

# تحليل احساسات مبتنی بر جنبه با استفاده از تولید تصاویر مصنوعی

محبوبه کریمی-۴۰۴۱۲۶۳۴

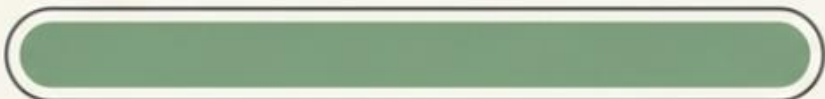
## چرا تحلیل احساسات کلی کافی نیست؟

محدودیت‌های تحلیل تک‌بعدی در مقابل پیچیدگی نظرات واقعی

### تحلیل مبتنی بر جنبه (ABSA)

«پیتزا عالی بود، اما گارسون کند رفتار می‌کرد.»

پیتزا (غذا)



مثبت ✓

گارسون (خدمات)



منفی ✗

### تحلیل سنتی (Traditional Sentiment Analysis)

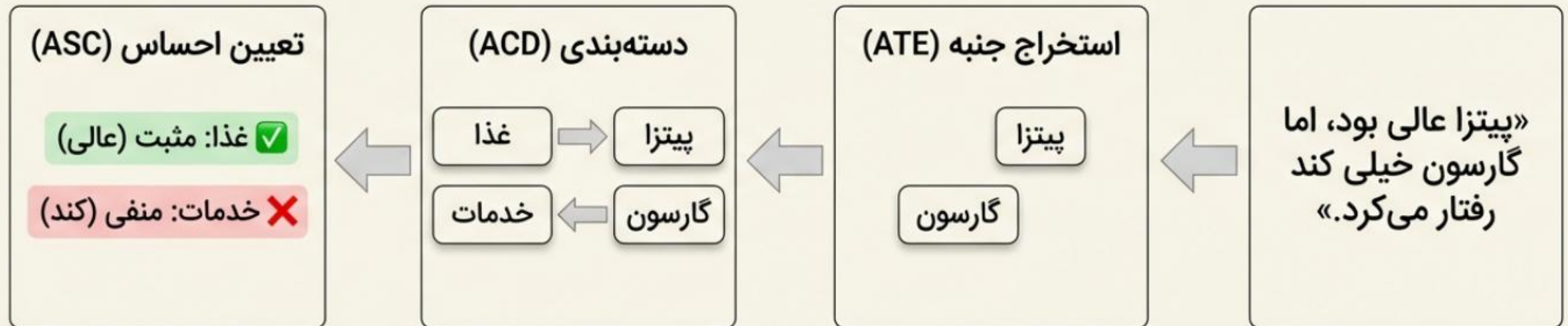
«پیتزا عالی بود، اما گارسون کند رفتار می‌کرد.»



نکات مثبت و منفی همدیگر را خنثی می‌کنند.

