

# خلاصه تحلیلی مقاله: «Towards Semantic Integration of Opinions: Unified Opinion Concepts Ontology and «Extraction Task

نویسندگان: Gaurav Negi, Dhairya Dalal, Omnia Zayed, Paul Buitelaar — LDK 2025

## (۱) مسئله اصلی چیست و چرا مهم است؟

مسئله اصلی مقاله این است که در ادبیات استخراج نظر، هر فرمول بندی فقط بخشی از مؤلفه های نظر را پوشش می دهد. مثلاً ABSA عمدتاً روی جنبه و قطبیت متمرکز است، ASTE سه تایی «هدف-احساس-قطبیت» را استخراج می کند، و ACOS چهار مؤلفه را هدف می گیرد. اما مؤلفه هایی مانند «دارنده نظر»، «شدت احساس»، «قید» و «دلیل» یا نادیده گرفته می شوند یا به صورت ضمنی در متن باقی می ماند.

این ناهماهنگی باعث می شود خروجی مدل ها از نظر معنا ناقص باشد و قابلیت اتصال به سیستم های عملیاتی را از دست بدهد. برای مثال، اگر کاربر نوشته باشد «باتری برای کارهای سنگین دوام ندارد»، بدون استخراج «قید» و «دلیل»، سیستم نمی فهمد این نظر برای همه کاربران صدق نمی کند و مشکل دقیقاً در سناریوهای خاص رخ می دهد.

اهمیت موضوع دقیقاً همین جاست: تحلیل احساسات صرفاً برچسب مثبت/منفی نیست؛ بلکه یک «دانش ساختاریافته» درباره تجربه کاربران است. هر چه این دانش دقیق تر استخراج شود، تصمیم گیری تجاری، طراحی محصول و حتی تحلیل های اجتماعی قابل اعتمادتر خواهد بود.

## (۲) ورودی ها و خروجی های مدل/سیستم چیست؟

ورودی سیستم یک متن طبیعی است؛ مثل نظر کاربر درباره یک محصول، رستوران، هتل یا کتاب. این متن ممکن است چند نظر هم زمان داشته باشد.

خروجی سیستم در وظیفه UOCE یک نمایش ساختاریافته از مؤلفه های نظر است که به طور کامل تری نسبت به فرمول بندی های قبلی ارائه می شود. مؤلفه های اصلی عبارت اند از:

- Aspect Term (جنبه/هدف)
- Aspect Category (دسته جنبه)

- Target Entity (موجودیت هدف)
- Sentiment Expression (عبارت احساس)
- Sentiment Polarity (قطبیت احساس)
- Sentiment Intensity (شدت احساس: weak < average < strong)
- Opinion Holder Span/Entity (دارنده نظر)
- Opinion Qualifier (قید)
- Opinion Reason (دلیل)

این خروجی به شکل «بردار نظر» قابل ذخیره در پایگاه دانش یا گزارش‌های تحلیلی است و می‌تواند به صورت عملیاتی وارد چرخه تصمیم‌گیری شود.

### ۳) داده مورد استفاده (نوع، منبع، اندازه)

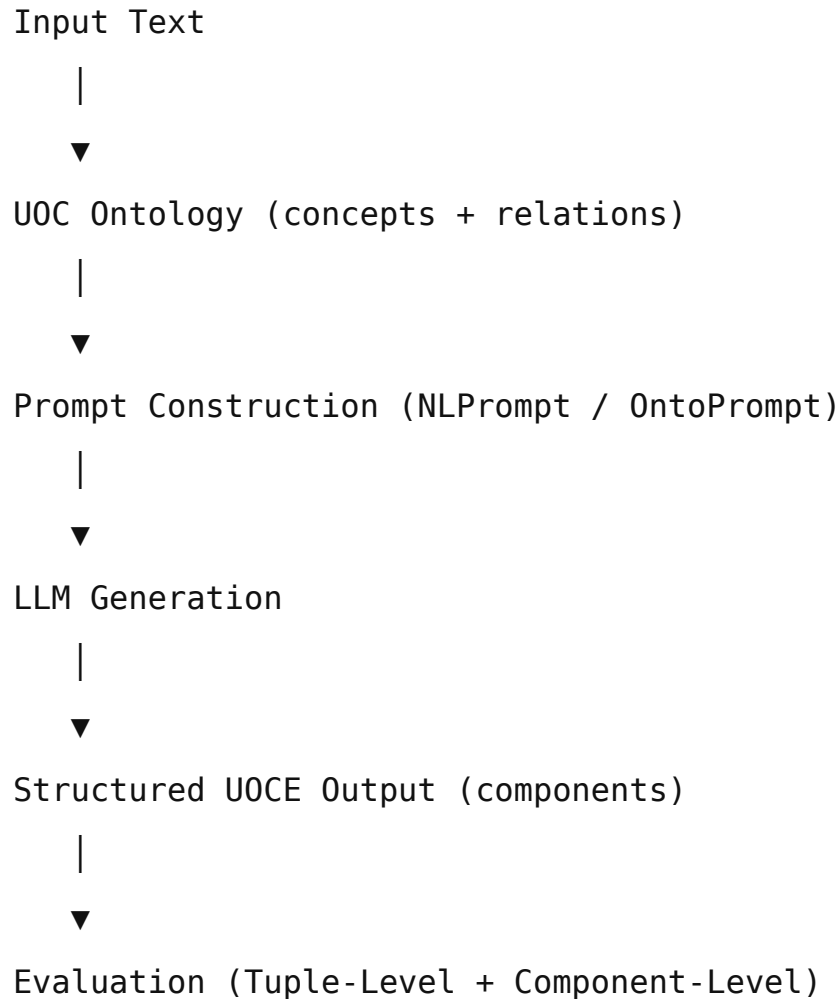
نویسندگان یک دیتاست ارزیابی توسعه داده شده معرفی می‌کنند که بر پایه مجموعه ME23 ساخته شده است. این دیتاست چندداده‌ای است و پنج حوزه (کتاب، پوشاک، هتل، رستوران و لپ‌تاپ) را پوشش می‌دهد. از هر حوزه ۲۰ نمونه انتخاب شده و در مجموع ۱۰۰ نمونه نهایی به عنوان بنچمارک ساخته شده است. سپس برچسب‌های جدید برای مؤلفه‌هایی که در دیتاست اولیه وجود نداشتند اضافه می‌شود؛ از جمله قید، دلیل، شدت احساس و دارنده نظر. این برچسب‌گذاری با اجماع سه متخصص نهایی شده است. بنابراین داده از نظر حجم کوچک است، اما از نظر «غنا معنایی» نسبت به دیتاست‌های رایج غنی‌تر محسوب می‌شود.

