

گزارش نهایی پروژه

تحلیل احساسات مبتنی بر جنبه با استفاده از تصاویر مصنوعی

نام: محبوبه کریمی

شماره دانشجویی: ۴۰۴۱۲۶۳۴

درس: پردازش زبان های طبیعی

استاد: دکتر غیاثی راد

۱. مقدمه

تحلیل احساسات مبتنی بر جنبه (Aspect-Based Sentiment Analysis یا ABSA) به شناسایی قطبیت احساس نسبت به جنبه‌های مشخصی که در یک متن ذکر شده‌اند می‌پردازد. در حالی که اکثر روش‌های ABSA صرفاً بر داده‌های متنی متکی هستند، پژوهش‌های اخیر به بررسی ادغام اطلاعات بصری برای بهبود درک احساسات پرداخته‌اند.

در این پروژه، یک سیستم نمونه‌ی چندوجهی برای ABSA ارائه می‌شود که از یک مقاله منتشرشده در EMNLP الهام گرفته شده است. با توجه به محدودیت منابع محاسباتی و در دسترس نبودن کد آموزش کامل، هدف پروژه بازتولید کامل مقاله نیست، بلکه پیاده‌سازی یک پایپ‌لاین اجرایی و قابل‌استفاده است که ایده‌ی اصلی تحلیل احساسات چندوجهی با استفاده از تصاویر مصنوعی را نشان دهد.

۲. نمای کلی سیستم

چارچوب پیشنهادی از یک پایپ‌لاین چندوجهی متوالی تشکیل شده است که شامل سه ماژول اصلی است:

۱. مولد تصویر مصنوعی

۲. ماژول پالایش تصویر

۳. دسته‌بند چندوجهی تحلیل احساسات

هر ماژول به صورت مستقل پیاده سازی شده است تا شفافیت و ماژولار بودن سیستم حفظ شود.

۳. تولید تصاویر مصنوعی

فایل: generator.py

این ماژول تصاویر مصنوعی را بر اساس موارد زیر تولید می کند:

- متن نظر (Review)

- جنبه (Aspect) هدف

- برچسب احساس (مثبت، منفی یا خنثی)

برای این منظور از یک مدل انتشار از پیش آموزش دیده (مانند Stable Diffusion) و پرامپت های متنی طراحی شده استفاده می شود. تصاویر تولید شده ذخیره شده و در مراحل بعدی برای پالایش یا دسته بندی مورد استفاده قرار می گیرند.

هدف:

شبیه سازی زمینه ی بصری که بتواند به صورت ضمنی نشانه های مرتبط با احساس نسبت به جنبه مورد نظر را منتقل کند.

۴. پالایش بصری

فایل: refiner.py

ماژول پالایش با هدف افزایش ارتباط تصاویر تولید شده با جنبه ی مورد نظر طراحی شده است. در صورت در دسترس بودن ابزارهای لازم، از روش های قطعه بندی و فیلترگذاری معنایی برای تمرکز بر نواحی مرتبط با جنبه استفاده می شود.

این مرحله اختیاری است و برای اجرای مطالعه ی Ablation (حذف اجزا) به منظور مقایسه تصاویر پالایش شده و پالایش نشده در نظر گرفته شده است.

هدف:

بررسی این موضوع که آیا پالایش تصاویر مصنوعی می تواند پیش بینی احساسات در سطح جنبه را بهبود دهد یا خیر.

۵. دسته‌بندی چندوجهی ABSA

فایل: classifier.py

دسته‌بند، نمایش‌های متنی و بصری را برای پیش‌بینی احساس ترکیب می‌کند:

- ویژگی‌های متنی با استفاده از مدل از پیش‌آموزش‌دیده BERT استخراج می‌شوند.
 - ویژگی‌های بصری با استفاده از رمزگذار تصویری CLIP استخراج می‌شوند.
 - نمایش تلفیقی حاصل برای پیش‌بینی قطبیت احساس (مثبت، خنثی یا منفی) استفاده می‌شود.
- این ماژول صرفاً در حالت استنتاج (Inference) اجرا می‌شود و هیچ آموزش یا Fine-tuning اضافی انجام نمی‌گیرد.

