

# Инструмент для анализа клиентских отзывов "КАПОТ"

Дополнительное задание

# Ход работы

## 1 Сбор отзывов

С помощью сервиса **Dataminer** мы создали базу данных в Excel из собранных нами отзывов. В качестве источников отзывов мы использовали [официальный сайт Тинькофф](#), а также [сервис сбора отзывов с открытых интернет-источников finsber.com](#)

## 2 Первичная обработка

Мы написали код, который с помощью [библиотеки openruhl](#) удаляет компьютерные элементы кодировки текстов из отзывов при первичной записи их в массив

## 3 Анализ и интерпретация

На данном этапе мы воспользовались open source моделью нейронной сети - [pipeline](#) из библиотеки **Transformers**. Преимущество этой библиотеки в том, что она позволяет работать со сложными моделями в пару строк кода. Запустив работу [pipeline](#), **Transformers** проводит обработку заданного текста после чего, на основе обученной модели, выявляет тональность отзыва (NEGATIVE/POSITIVE). После этого тональности сохраняются в исходную базу данных

# Результаты анализа

В ходе работы над дополнительным заданием мы обучили нейросеть выявлять эмоциональную окраску отзыва (тональность), используя собранную нами базу данных из 50 отзывов. В результате **80% отзывов** оценены нейросетью верно. Использование нейросетью рейтингов отзывов поможет получать более точный результат анализа. Для этого требуется дальнейшее ее обучение и доработка

Ранее мы уже представляли идею разделения отзывов по темам. Для этой цели мы предлагаем использовать дополнительную нейросеть, которая будет выделять в каждом отзыве ключевые слова и на их основе разделять отзывы по темам: продукты Тинькофф, работа сотрудников компании, функционирование приложения и сайта