

Мясоедова Юлия Витальевна

магистрант, Белгородский государственный национальный

исследовательский университет,

Россия, г. Белгород

**ТЕОРЕТИЧЕСКИЕ ОСНОВЫ АНАЛИЗА ЭМОЦИОНАЛЬНОЙ
ОКРАСКИ СООБЩЕНИЙ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ С
ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

АННОТАЦИЯ

В статье рассматриваются теоретические основы и методы анализа эмоциональной окраски (сентимент-анализа) текстовых сообщений из социальных сетей с применением технологий искусственного интеллекта. Проведен анализ эволюции подходов к решению задачи: от лексических методов и классических алгоритмов машинного обучения до современных архитектур глубоких нейронных сетей (трансформеров). Особое внимание уделено специфике обработки неформального русскоязычного контента, содержащего сленг, эмодзи и орфографические ошибки. Автором обосновывается необходимость комбинированных алгоритмов, объединяющих глубокий семантический анализ, извлечение явных лингвистических признаков и поведенческий анализ во временной динамике. Подчеркивается теоретическое разграничение клинической депрессии и статистически значимых лингвистических паттернов, что исключает подмену алгоритмического анализа клинической диагностикой. Результаты работы могут быть использованы при создании адаптивных систем мониторинга психологического состояния пользователей на основе данных социальных медиа.

ABSTRACT

The article discusses the theoretical foundations and methods of analyzing the emotional coloring (sentiment analysis) of text messages from social networks using artificial intelligence technologies. The analysis of the evolution of approaches to solving the problem is carried out: from lexical methods and classical machine learning algorithms to modern architectures of deep neural networks (transformers). Special attention is paid to the specifics of processing informal Russian-language content containing slang, emojis and spelling errors. The author substantiates the need for combined algorithms combining deep semantic analysis, extraction of explicit linguistic features and behavioral analysis in time dynamics. The theoretical distinction between clinical depression and statistically significant linguistic patterns is emphasized, which excludes the substitution of algorithmic analysis for clinical diagnosis. The results of the work can be used to create adaptive systems for monitoring the psychological state of users based on social media data.

Ключевые слова: анализ тональности, социальные сети, обработка естественного языка (NLP), машинное обучение, выявление депрессии, нейронные сети, трансформеры, RuBERT, лингвистические маркеры, психологическая диагностика.

Keywords: tonality analysis, social networks, natural language processing (NLP), machine learning, depression detection, neural networks, transformers, RuBERT, linguistic markers, psychological diagnostics.

Современное общество характеризуется быстрым развитием цифровых технологий и широким распространением социальных сетей, которые стали неотъемлемой частью коммуникации и самовыражения миллионов людей. Однако эта цифровая среда становится не только площадкой для общения, но и мощным источником данных о психологическом состоянии пользователей. По данным Всемирной организации здравоохранения, более 280 миллионов человек в мире страдают от депрессии, причем многие из них не получают

своевременной помощи из-за стигматизации, нехватки специалистов или нежелания обращаться к ним лично.

Под анализом эмоциональной окраски текста понимается процесс определения эмоционального состояния автора на основе текста. Этот метод машинного обучения и обработки естественного языка анализирует тексты, чтобы определить их тональность: положительную, отрицательную или нейтральную.

Одним из ключевых аспектов анализа эмоциональной окраски сообщений является поиск признаков депрессии. Депрессия – психическое расстройство, которое характеризуется длительными периодами подавленного настроения либо утраты интереса к привычной деятельности или способности получать от неё удовольствие.

Ключевым теоретическим ограничением в данной области является строгое разграничение понятий «клиническая депрессия» и «лингвистические признаки депрессивного состояния». Интеллектуальный анализ данных не способен и не должен заменять клиническую диагностику. Выявляемые алгоритмами паттерны представляют собой лишь статистические корреляты, которые могут свидетельствовать о ситуативном стрессе, усталости или специфике коммуникативного стиля.

Обработка текстовых сообщений с использованием инструмента анализа настроений позволяет значительно ускорить процесс обработки данных и повысить точность определения эмоциональной окраски сообщений.

