شده است، CNN پیشنهادی عملکرد قوی ای در طول آموزش نشان می دهد و به دقت آموزشی ۹۴.۳۸٪ و زبان آموزشی پایین ۰.۱۶۹۰ دست یافته است. با این حال، LSTM پیشنهادی قدرت خود را در عملکرد آزمایشی با دقت ۸۵.۰۸٪ و زبان آزمایشی پایین تر ۰.۳۸۲۷ نشان می دهد که نشان دهنده تعمیم بهتر به داده های نادیده است.

هر دو مدل پیشنهادی از شبکه های استاندارد که دقت آزمایشی زیر ۸۲٪ داشتند، عملکرد بهتری نشان دادند و کارایی معماری های پیشنهادی در ارزیابی دقیق احساسات در مجموعه داده مورد نظر را برجسته کردند. تحقیقات آینده می توانند بر بهینه سازی بیشتر این مدل ها، احتمالاً از طریق روش هایی مانند منظم سازی، تنظیم پارامترهای بیشینه یا ترکیب مزایای هر دو معماری برای نتایج بهتر، تمرکز کنند.



شکل ۲. نتایج روی مجموعه داده Snappfood.

۵. نتیجه گیری

در این مطالعه، یک معماری نوین برای تحلیل احساسات با استفاده از مدل‌های زبانی بزرگ (LLMs) و تکنیک‌های یادگیری عمیق ارائه شد. سیستم پیشنهادی ما به‌طور مؤثری شبکه‌های عصبی کانولوشنی (CNN) و شبکه‌های حافظه کوتاه‌مدت بلند (LSTM) را برای شناسایی احساسات در مجموعه داده‌های متنوع ادغام می‌کند. نتایج به‌دست‌آمده از آزمایش روی مجموعه داده‌های Tweeter_en_db و Snappfood کارایی رویکرد ما را نشان می‌دهد و به نرخ‌های دقت چشمگیر ۹۰.۷۰٪ و ۸۸.۴۷٪ برای مدل CNN و ۹۱.۳۷٪ و ۸۵.۰۸٪ برای مدل LSTM دست یافته است.

بهبودهای قابل توجه عملکرد نسبت به روش‌های سنتی، پتانسیل معماری ما را در طبقه‌بندی دقیق احساسات در متون انگلیسی و فارسی برجسته می‌کند. توانایی مدل CNN در برتری در دقت آموزشی و مدل LSTM در نشان دادن تعمیم برتر در دقت آزمایشی، نقاط قوت ترکیب هر دو معماری را نشان می‌دهد. این رویکرد دوگانه نه تنها استحکام طبقه‌بندی احساسات را افزایش می‌دهد، بلکه بینش‌هایی در مورد پویایی‌های شناسایی احساسات در داده‌های متنی ارائه می‌دهد.

مراجع

1. Casheekar, A., et al., A contemporary review on chatbots, AI-powered virtual conversational agents, ChatGPT: Applications, open challenges and future research directions. Computer Science Review, ۲۰۲۴, ۵۲: p. ۱۰۰۶۳۲.
2. Kumar, V., et al. Large-Language-Models (LLM)-Based AI Chatbots: Architecture, In-Depth Analysis and Their Performance Evaluation. in International Conference on Recent Trends in Image Processing and Pattern Recognition. ۲۰۲۳. Springer.
3. Løvås, S.E., AI Chatbots in Health: Implementing an LLM-Based Solution to Promote Physical Activity. ۲۰۲۴, UiT Norges arktiske universitet.
4. Kamath, U., et al., LLMs in Production, in Large Language Models: A Deep Dive: Bridging Theory and Practice. ۲۰۲۴, Springer. p. ۳۱۵-۳۷۳.
5. Pan, R., et al., Understanding the effectiveness of large language models in code translation. CoRR, ۲۰۲۳.
6. Miah, M.S.U., et al., A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM. Scientific Reports, ۲۰۲۴, ۱۴(۱): p. ۹۶۰۳.
7. Goel, A., et al., Beyond Binary: A Multi-Class Approach to Large Language Model-Generated Text Percentage Estimation. ۲۰۲۴.
8. Liu, X., et al., MTR-SAM: Visual Multimodal Text Recognition and Sentiment Analysis in Public Opinion Analysis on the Internet. Applied Sciences, ۲۰۲۳, ۱۳(۱۲): p. ۷۳۰۷.

-
-
۹. dos Santos, V.G., et al. Identifying Citizen-Related Issues from Social Media Using LLM-Based Data Augmentation. in International Conference on Advanced Information Systems Engineering. ۲۰۲۴. Springer.
 ۱۰. Zhang, T., et al., Text-image semantic relevance identification for aspect-based multimodal sentiment analysis. PeerJ Computer Science, ۲۰۲۴. ۱۰: p. e۱۹۰۴.
 ۱۱. Alslaity, A. and R. Orji, Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions. Behaviour & Information Technology, ۲۰۲۴. ۴۳(۱): p. ۱۳۹-۱۶۴.
 ۱۲. Wang, Z., et al., Llm for sentiment analysis in e-commerce: A deep dive into customer feedback. Applied Science and Engineering Journal for Advanced Research, ۲۰۲۴. ۳(۴): p. ۸-۱۳.