**Государственное бюджетное образовательное учреждение**

**Школа № 1533 «ЛИТ»**

**Анализ отзывов клиентов на банковские каналы обслуживания и продукты/услуги**

Рукавица Артём Кириллович, группа 10.5

Федченко Анастасия Кирилловна, группа 10.4

ГБОУ Школа № 1533 «ЛИТ»

Чамров Михаил Валерьевич

Вице-президент, Директор департамента кредитного бизнеса ПАО Банк «ФК Открытие»

Москва, 2021

Оглавление

[Введение 3](#_Toc64719119)

[Цель и задачи работы 3](#_Toc64719120)

[Гипотеза 3](#_Toc64719121)

[Методика выполнения работы 3](#_Toc64719122)

[Результаты и обсуждение 3](#_Toc64719123)

[Выводы 3](#_Toc64719124)

[Список используемой литературы 3](#_Toc64719125)

# Введение

Причиной актуальности нашего продукта является отсутствие в открытом доступе программ/сайтов, занимающихся сравнением клиентского опыта в различных банках. Мы предполагаем, что целевой аудиторией нашего сайта будут две группы людей. Первая — пользователи, которые смогут увидеть сильные и слабые стороны банков и выбрать банк, подходящий для них. Вторая — это представители компаний, которые хотят улучшить процесс предоставления услуг в условиях постоянно изменяющихся предпочтений пользователей конкурентной среды.

Мы решили выбрать данную тему проекта, так как нам интересно попробовать себя в работе с большими данными и навыки, полученные в ходе выполнения проекта, с большой вероятность потребуются нам в будущем.

# Цель и задачи работы

Цель работы — создание эффективного и объективного информационного поля для быстрого сравнения банков по основным факторам, полученным на основе отзывов клиентов на отделения с Яндекс.Карт.

# Задачи:

1. Собрать данные с картографического сервиса Яндекс.Карты;
2. Обработать собранные отзывы;
3. Создать модель для анализа текста;
4. Визуализировать результат.

Гипотеза

Выбор клиентов в пользу того или иного банка будет гораздо быстрее и удобнее, если разработать специальный сайт с интуитивно понятным интерфейсом, помогающий с анализом клиентского опыта в разных банках, в основании которого будут отзывы реальных людей, взятые с картографического сервиса.

# Методика выполнения работы

В качестве языка программирования мы выбрали язык Python, так как на нем легко писать аналитику и визуализацию сайта. IDE — PyCharm. Для работы с БД мы решили использовать программу SQLiteStudio. Для первичного анализа данных — KNIME Analytics Platform. В качестве API мы использовали API Поиска по организациям от компании Яндекс.

Мы начали использовать такие библиотеки на Python как:

1. requests -- для работы с API и сайтами;
2. BeautifulSoup4 -- для скреппинга отзывов через HTML-структуру сайта Яндекс.Карт;
3. sqlite -- для работы с БД;
4. nltk — для обработки естественного языка;
5. pymorphy2 — для работы с русскими словами (приведения в начальную форму);

Через код на Python мы обращались к двум сайтам: 1000bankov.ru (это источник отделений банков) и yandex.ru/maps (источник отзывов).

# Результаты и обсуждение

* Собрана база данных (300+ тысяч отзывов с Яндекс.Карт на 21500+ отделений 24 банков);
* Произведена отчистка текста от знаков препинания, цифр и стоп-слов, а также все слова приведены в начальную форму.

# Выводы

1. Собрали данные с картографического сервиса Яндекс.Карты;
2. Обработали собранные отзывы.

Исходя из всего выше сказанного, была проделана половина работы и продолжается активная работа над проектом.

# Список используемой литературы

* ООО «[Яндекс](https://yandex.ru/)»: О сервисе. API поиск по организациям — 2020 — URL: <https://yandex.ru/dev/maps/geosearch/doc/concepts/about.html?from=geosearch> (Дата обращения 15.11.2020)
* [Voximplant](https://habr.com/company/Voximplant/): Основы Natural Language Processing для текста / Блог компании [Voximplant](https://habr.com/company/Voximplant/) / Хабр — 15.04.2019 — URL: <https://habr.com/ru/company/Voximplant/blog/446738/> (Дата обращения 21.02.2021)
* Mikhail Korobov Revision: Документация — Морфологический анализатор pymorphy2 — 2013-2020 — URL: <https://pymorphy2.readthedocs.io/en/latest/user/index.html> (Дата обращения 21.02.2021)
* Список регионов (субъектов, областей) России 2021 РФ с кодами согласно данным ФНС по алфавиту — 2020 — URL: <https://www.sites.google.com/site/ruregdatav1/spisok-regionov-rossii-s-kodamy> (Дата обращения 12.12.2020)
* Udemy: Бесплатное учебное руководство по теме “Обработка и анализ данных” — 2021 — URL: <https://www.udemy.com/course/knime-bootcamp/> (Дата обращения 05.02.2021)

оформление в соответствии с ГОСТ Р 7.0.100 – 2018