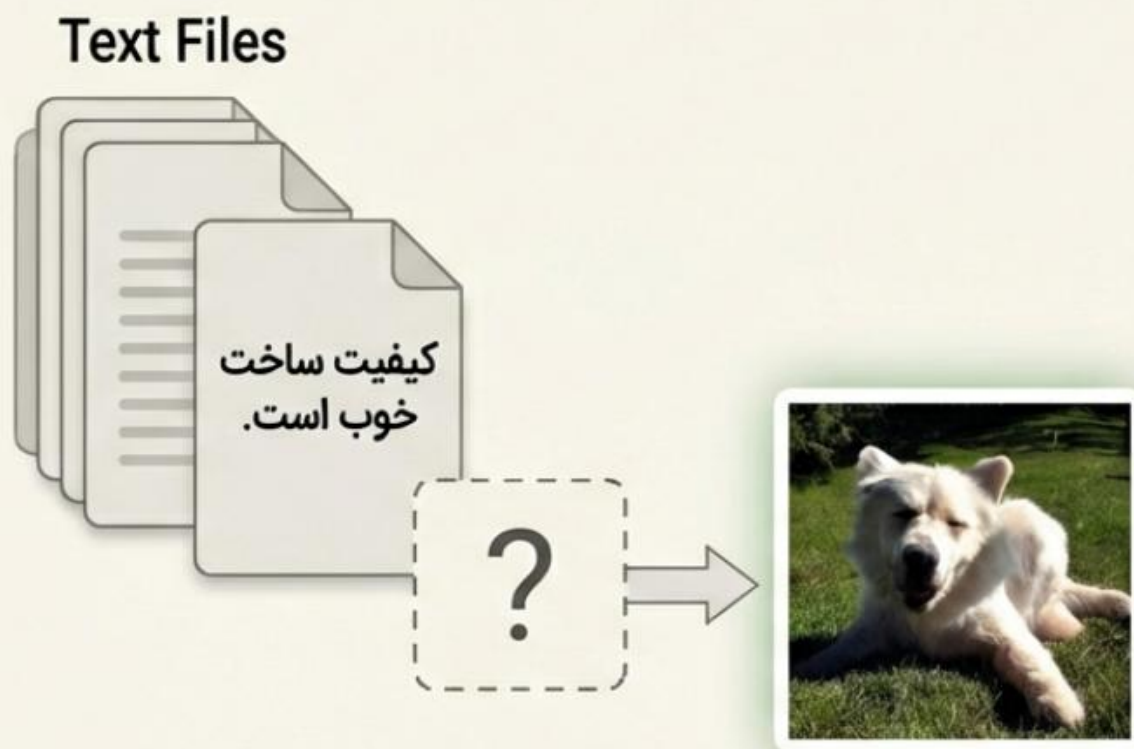
# کالبدشکافی یک سیستم ABSA

سه وظیفه اصلی: استخراج، دسته‌بندی و تعیین احساس



هدف نهایی: تحلیل ریزدانه (Fine-grained) برای تبدیل نظرات کیفی به داده‌های کمی و عملیاتی.

# چالش ابهام در متن و کمبود داده‌های تصویری

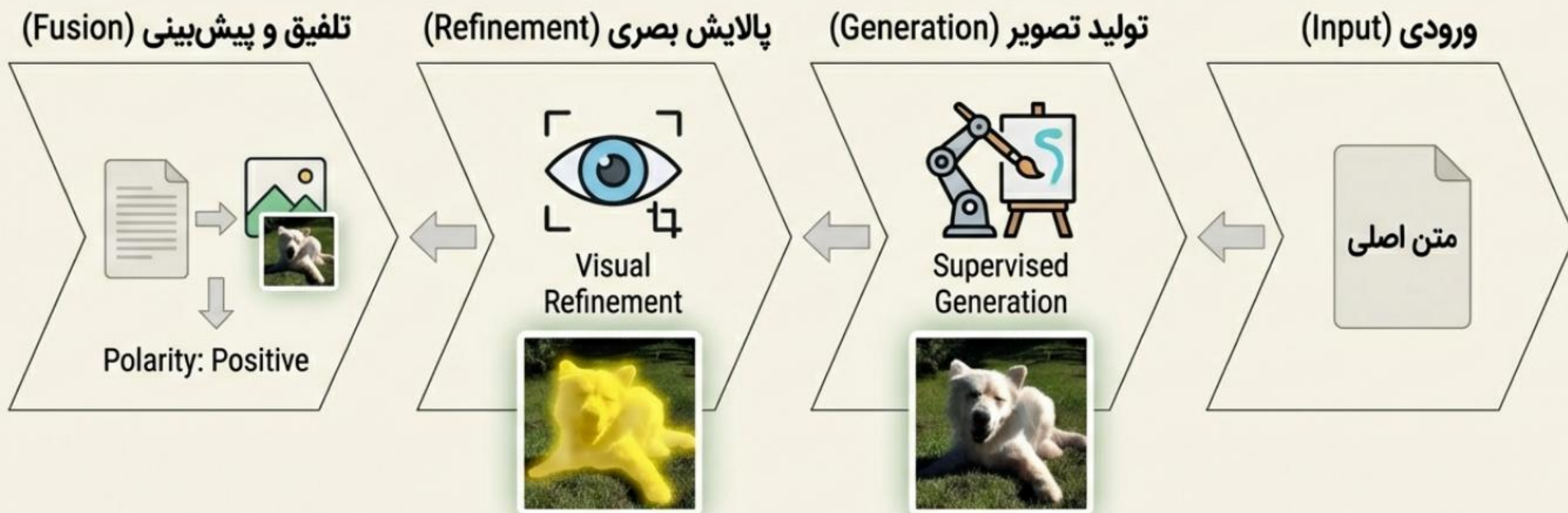


تولید تصویر از روی متن برای رفع ابهام

- متن به تنهایی می‌تواند مبهم باشد (کنایه، جنبه‌های مخفی).
- مدل‌های چندوجهی (Multimodal) دقیق‌ترند، اما اکثر دیتاست‌ها فقط متنی هستند.
- جمع‌آوری تصاویر واقعی برای هر نظر، دشوار یا غیرممکن است.
- راه‌حل پیشنهادی: تولید تصاویر مصنوعی (Synthetic Image Generation)

