|  |
| --- |
| Секция: цифровые технологии |
| ГБОУ “Школа №1533 “ЛИТ” |
| 119296, г. Москва, Ломоносовский проспект, 16, тел.: (495) 133-2435; E-mail: info@lit.msu.ru |
| Анализ отзывов клиентов на банковские каналы обслуживания и продукты/услуги |
| Рукавица Артем, Федченко Анастасия |
| Класс: 11 |
| 119334, г.Москва, Ленинский пр-кт, 40-153; 119526, г.Москва, пр-кт Вернадского, 97-32 |
| Тел.: (985) 491-35-44; E-mail: [rukavitsa.a@yandex.ru](mailto:rukavitsa.a@yandex.ru); Тел.: (985)855-20-04; E-mail: [nastyaf1504@gmail.com](mailto:nastyaf1504@gmail.com) |
| Научный руководитель: Чамров Михаил Валерьевич, Вице-президент, Директор департамента кредитного бизнеса ПАО Банк «ФК Открытие» |

Цифровые технологии открыли новые возможности и установили абсолютно новые правила игры для компаний и пользователей, в результате чего конкуренция сместилась от создания лучшего продукта/услуги к созданию лучшего клиентского опыта.

Digital technologies have opened up new opportunities and established completely new rules of the game for companies and users; as a result, the competition has shifted from creating a better product/service to creating a better customer experience.

**Актуальностью** нашего продукта является отсутствие в открытом доступе программ/сайтов, занимающихся сравнением клиентского опыта в различных банках. Наша программа упрощает пользователю процесс сравнения нескольких банков по нескольким критериям при помощи различных интерактивных диаграмм и карты банковских отделений.

The relevance of our product is the absence of publicly available programs/ sites that compare customer experience in various banks. Our program makes it easier for users to compare several banks by several criteria using various interactive charts and maps of bank offices.

**Целью** нашей работы является создание web-сайта, позволяющего пользователю сравнить любой банк с рынком в среднем и с лучшими игроками рынка, определить области улучшения и дальнейшего развития каналов обслуживания, продуктов и услуг по определенным критериям: удобство офиса, банкоматы, касса, уровень сервиса, персонал, продукты и услуги, дистанционные каналы обслуживания.

The purpose of our work is to create a website that allows the user to compare any bank with the market on average and with the best market players. Besides­­­, banks can identify areas for improvement and further development of service channels, products and services according to certain criteria: convenience of the office, ATMs, level of service, staff, products and services, remote service channels.

В качестве входных данных мы используем отзывы с картографического сервиса Яндекс.Карты (www.yandex.ru/maps) на банковские отделения. На основе полученных данных создаем модель оценки качества банковских каналов, продуктов и услуг с визуализацией результата.

As input data, we use reviews from Yandex.Maps cartographic service (www.yandex.ru/maps) to bank offices. Based on the collected data, we create a model for estimating the quality of banking channels, products and services with visualization of the result.

Для создания нашего проекта мы воспользовались языком программирования Python, фреймворком для визуализации приложений с использованием машинного обучения streamlit, фреймворком scikit-learn для предиктивного анализа данных, библиотекой plotly для создания интерактивных диаграмм и модулем folium для создания карты банковских отделений.

To create our project, we have used Python as programming language and several frameworks: streamlit (for visualizing apps using machine learning), scikit-learn (for predictive data analysis), plotly (for creating interactive diagrams) and folium (for creating bank offices map).

Кратко опишем, как решается наша задача.

1. Сбор адресов отделений банков с сайта <https://1000bankov.ru>;
2. Сбор отзывов с сайта <https://yandex.ru/maps>;
3. Подготовка отзывов к анализу: удаление стоп-слов, знаков препинания, приведение слов в начальную форму;
4. Ручной лейблинг отзывов на случайной выборке, состоящей из 1600 отзывов, на основе категоризации из «Постановки задачи»;
5. Обучение моделей бинарной классификации на основе пролейбленных отзывов из пункта 4 при помощи метода опорных векторов;
6. Категоризация всех отзывов при помощи обученных моделей;
7. Визуализация: создание горизонтальных гистограмм для отображения средних оценок и важных слов (на основе градиентного бустинга), диаграммы в форматах «паутина» и «торнадо» (с применением критерия согласия Пирсона), интерактивной карты банковских отделений.

