

Задача: Создание системы рекомендации статей пользователю, на основании понравившихся.

Решение: Решение с использованием nlp методов из классического ML. Использую CountVectorizer из sklearn. Использую простые Леммы токенов (корни слов) для векторизации. Полученный эмбединг статьи сравниваю просто Эвклидовой метрикой со всеми эмбедингами корпуса статей.

Stack: DS, Python, scipy, sklearn