

**Анализ «цифровых следов» как инструмент предиктивного маркетинга на примере кофеен Санкт-Петербурга**

**Егоров Н.С. (ИТМО), Зайцева И.А. (ИТМО)**

**Научный руководитель – старший преподаватель Бурков И.П. (ИТМО)**

**Введение.** Выбор кофейни для посещения является сложным процессом, зависящим от множества факторов: расположения, ценовой политики, формата заведения, отзывов и предпочтений потребителей [1]. Несмотря на активное развитие рынка кофеен, данная проблема остается малоизученной. Важно понимать, какие параметры играют ключевую роль в принятии решений посетителями, что позволяет бизнесу более точно ориентироваться на целевую аудиторию [2].

**Основная часть.** Цель данного исследования заключается в оценке влияния отзывов пользователей на выбор кофеен, используя анализ цифровых следов и методы машинного обучения.

В рамках исследования решались следующие задачи:

1. Определить ключевые темы, которые наиболее часто встречаются в пользовательских отзывах на Яндекс Картах, и выявить их влияние на формирование восприятия кофейни.
2. Исследовать, каким образом тональность отзывов коррелирует с популярностью заведений, применяя методы машинного обучения.
3. Определить, в какой степени качество напитков, уровень обслуживания и ценовая стратегия заведения отражаются на его рейтинге среди пользователей.

Для анализа были собраны данные 1000 кофеен Санкт-Петербурга, имеющих значительное число отзывов на Яндекс Картах. Эти заведения охватывают различные форматы кофеен (coffee-to-go, американский формат, формат полного цикла), что позволяет собрать репрезентативные данные для исследования.

Методология исследования включала сбор данных с использованием Selenium для автоматизированного парсинга веб-страниц и API Яндекс Карт [3]. В общей сложности было собрано более 200 000 отзывов, которые затем прошли предобработку.

Анализ отзывов проводился с применением алгоритма NLP – BERT (Bidirectional Encoder Representations from Transformers), обученный на русскоязычных текстах. Этот алгоритм позволил выполнить тематическое моделирование отзывов и определение их тональности.

**Выводы.** Исследование показало следующее:

1. Качество напитков является наиболее значимым фактором для положительных отзывов, что подтверждает его влияние на репутацию кофейни.
2. Высокое время ожидания заказов вызывает наибольшее количество негативных отзывов, что указывает на необходимость оптимизации процессов в кофейнях.
3. Ценовая политика играет важную роль в оценках пользователей, но ее влияние уступает таким параметрам, как сервис и удобство расположения.

Таким образом, анализ цифровых следов пользователей и применение методов машинного обучения позволяют более точно прогнозировать потребительские предпочтения. Результаты исследования могут быть использованы для корректировки маркетинговых стратегий и улучшения клиентского опыта в кофейнях.

**Список использованных источников:**

1. Попова Е.А. Анализ потребительских предпочтений в сегменте кофеен // Журнал маркетинговых исследований. – 2021. – №3. – С. 25–32.

2. Волобуев С.С. Социально ориентированные бизнес-сообщества в виртуальных социальных сетях // Общество: социология, психология, педагогика. – 2021. – №6 (86).
3. Яндекс Карты: открытые данные и API: [сайт]. – URL: <https://yandex.ru/maps> (дата обращения: 10.02.2025).