**Результатом** нашей работы является сайт, предоставляющий визуализацию проведённого анализа на собранных данных. На нашем сайте пользователю доступно три типа диаграмм и интерактивная карта:

1. горизонтальная гистограмма (для отображения средних оценок и важных слов);
2. диаграмма в формате “паутинка” (для удобного отображения средних оценок по категориям);
3. диаграмма в формате “торнадо” (для отображения всех распределений положительных и отрицательных отзывов по категориям и в среднем, а также только тех распределений, которые сильно отличаются от среднего распределения и от распределения по категориям между банками);
4. интерактивная карта банковских отделений РФ с подробной аналитикой для каждого отделения из базы данных.

The result of our work is the website that provides a visualization of the analysis performed on the collected data. On our website users have access to three types of diagrams and interactive map:

1. Horizontal bar graph (displays average ratings, important words);

2. Radar chart (conveniently displays average ratings by categories);

3. Tornado chart (displays all distributions of positive and negative reviews by category and on average);

4. Interactive map of bank offices of the Russian Federation with detailed analytics for each office from the database (detailed analytics contains average overall rating of bank office and average ratings by each category)

We have created an approach that can be applied to assessment and comparison of different types of service (e.g. banks' & delivery apps, food supermarket chains, etc.). To build analytics, you only need reviews and their categorization (that is, what users write about).

Prospects for further development:

