

Методы использования больших языковых моделей для выявления провокационных комментариев в социальных сетях во время протестов

Уртнасан Батнасан

Академия управления МВД России, Москва

Аннотация: В статье рассматриваются методы использования больших языковых моделей (БЯМ) для выявления провокационных комментариев в социальных сетях во время массовых протестов. Исследование включает анализ существующих подходов к обработке текстовых данных, разработку методики применения моделей Llama3 8B, Mistral 7B и Gemma 7B, а также оценку их эффективности на примере комментариев, связанных с протестами в Монголии в декабре 2022 года. Результаты показывают, что модель Mistral 7B является наиболее точной и эффективной для классификации провокационных комментариев. Выводы исследования подтверждают, что использование больших языковых моделей значительно улучшает точность и оперативность анализа контента в социальных сетях, что важно для управления общественными настроениями и предотвращения конфликтов.

Ключевые слова: большая языковая модель, социальная сеть, провокационный комментарий, массовый протест, анализ текста, обработка естественного языка, Llama3 8B, Mistral 7B, Gemma 7B, классификация, точность, полнота, F1-мера, политическая нестабильность, управление общественным мнением.

Введение

В современном мире социальные сети играют ключевую роль в формировании общественного мнения, особенно в контексте массовых протестов. Особенно актуальным становится анализ провокационных комментариев, которые могут как способствовать эскалации конфликта, так и служить инструментом для манипуляции общественным мнением. Например, в ходе протестов в Монголии прошел протест под названием "Кража угля", организованный против ущерба в размере 44 триллионов тугриков, нанесенного государству, активно использовались социальные сети для распространения как поддерживающих, так и провокационных комментариев. Протестующие граждане и пользователи социальных сетей активно призывали полицейских бросить работу и встать на сторону протестов. Это существенно влияло на динамику событий, демонстрируя

важность анализа контента социальных сетей в условиях политической нестабильности.

Цель данного исследования - разработать методы использования больших языковых моделей для выявления и анализа провокационных комментариев в социальных сетях во время массовых протестов. В рамках исследования ставятся следующие задачи:

- Анализ существующих подходов к обработке текстовых данных в социальных сетях.
- Разработка методики применения больших языковых моделей для выявления провокационных комментариев.
- Оценка эффективности предложенной методики на примере конкретных данных.

В исследовании предполагается, что применение языковых моделей, таких как не требующие больших вычислительных мощностей Llama-3 8B, Gemma 7B и Mistral 7B, может значительно повысить точность и оперативность анализа текстов в социальных сетях, что окажет значительное влияние на управление общественными настроениями и предотвращение возможных конфликтов.

Обзор литературы

В современном анализе текстов в социальных сетях используются как традиционные, так и современные методы. Традиционные методы, такие, как частотный анализ слов и контент-анализ, обеспечивают высокую точность и позволяют обрабатывать большие объемы данных. Однако они требуют значительных затрат времени на ручную классификацию данных и имеют ограниченные возможности по выявлению контекста.

Влияние цифровой среды на поведение молодежи в интернете становится всё более значимой темой для исследований. Ларионова и Горчакова отмечают, что социальные сети, такие, как ВКонтакте,

способствуют формированию деструктивных коммуникативных паттернов, особенно в контексте социально-политических обсуждений [1]. Эти наблюдения важны для понимания особенностей взаимодействия пользователей с контентом в условиях политической нестабильности.

Применение больших языковых моделей (БЯМ) активно изучается в различных областях, таких, как управление рисками, где они позволяют анализировать большие объемы текстовой информации и находить скрытые зависимости в данных [2]. Эти модели предоставляют возможности для более эффективного анализа и автоматизации процессов обработки текстов, что делает их перспективными для использования в мониторинге социальных сетей.

В работе Центра демократии и технологий [3] обсуждаются проблемы и возможности применения больших языковых моделей для анализа текстов на языках, отличных от английского. Особое внимание уделено тому, как компании уже используют эти модели для модерации контента в социальных сетях. Работа подчеркивает важность адаптации больших языковых моделей для работы с различными языками, что способствует улучшению точности и эффективности анализа текстов в глобальном масштабе.