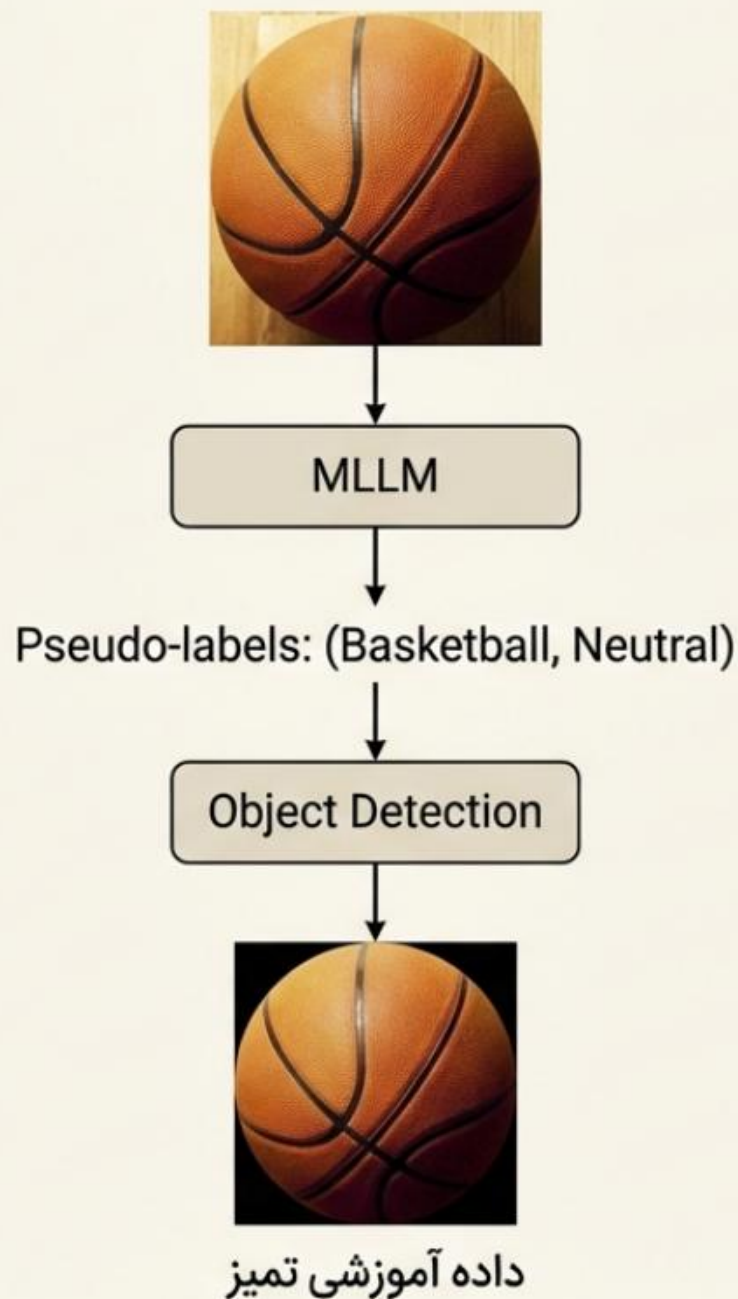


# معماری پیشنهادی: از متن خام تا تصویر مصنوعی



فرآیند کلی: تولید تصویر نظارت‌شده ← پالایش دقیق ← تحلیل چندوجهی

## مرحله ۱: تولید تصویر آگاه از احساس (Sentiment-Aware)



- چالش: تصویر تولید شده باید دقیقاً با «جنبه» و «احساس» متن هماهنگ باشد.

### • روش کار:

۱. استفاده از MLLM برای تولید برچسب‌های مصنوعی.
۲. استفاده از Object Detection برای برش (Crop) ناحیه مرتبط.
۳. تنظیم دقیق (Fine-tuning) مدل Stable Diffusion با تکنیک **LoRA\*\***.