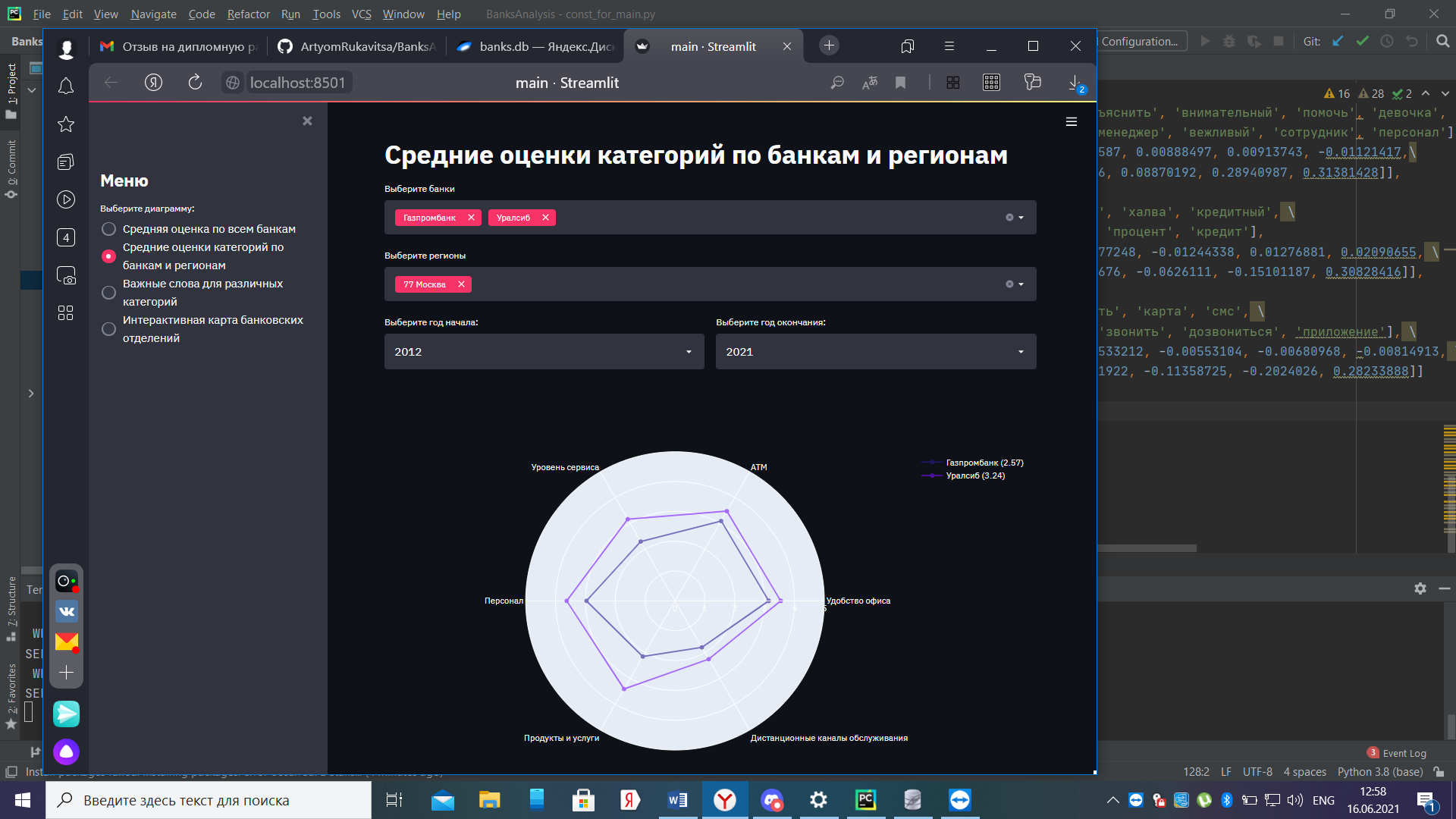
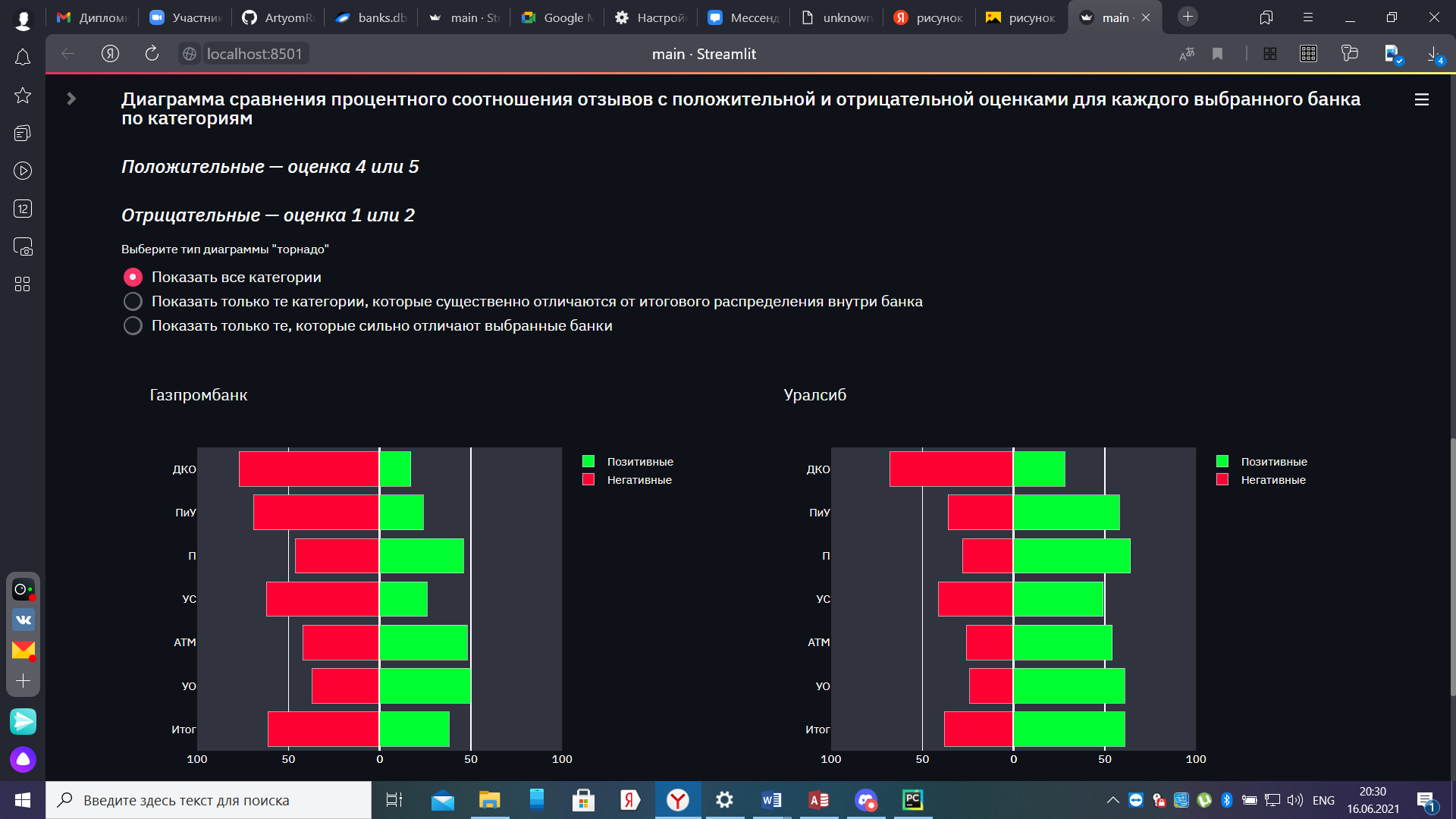
