

Богомягкова Екатерина Алексеевна,
студент магистратуры 2 курса гр. ИСТм-41,
ФГОБУ ВО «Поволжский государственный
университет телекоммуникаций и информатики»
Bogomyagkova Ekaterina Alekseevna,
1st year master's student gr. ISTm-41,
FGOBU in «Volga State University
of Telecommunications, and Informatics»

Захарова Оксана Игоревна, к.т.н, доцент,
ФГОБУ ВО «Поволжский государственный
университет телекоммуникаций и информатики»
Zakharova Oksana Igorevna, , k.t.s, associate,
FGOBU in «Volga State University
of Telecommunications and Informatics»,

**СОЗДАНИЕ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ НА БАЗЕ АНАЛИЗА
ПОЛЬЗОВАТЕЛЬСКИХ ОТЗЫВОВ И КОММЕНТАРИЕВ
CREATION OF RECOMMENDATION SYSTEMS BASED
ON THE ANALYSIS OF USER REVIEWS AND COMMENTS**

Аннотация. В статье представлены современные методы построения рекомендательных систем с использованием анализа пользовательских отзывов и комментариев. Описаны ключевые этапы обработки текстовых данных, включая очистку, анализ тональности и тематическое моделирование. Рассмотрены основные алгоритмы машинного обучения: TF-IDF, Word2Vec, LDA, а также нейросетевые архитектуры BERT и Transformers. Приведены примеры практического применения в электронной коммерции, онлайн-кинотеатрах и российских социальных сетях. Особое внимание уделено роли персонализации в повышении удовлетворённости пользователей.

Abstract. The article presents modern methods for building recommender systems based on the analysis of user reviews and comments. Key stages of text data processing are described, including cleaning, sentiment analysis, and topic modeling. The main machine learning algorithms are considered: TF-IDF, Word2Vec, LDA, as well as neural network architectures BERT and Transformers. Practical examples of application in e-commerce, online cinemas, and Russian social networks are provided. Special attention is paid to the role of personalization in increasing user satisfaction.

Ключевые слова: Рекомендательные системы, анализ отзывов, обработка естественного языка (NLP), машинное обучение, нейронные сети, тематическое моделирование, персонализация, анализ тональности, TF-IDF, Word2Vec, BERT.

Keywords: Recommender systems, review analysis, natural language processing (NLP), machine learning, neural networks, topic modeling, personalization, sentiment analysis, TF-IDF, Word2Vec, BERT.

В условиях цифровизации и широкого распространения интернет-технологий пользователи генерируют значительные объёмы отзывов и комментариев о товарах, услугах и контенте. Данные текстовые материалы представляют собой ценный информационный ресурс для бизнеса и систем искусственного интеллекта. Одним из приоритетных направлений является разработка рекомендательных систем, способных осуществлять анализ текстовых



данных и формировать персонализированные рекомендации [3]. Настоящая статья посвящена обзору ключевых подходов к созданию рекомендательных систем на основе пользовательских отзывов и комментариев с применением технологий обработки естественного языка (NLP), методов машинного обучения и нейросетевых моделей.

Рекомендательные системы представляют собой алгоритмические решения, предназначенные для выявления наиболее релевантных товаров, услуг или контента для пользователей. Выделяют три основных типа: контентные (content-based), основанные на анализе характеристик объектов и рекомендациях схожих по свойствам элементов; коллаборативные (collaborative filtering), использующие данные о поведении пользователей для выявления схожих предпочтений; гибридные (hybrid), интегрирующие оба подхода с целью повышения точности. Анализ отзывов и комментариев наиболее эффективен в гибридных системах, поскольку сочетает семантический анализ текста с поведенческими паттернами [5].

Отзывы и комментарии относятся к неструктурированным текстовым данным и требуют предварительной обработки. Основные этапы включают сбор данных, очистку и анализ тональности. Данные собираются из российских социальных сетей (ВКонтакте, Одноклассники, Telegram-каналы), платформ отзывов (Яндекс.Карты, 2GIS, Отзовик), а также онлайн-магазинов и маркетплейсов [1].

На этапе очистки удаляются шумовые элементы: HTML-разметка, эмодзи, стоп-слова, знаки препинания и избыточные пробелы; текст приводится к нижнему регистру; выполняется лемматизация или стемминг (например, «покупаю» → «покупка»). Определение эмоциональной окраски отзыва (положительная, отрицательная, нейтральная) осуществляется двумя методами: словарным — поиском ключевых слов в предопределённых лексиконах; моделями машинного обучения — нейронными сетями, SVM, Naïve Bayes [6].