## مرحله ۲: پالایش بصری (Visual Refinement)

حذف نویز و تمرکز بر جزئیات کلیدی

### Before and After

Noisy Image



تصویر اولیه  
(Noisy/Raw)

SAM &  
Masking

Masked Image



ماسک‌گذاری  
(Masked)

Attention

After Image



تصویر نهایی  
(Refined)

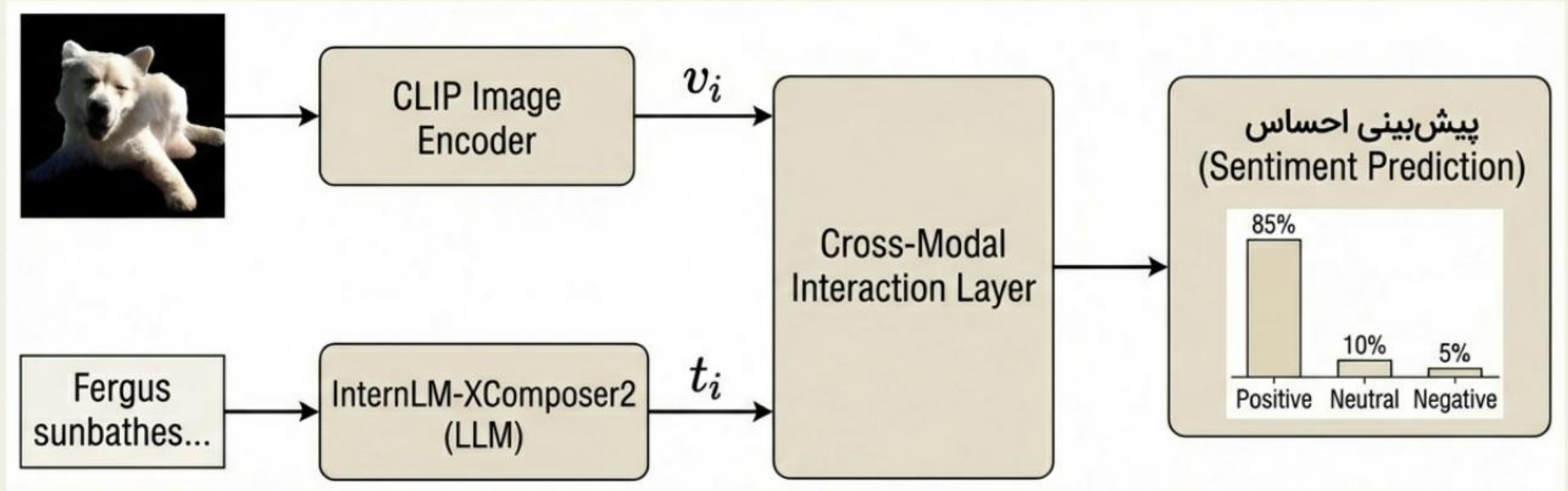
- هدف: حذف پس‌زمینه نامرتب و برجسته‌سازی ویژگی‌های حسی.

### • تکنیک‌ها:

- Fine-Grained Visual Prompting (FGVP): ماسک‌گذاری با مدل SAM.
- Attention Prompt: هدایت توجه مدل با نقشه حرارتی.