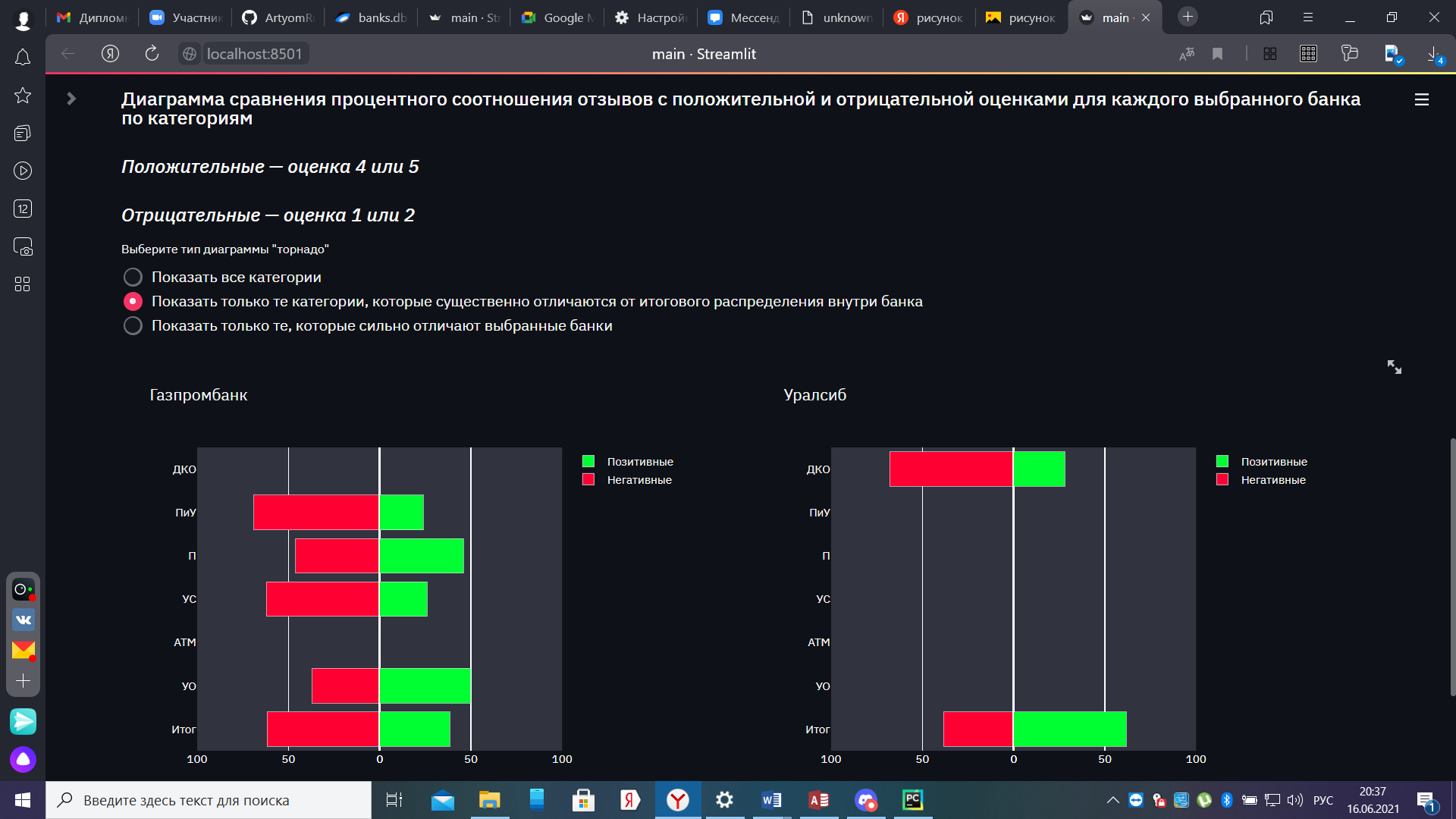
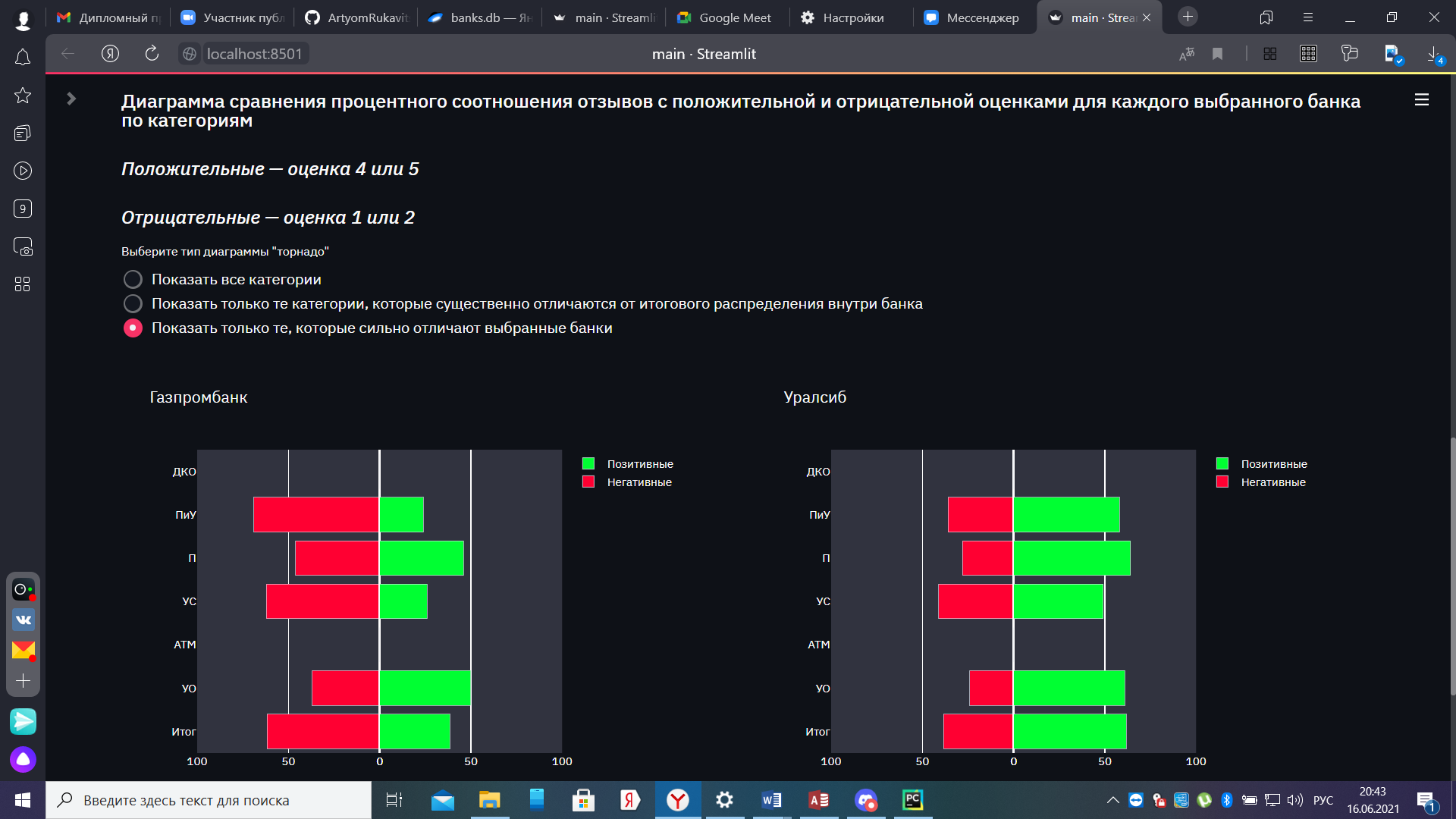
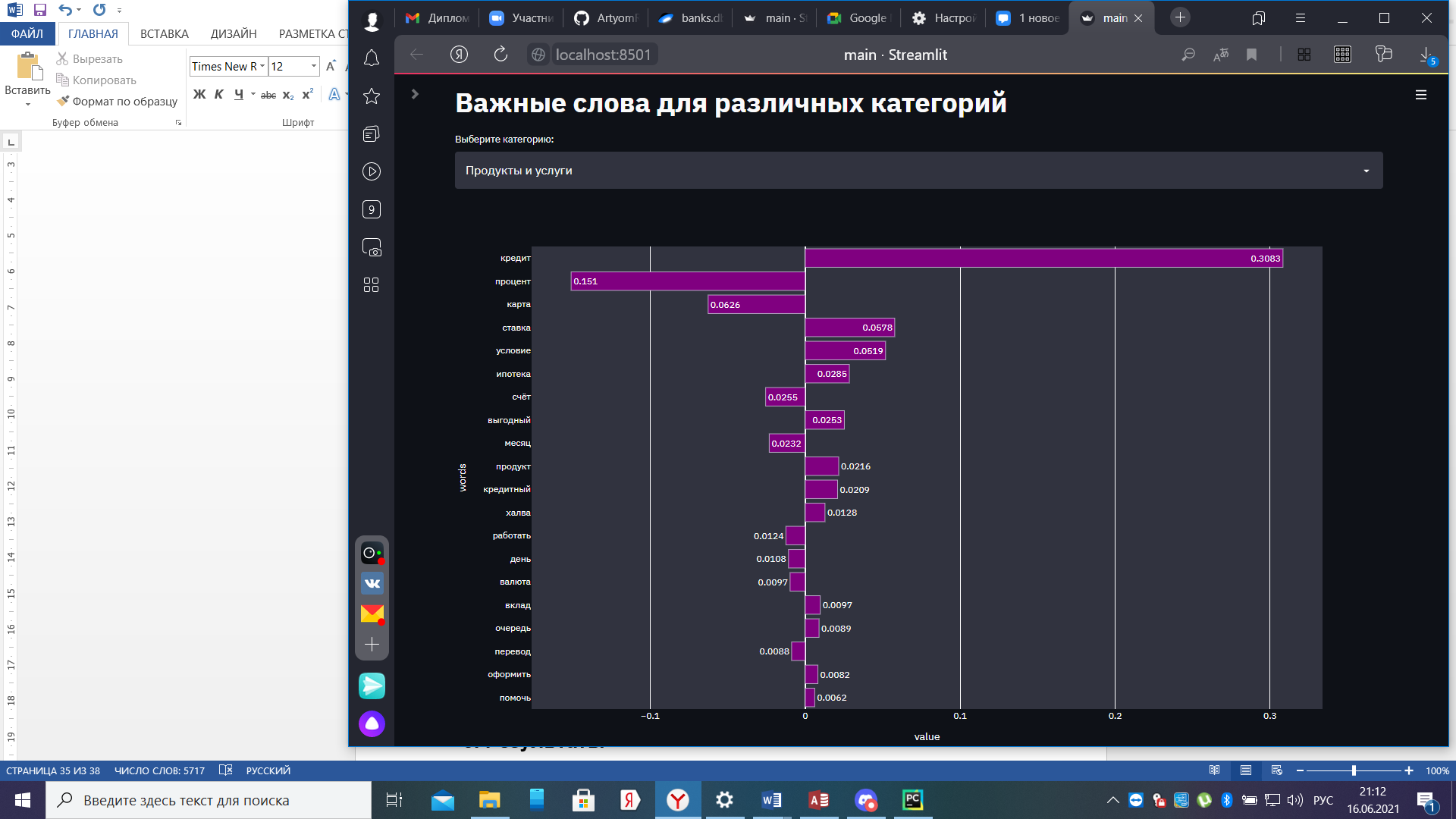
