

Тэгирование генеративных картинок с помощью VLM

Цель: построить модель, которая обеспечит разнообразие ленты через эффективное тэгирование, генерируемых на основе текстовых промптов и сгенерированного изображения.

Задачи

1. Построить модель выделения ключевых сущностей из картинок и текстовых промптов.
2. Создать систему ранжирования изображений на основе полученных тегов для улучшения разнообразия ленты.

Исследуемая проблема

1. Обеспечение единства и обобщенности тегов без декоративных и несущественных слов. Преобразование специфических обозначений в более общие теги (например, «котенок» в «кот»).

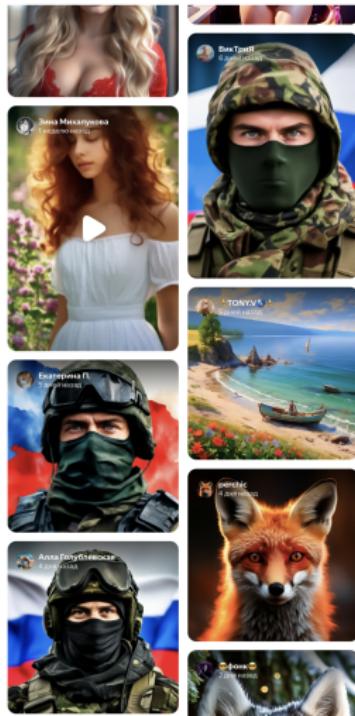
Метод решения

Алгоритм тэгирования картинок и промптов основан на применении визуальноязыковой модели для выделения ключевых сущностей и гиперонимов. Это позволит игнорировать стилистические изменения и будет способствовать созданию обобщенного представления тегов.

Как это выглядит сейчас¹

Типичная лента, до внедрения подобной модели

- 1) Что происходило с персональными рекомендациями на 23 февраля



¹М. Ю. Назарько. Кадр с трудовых будней // Работа, 2024.