

ArcFace вместо Contrastive-Learning для обучения мультимодальных векторных представлений

Специфика данных

- данные о компаниях
- 1 класс = 1 компания
- основные модальности:
 - текст
 - табличные данные
 - метки
 - связи и взаимодействия

Применение

- retrieval (поиск)
- look-a-like (поиск похожих)
- clusterization (сегментация рынка и выделения идеального профиля клиента)

Требования

- вектор небольшого размера (<512)
- соблюдение семантической близости
- mAP@5 > 60