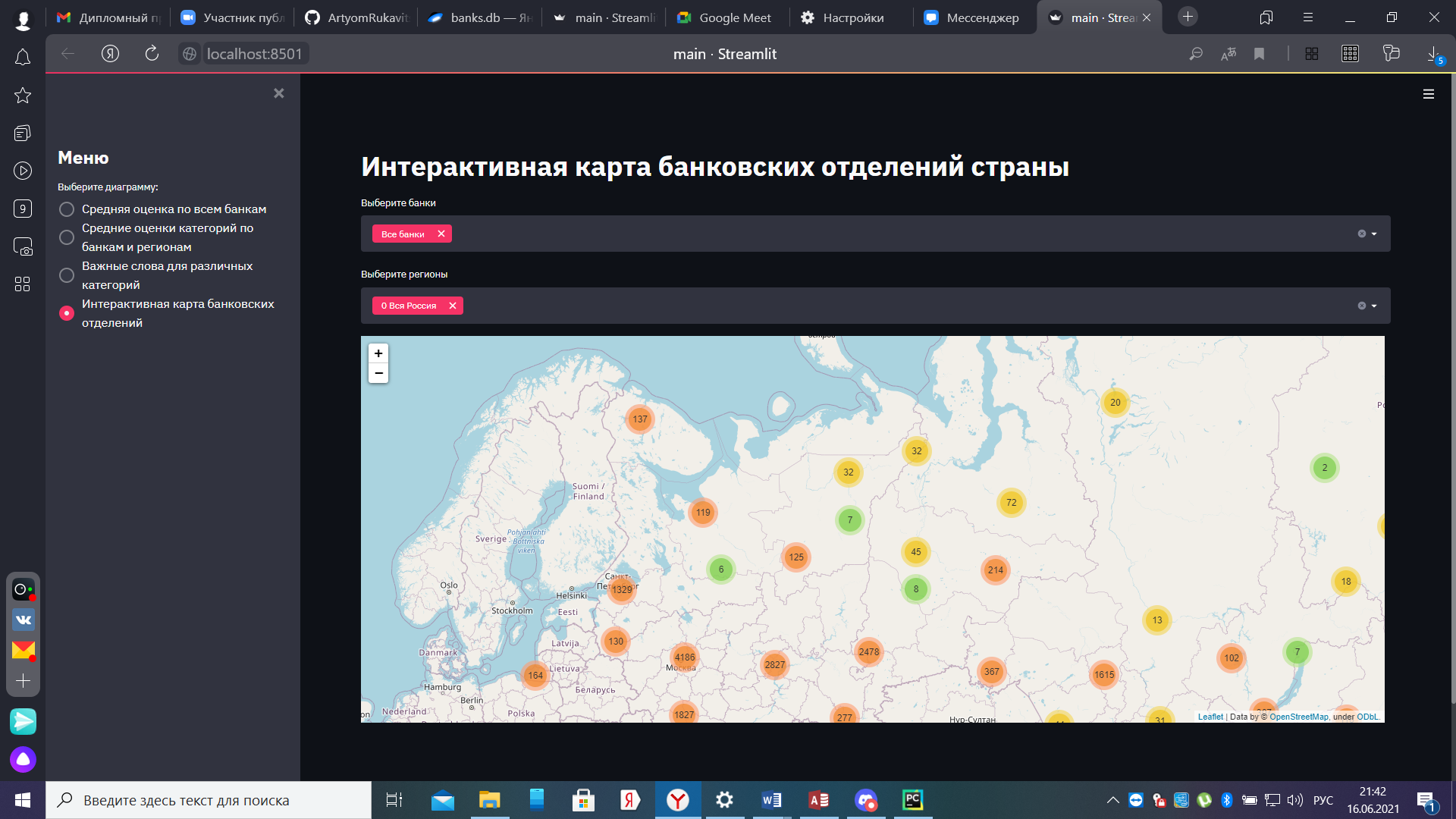
1. Analysis of user reviews of banks' mobile apps from Google Play;
2. Analysis of detailed customer reviews of banking services from banki.ru.

