

Рецензия на статью "Нейросетевой подход к обучению векторных представлений мультимодальных данных на основе угловых аддитивных функций потерь"

В статье предлагается новый способ построения мультимодальных эмбедингов с помощью специальной аддитивной функции потерь и двухуровневого трансформера.

Из плюсов хочется отметить

- Актуальность и новизну работы
- Хорошо изложенная и понятная теория
- Подробно описана архитектура решения
- Использование таких модальностей, как табличные данные и категориальные признаки, которые редко встречаются в других статьях, но при этом часто используются в прикладных задачах

К недостаткам можно отнести

- Нет сравнения результатов с аналогичными архитектурами
- Отсутствие таких модальностей, как изображения и аудио
- Эксперименты проводились только на одном датасете
- В бейзлайне присутствует только текстовая модальность

В целом, статья мне показалась интересной, так как построение мультимодальных эмбедингов достаточно сложная задача, и эта тема сейчас активно исследуется