Основу предметной области составляют методы машинного обучения и обработки естественного языка. Эволюция подходов прошла путь от простых словарных методов к алгоритмам классического машинного обучения и далее к глубоким нейронным сетям. Однако для неформальных текстов социальных сетей, где присутствуют сленг, опечатки и специфические сокращения, требуются более адаптивные подходы.

Теоретический базис анализа эмоций прошел путь от детерминированных правил к вероятностным нейросетевым моделям.

Сравнение основных классов методов выявляет их фундаментальные ограничения:

1) Лексические методы. Недостаток: абсолютная слепота к контексту. Слова «убийственно» и «классно» в разных контекстах могут нести противоположную эмоциональную нагрузку, что игнорируется словарями.

2) Классическое машинное обучение. Недостаток: зависимость от ручной инженерии признаков и слабая адаптивность к морфологии русского языка.

3) Нейросетевые модели с фиксированными эмбедингами. Недостаток: неспособность учитывать глобальный контекст предложения и длинные семантические зависимости.

4) Трансформерные модели. Базируются на механизме самовнимания, что позволяет учитывать двунаправленный контекст. Несмотря на наивысшую точность, они функционируют как «черный ящик», что снижает интерпретируемость результатов для специалистов.

Для преодоления ограничений изолированных методов современные теоретические разработки опираются на комбинированные алгоритмы.

Ранние исследования в области анализа эмоциональной окраски русскоязычных текстов были сосредоточены на создании лексических ресурсов и применении классических алгоритмов. Современный этап характеризуется переходом к использованию предобученных языковых моделей на архитектуре трансформеров. Исследования показывают превосходство моделей BERT и их адаптированных для русского языка версий в задачах сентимент-анализа, что связано с их способностью учитывать глубокий контекст и семантику предложений.

Тексты в социальных сетях характеризуются высокой степенью неформальности, наличием сленга, эмодзи, орфографических ошибок и специфических сокращений, что создает дополнительные сложности для решения задач эмоционального анализа интернет-текстов. Для решения этих проблем требуется как специфическая предобработка данных, так и

использование моделей, обученных на соответствующих корпусах. Особое направление в рамках предметной области – автоматическое выявление рисков психических расстройств, в частности депрессии. Исследования установили статистически значимую связь между определенными лингвистическими паттернами в соцсетях и клинически диагностированной депрессией. Работы по диагностике депрессивных состояний на основе данных социальных медиа развивают это направление, адаптируя методы машинного обучения для русскоязычного контента и исследуя маркеры депрессии.

Базовая архитектура такого подхода включает три параллельных контура анализа:

1) Глубокий семантический анализ. Использование дообученных трансформерных моделей для формирования контекстных векторных представлений и оценки скрытых смыслов.

2) Извлечение явных лингвистических признаков. Статистический анализ частоты негативной лексики, маркеров тревожности, аномалий пунктуации и длины предложений. Это компенсирует «непрозрачность» нейросетей.

3) Поведенческий и временной анализ. Эмоция не всегда содержится в одном сообщении. Теоретически обоснованным является анализ динамических окон. Это позволяет отслеживать кумулятивный эффект, частоту публикаций и устойчивые паттерны апатии во времени.

Анализ эмоциональной окраски текстов представляет собой междисциплинарное направление, в котором прослеживается устойчивая эволюция от простых словарных подходов к нейросетевым архитектурам. Ключевым вектором является адаптация моделей к специфике неформальной интернет-коммуникации для решения прикладных задач, в частности – выявления лингвистических маркеров депрессии на основе данных социальных сетей.

Литература

1. Анализ тональности текста [Электронный ресурс]. — URL <https://sysblok.ru/courses/kak-provesti-analiz-tonalnosti-teksta/> (дата обращения: 19.03.2026).
2. Человек и системы искусственного интеллекта [Текст]/ В.А. Лекторский — Санкт-Петербург : Юридический центр Пресс, 2022. — 328 с.
3. Шевченко, А. С. Методы машинного обучения : учебное пособие [Текст]/ А. С. Шевченко, О. В. Самарина. — Москва : Ай Пи Ар Медиа, 2025. — 143 с.