### ۴) روش پیشنهادی به زبان ساده + شماتیک یا شبه‌کد

#### ۴-۱) ایده روش

مقاله یک هستان‌شناسی به نام UOC تعریف می‌کند که مجموعه مفاهیم و روابط نظر را به صورت رسمی مشخص می‌کند. UOC سعی می‌کند مفاهیم پراکنده در پژوهش‌های NLP را با ساختارهای نمادین یکپارچه کند. سپس یک وظیفه جدید به نام UOCE معرفی می‌شود که هدف آن استخراج همه این مؤلفه‌ها از متن است.

برای استخراج، از مدل‌های زبانی بزرگ به صورت پرامپت‌محور استفاده می‌شود. دو مسیر اصلی وجود دارد: (۱) پرامپت زبانی (NL Prompt) با توضیح و مثال، و (۲) پرامپت مبتنی بر نمایش هستان‌شناسی (Onto Prompt) که در قالب‌هایی مثل JSON-LD یا OWL ارائه می‌شود. خروجی مدل باید ساختاریافته و مطابق قالب تعریف شده باشد.



## ۲-۳) شبه کد خلاصه

```

for each input text T:
    build prompt P using UOC definitions
    response R = LLM(P, T)
    parse R into components:
        aspect_term, category, target_entity,
        sentiment_expression, polarity, intensity,
        holder_span/entity, qualifier, reason
    compare with gold annotations
  
```

## (۵) نتایج اصلی

نتایج نشان می‌دهد که مدل‌های زبانی بزرگ در استخراج مؤلفه‌های UOCE عملکرد قابل قبول دارند، اما عملکرد به شدت به طراحی پرامپت وابسته است. تغییر ترتیب «تعریف‌ها-مثال‌ها-قالب خروجی» یا نوع نمایش هستان‌شناسی می‌تواند F1 را چند واحد جابه‌جا کند.

در گزارش مقاله، برخی مدل‌ها (مثل GPT-4o و Gemma2-27B) به میانگین F1 حدود ۵۷-۵۸ می‌رسند و مدل‌های کوچک‌تر معمولاً پایین‌تر قرار می‌گیرند. این تفاوت نشان می‌دهد که استخراج کامل مؤلفه‌ها هنوز یک چالش سخت برای مدل‌های سبک است.

یکی از پیام‌های کلیدی مقاله این است که معیار Tuple-Level برای وظایف چندمؤلفه‌ای بسیار سخت‌گیرانه است و ممکن است کارایی سیستم‌ها را کمتر از واقع نشان دهد؛ در مقابل، معیار Component-Level امکان سنجش جزئی مؤلفه‌ها را فراهم می‌کند و تصویر دقیق‌تری ارائه می‌دهد.

## (۶) محدودیت‌ها

محدودیت اصلی، اندازه کوچک دیتاست ارزیابی است که بیشتر برای سنجش مناسب است تا آموزش. علاوه بر این، مدل‌ها آموزش اختصاصی ندیده‌اند و تنها با in-context prompting ارزیابی شده‌اند. همچنین، دامنه‌ها محدود است و پوشش زبان‌ها یا حوزه‌های تخصصی دیگر در حداقل قرار دارد.

از نظر عملی، خروجی‌ها هنوز به شدت به قالب پرامپت وابسته‌اند و برای استفاده صنعتی نیازمند استانداردسازی و مقاومت در برابر تغییرات ورودی هستند.

## (۷) ایده‌های ادامه (Future Work)

- گسترش دیتاست با دامنه‌ها و زبان‌های بیشتر (مثلاً فارسی) و اندازه بزرگ‌تر.
- آموزش یا تنظیم دقیق مدل‌ها برای استخراج پایدارتر مؤلفه‌های UOCE.
- تعریف معیارهای ترکیبی که هم سخت‌گیری داشته باشند و هم حساس به استخراج جزئی باشند.
- اتصال خروجی UOCE به پایگاه‌های دانش یا سیستم‌های تصمیم‌یار برای کاربرد عملی.

- مطالعه اثرات طراحی پرامپت و ارائه راهنماهای استاندارد برای پرامپت نویسی.

## ۸) جمع بندی نهایی

این مقاله یک چارچوب مفهومی و عملی برای یکپارچه سازی معنایی «نظر» در NLP ارائه می کند. با تعریف UOC و UOCE، مسیر روشنی برای استخراج مؤلفه های کامل نظر معرفی شده است. خروجی های ساختاریافته می توانند به سیستم های عملیاتی (از تحلیل تجربه مشتری تا پایش رسانه های اجتماعی) متصل شوند. اگرچه داده کوچک و مسیر هنوز ابتدایی است، اما چارچوب مفهومی مقاله ظرفیت توسعه گسترده در پژوهش و صنعت را دارد.