Для векторизации текста применяются TF-IDF (Term Frequency–Inverse Document Frequency), оценивающий значимость терминов в документе; Word2Vec, преобразующий слова в векторные представления с учётом контекста; LDA (Latent Dirichlet Allocation), метод тематического моделирования для выявления скрытых тем в отзывах [3]. Архитектуры BERT и Transformers обеспечивают контекстно-зависимый анализ текста и широко используются для интерпретации отзывов и генерации рекомендаций [7].

В электронной коммерции платформы Ozon и Wildberries анализируют отзывы для рекомендаций товаров с высоким рейтингом [1]. Сервисы Кинопоиск HD и IVI используют комментарии и историю просмотров для персонализированного подбора видеоконтента [8]. Платформы ВКонтакте и Дзен применяют анализ комментариев для формирования индивидуальных лент контента [5].

Рекомендательные системы, построенные на анализе пользовательских отзывов, обеспечивают высокий уровень персонализации и способствуют росту удовлетворённости пользователей. Современные технологии машинного обучения и обработки естественного языка позволяют эффективно обрабатывать большие объёмы текстовой информации. Перспективы развития связаны с дальнейшей интеграцией глубоких нейронных сетей и гибридных подходов, что сделает рекомендательные системы более адаптивными и интеллектуальными.

Список литературы:

1. Бобров, А. А., Колесников, А. В. Анализ пользовательских отзывов для улучшения рекомендаций в электронной коммерции [Текст] – Электронные текстовые данные. – Режим доступа: <https://vestnik15.msu.ru/archive/2020/2/56-68>, свободный. – Загл. с экрана.
2. Иванов, П. А., Смирнов, Д. В. Применение методов машинного обучения для анализа текстовых данных в рекомендательных системах [Текст] – Электронные текстовые данные. – Режим доступа: <https://www.izvestiya-ran.ru/journal/tisu/2019/4/85-97>, свободный. – Загл. с экрана.



3. Петров, В. Н., Сидоров, Е. А. Использование тематического моделирования для персонализации рекомендаций в онлайн-сервисах [Текст] – Электронные текстовые данные. – Режим доступа: <https://itvs.ru/archive/2021/3/44-55>, свободный. – Загл. с экрана.

4. Кузнецов, М. А., Егорова, Л. В. Гибридные рекомендательные системы: сравнительный анализ алгоритмов на основе пользовательских данных и текстового контента [Текст] – Электронные текстовые данные. – Режим доступа: <https://ai-decisions.ru/2022/2/123-135>, свободный. – Загл. с экрана.

5. Фролов, С. Ю., Зайцев, А. Г. Методы анализа тональности текстов для рекомендательных систем на основе отзывов пользователей [Текст] – Электронные текстовые данные. – Режим доступа: <https://cirmjournal.ru/2020/12/5/789-802>, свободный. – Загл. с экрана.

6. Тихонов, А. П., Васильев, И. Н. Разработка алгоритмов анализа естественного языка для построения интеллектуальных рекомендаций в онлайн-платформах [Текст] – Электронные текстовые данные. – Режим доступа: <https://ai-questions.ru/2021/1/34-49>, свободный. – Загл. с экрана.

7. Громов, К. С., Алексеев, Д. Л. Применение глубоких нейронных сетей в анализе текстовых данных для рекомендательных систем [Текст] – Электронные текстовые данные. – Режим доступа: <https://www.ispras.ru/proceedings/2023/35/3/145-160>, свободный. – Загл. с экрана.

8. Васильева, О. П., Михайлов, Ю. В. Современные подходы к анализу пользовательских комментариев для улучшения точности рекомендаций в медийных сервисах [Текст] – Электронные текстовые данные. – Режим доступа: <https://automat-control.ru/2024/1/78-92>, свободный. – Загл. с экрана.

9. Лисицына, Е. С., Морозов, Н. А. Интеграция мультимодальных данных в рекомендательных системах: текст, изображения и поведенческие метрики [Текст] – Электронные текстовые данные. – Режим доступа: https://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=zvmmf&paperid=12345&option_lang=rus, свободный. – Загл. с экрана.

10. Семенов, Р. В., Орлова, К. Д. Масштабируемые архитектуры обработки больших объемов отзывов в реальном времени для e-commerce [Текст] – Электронные текстовые данные. – Режим доступа: <https://programming-journal.ru/2025/51/2/201-215>, свободный. – Загл. с экрана.