## مرحله ۳: تلفیق چندوجهی (Multimodal Fusion)





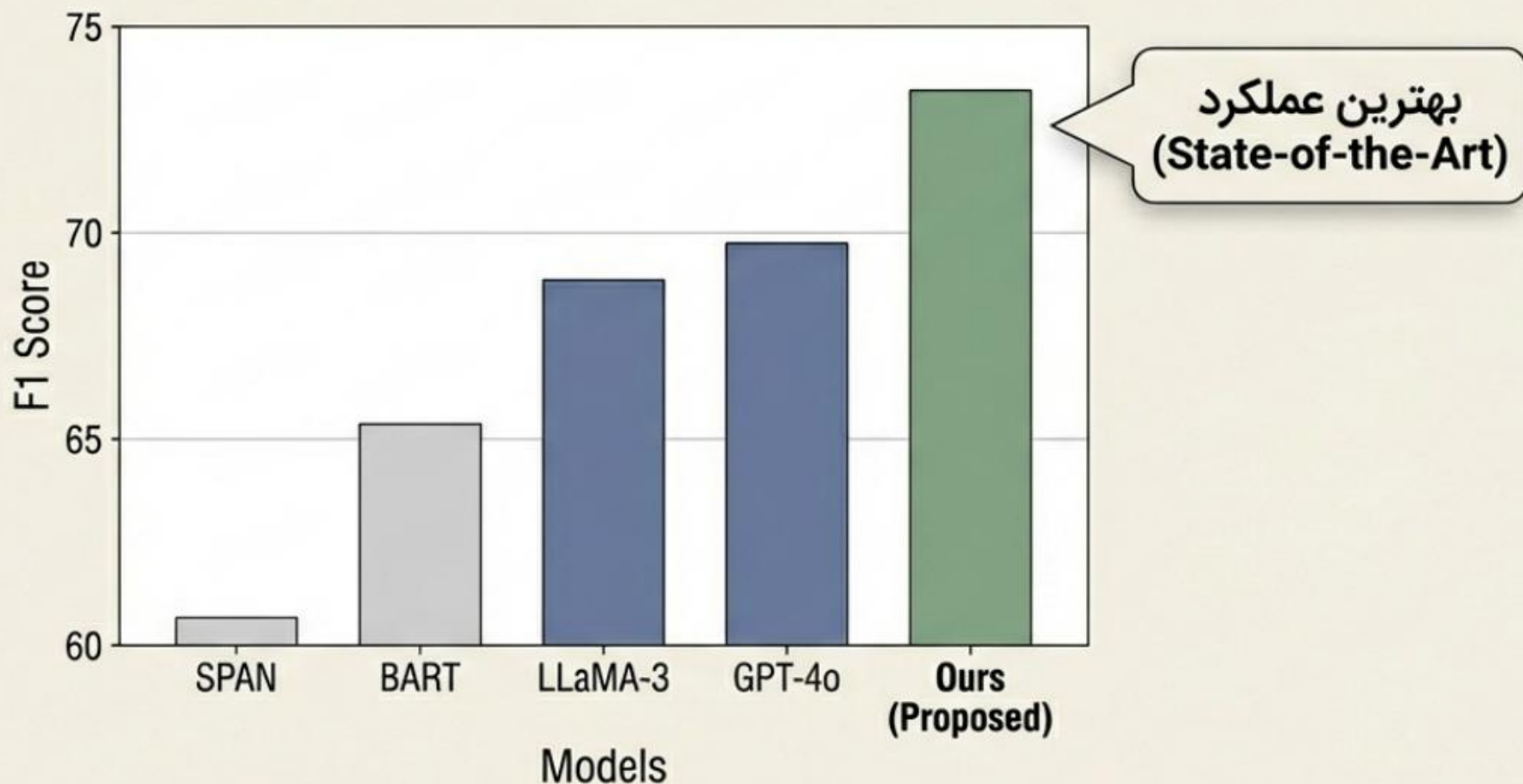
## تنظیمات آزمایش و داده‌ها

دیتاست‌ها	مدل‌های پایه	معیارهای ارزیابی
<b>Twitter-2015</b>	<b>LLMs:</b> GPT-4o, LLaMA-3, Qwen-2.5	<b>Accuracy Micro-F1</b>
<b>Twitter-2017</b>	<b>ABSA Models:</b> FaiMA, OTG, Paraphrase	
شامل تویتهای چندجنبه‌ای (Multi-aspect) و پیچیده.	<b>Standard:</b> BART, SPAN	



# نتایج: برتری تصاویر مصنوعی بر مدل‌های زبانی بزرگ

مقایسه عملکرد F1-Score در دیتاست Twitter-2017



مدل پیشنهادی با  
اختلاف ۳.۷  
واحد GPT-4o را  
پشت سر گذاشت.

# مطالعه موردی: حل ابهام با تصویر



Ryan Howard

Ryan Howard turns **painful lessons** into new [#startup](#)



Path A - Text Only

پیش‌بینی: خنثی (Neutral)