Нами создан продукт, подход которого применим к оцениванию и сравнению различных видов обслуживания, например, приложения банков и доставки продуктов, сети продовольственных супермаркетов и пр. Для построения аналитики нужны лишь отзывы, а также их категоризация (т.е. про что пользователи пишут).

**Литература:**

1. Критерий согласия Пирсона χ2 (Хи-квадрат) | statanaliz.info [Электронный ресурс]: 07.10.2019 — Режим доступа: https://statanaliz.info/statistica/proverka-gipotez/kriterij-soglasiya-pirsona-khi-kvadrat/ (Дата обращения 04.06.2021)
2. Elizabeth Mays. «Credit Scoring for Risk Managers. The Handbook for Lenders» 2004 г., с. 63-89
3. Краткий обзор алгоритма машинного обучения Метод Опорных Векторов (SVM) [Электронный ресурс]: 01.11. 2018 — Режим доступа: https://habr.com/ru/post/428503/ (Дата обращения 05.04.2021)

**Дополнительные материалы к тезисам**

1.  Горизонтальная гистограмма
2.  Диаграмма в формате “паутинка”
3. Диаграммы в формате «торнадо» с применением критерия согласия Пирсона
4. Важные слова для различных категорий
5. Интерактивная карта банковских отделений РФ
6. Подробная информаиция об одном из отделений на карте