۶. تنظیمات آزمایشی و پروتکل ارزیابی

با توجه به محدودیت منابع و نبود مجموعه‌داده‌های برچسب‌خورده‌ی بزرگ، آزمایش‌ها روی تعداد محدودی نمونه و با استفاده از مدل‌های از پیش‌آموزش‌دیده و وزن‌های ثابت انجام شده‌اند.

اهداف اصلی آزمایش‌ها عبارت‌اند از:

- بررسی صحت عملکرد پایپ‌لاین چندوجهی
 - تحلیل تأثیر اطلاعات بصری
 - نمایش امکان‌پذیری معماری پیشنهادی
- هیچ ادعایی مبنی بر رقابت با روش‌های پیشرفته‌ی روز ارائه نمی‌شود.

۷. خط مبنا و مطالعات Ablation

۷.۱ خط مبنا فقط-متن (Text-only Baseline)

برای اندازه‌گیری تأثیر اطلاعات بصری، یک خط مبنا فقط-متن پیاده‌سازی شده است.

در این حالت، معماری دسته‌بند چندوجهی بدون تغییر باقی می‌ماند، اما ورودی تصویری با یک تنسور صفر جایگزین می‌شود.

این روش امکان مقایسه‌ی منصفانه بین تحلیل فقط-متنی و چندوجهی را بدون تغییر ظرفیت مدل فراهم می‌کند.

۷.۲ مطالعه Ablation: تصاویر خام در برابر تصاویر پالایش‌شده

در این مطالعه، دو نوع ورودی تصویری مقایسه می‌شوند:

۱. تصاویر مصنوعی خام تولیدشده توسط مولد

۲. تصاویر پالایش‌شده پس از ماژول پالایش

این مقایسه فقط در مرحله استنتاج انجام شده و شامل هیچ‌گونه آموزش مجدد نیست. هدف، بررسی تأثیر پالایش بصری بر دقت پیش‌بینی احساسات در سطح جنبه است.

۸. نتایج و بحث

مشاهدات تجربی نشان می‌دهند که:

- معماری چندوجهی قادر است اطلاعات بصری مصنوعی را بدون مشکل ساختاری ادغام کند.
- خط مبنای فقط-متن یک حد پایین قابل اعتماد برای عملکرد مدل فراهم می‌کند.
- پالایش بصری در برخی موارد که جنبه‌ها دارای ارتباط بصری واضح هستند، می‌تواند بهبود کیفی ایجاد کند، هرچند این بهبود به کیفیت تصاویر مصنوعی وابسته است.

این نتایج امکان‌پذیری پایپ‌لاین پیشنهادی را تأیید می‌کنند، اما اهمیت کیفیت بالای زمینه‌ی بصری را نیز نشان می‌دهند.

۹. محدودیت‌ها

این پروژه دارای محدودیت‌های زیر است:

- Fine-tuning نظارت‌شده برای مدل‌های انتشار انجام نشده است.

- ارزیابی کمی در مقیاس بزرگ امکان پذیر نبوده است.
- سیستم ارائه شده باید به عنوان یک نمونه‌ی مفهومی (Proof-of-Concept) در نظر گرفته شود، نه یک مدل نهایی بهینه شده.

۱۰. ارتباط با مقاله EMNLP

مقاله اصلی (Aspect-based Sentiment Analysis via Synthetic Image Generation) از تولید تصاویر مصنوعی آگاه از احساس با آموزش نظارت شده و ارزیابی گسترده استفاده می کند.

در مقابل، این پروژه نسخه ای ساده شده و اجرایی از ایده ی اصلی را پیاده سازی کرده است. برخی اجزا عمداً ساده سازی شده اند تا امکان اجرای آفلاین، بازتولیدپذیری و سازگاری با منابع محدود فراهم شود.

۱۱. نتیجه گیری

در این پروژه، یک چارچوب چندوجهی ماژولار و قابل اجرا برای تحلیل احساسات مبتنی بر جنبه ارائه شد که از تصاویر مصنوعی استفاده می کند. علی رغم محدودیت ها، سیستم نشان می دهد که ایده ی استفاده از اطلاعات بصری در ABSA از نظر منطقی و فنی امکان پذیر است و می تواند مبنایی برای توسعه های آینده باشد.