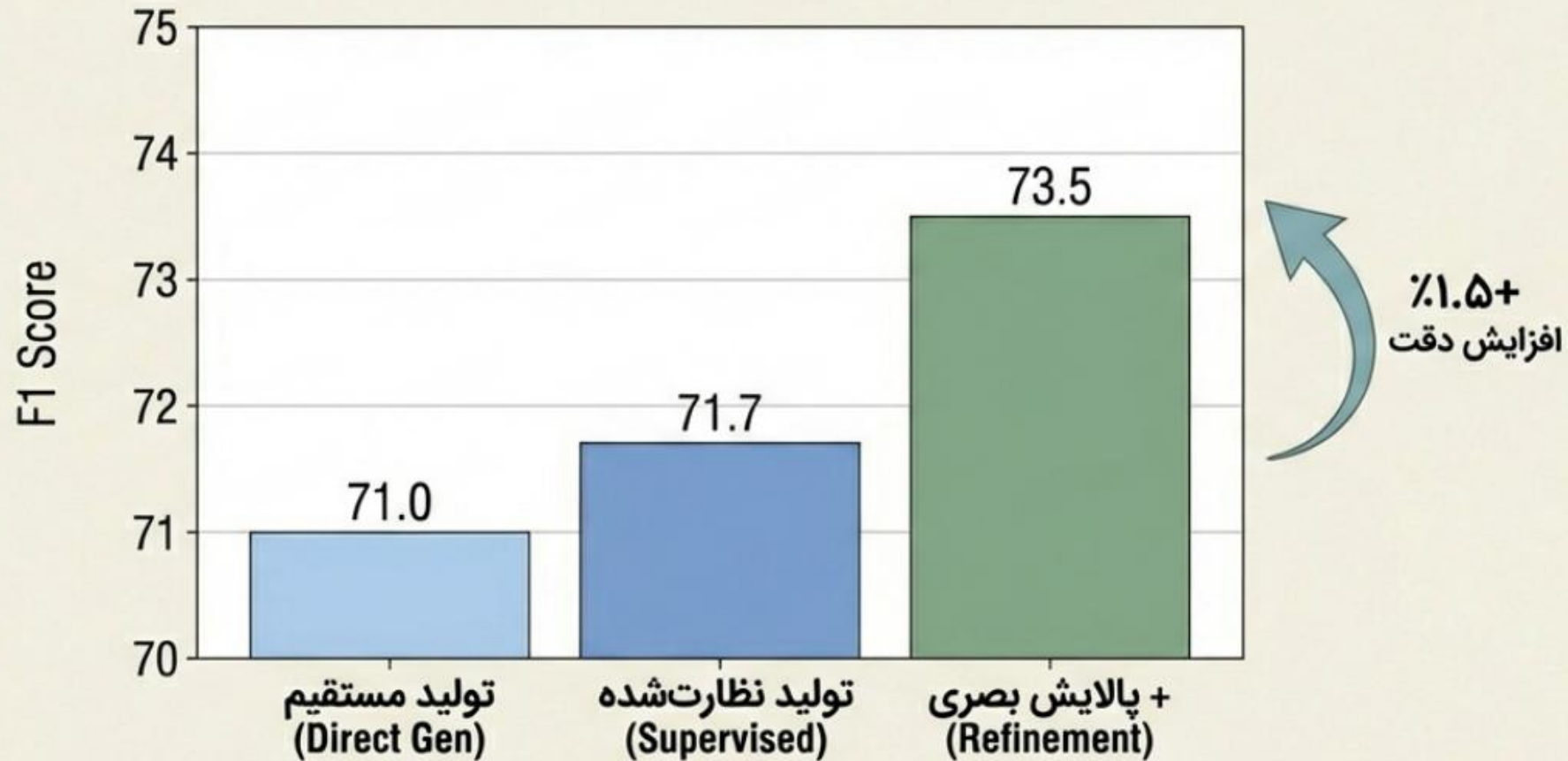


Path B - Multimodal/Ours

پیش‌بینی: مثبت (Positive)

تصویر زمینه را شفاف می‌کند: تویت درباره «موفقیت» است، نه «درد».

## تأثیر پالایش تصویر بر دقت مدل (Ablation Study)



کیفیت تصویر مصنوعی و حذف نویز، نقش کلیدی در عملکرد نهایی دارد.



# چالش‌ها و مسیرهای آینده



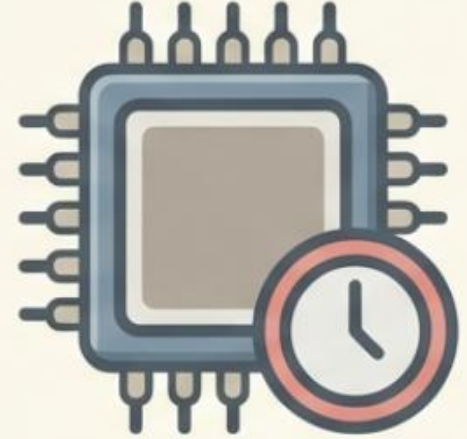
## توسعه زبان فارسی

کاربرد ABSA در زبان فارسی و  
چالش کمبود داده‌های آموزشی.



## تفسیرپذیری (XAI)

چرا مدل این تصویر خاص را تولید  
کرد؟ درک منطق پشت تصویر.



## هزینه محاسباتی

تولید تصویر زمان‌بر است. نیاز به  
بهینه‌سازی برای کاربرد Real-time.