Манипуляция общественным мнением в социальных сетях стала одной из ключевых тем современных исследований в области информационной безопасности. Как отмечает Тихон, распространение практик, таких, как астротурфинг и массовые информационные вбросы, позволяет создавать искусственное восприятие массовой поддержки или критики в социальных медиа [4]. Эти методы становятся особенно актуальными в условиях политической нестабильности и могут существенно влиять на общественное мнение.

Большие языковые модели (БЯМ) все чаще используются для анализа социальных сетей, особенно в контексте выявления тем и настроений,

обсуждаемых пользователями. Как указывает Брагин, БЯМ способны выявлять скрытые паттерны и тенденции в общественном восприятии через анализ текстов, опубликованных в социальных сетях, таких как ВКонтакте [5]. Это делает их важным инструментом для понимания настроений пользователей и предсказания социально значимых тем.

Методы обработки естественного языка (NLP) стали неотъемлемой частью анализа текстов, предоставляя возможности для выявления скрытых связей и анализа больших объемов неструктурированных данных. Как отмечают Белов и его коллеги, современные технологии NLP, такие как Word2vec и глубокие нейронные сети, позволяют преобразовывать текстовые данные в векторные представления, которые эффективно отражают семантическую близость слов и контекстов [6].

Современные методы моделирования социальных взаимодействий включают в себя применение многоагентного моделирования (МММ), которое позволяет учитывать взаимодействие множества факторов и анализировать динамику изменений поведения агентов в сложных системах. Тымчук и Свечкарев указывают, что МММ стремится к представлению реальных социальных динамик через использование программных агентов, которые имитируют поведение реальных объектов в изменяющейся среде [7]. Это делает МММ эффективным инструментом для изучения поведения пользователей в цифровой среде.

Помимо этих методов, важную роль в понимании социальных процессов играют модели, основанные на теории идентичности и многоагентном моделировании. В работе Тымчука [8] проводится анализ многоагентных социальных моделей террористически опасных ситуаций на основе теории идентичности. Авторы рассматривают ключевые критерии моделей, такие как стабильность, доступность и соответствие идентичностей,

что позволяет лучше понять механизмы формирования конфликтов и распространения провокационного контента в социальных сетях.

Современные методы, такие, как машинное обучение и обработка естественного языка, позволяют автоматизировать процесс анализа, улучшить выявление контекста и значительно сократить время обработки данных, что делает их более эффективными для анализа больших массивов информации. В работе Василия Ванюкова [9] обсуждаются основные различия между традиционными методами обработки естественного языка (NLP) и большими языковыми моделями (БЯМ), включая их применение для анализа текстов. В статье рассматриваются преимущества и недостатки каждого подхода, а также обсуждаются случаи, когда использование БЯМ является более целесообразным и эффективным.

Несмотря на значительные достижения, текущие исследования все еще имеют ряд пробелов, таких как высокая вычислительная сложность и необходимость большого объема данных для обучения моделей. Дальнейшие исследования в области применения БЯМ для анализа текстов в социальных сетях могут заполнить эти пробелы и предложить более эффективные и менее ресурсоемкие методы.

Методология

1. Выбор моделей и обоснование:

В исследовании применяются три большие языковые модели: LLaма3 8B, Mistral 7B и Gemma 7B. Эти модели выбраны из-за их передовых способностей в обработке естественного языка и специфической адаптации к задачам текстового анализа в социальных сетях. LLaма3 8B и Mistral 7B известны своей эффективностью в генерации и понимании сложного текста, в то время как Gemma 7B обладает способностями к более глубокому контекстуальному анализу и разграничению намерений пользователя.

В рамках исследования были выбраны три современные большие языковые модели: Llama3 8B, Mistral 7B и Gemma 7B, на основе ключевых метрик, критичных для анализа текста в социальных сетях. Для сравнения их возможностей использовались результаты тестов, представленные на сайте Ollama, которые оценивают общие способности моделей к пониманию и обработке естественного языка:

- MMLU (Massive Multitask Language Understanding): Эта метрика оценивает общие способности моделей к пониманию языка в многозадачном режиме. Согласно данным, Llama3 8B показала наивысшую точность в 68.4%, что свидетельствует о её высоких общих способностях. Mistral 7B и Gemma 7B также показали хорошие результаты с 58.4% и 53.3% соответственно.

- GPQA (General Purpose Question Answering): Метрика оценивает способность моделей отвечать на общие вопросы. Llama3 8B достигла 34.2%, что на 12.8 процентных пунктов выше, чем у Gemma 7B (21.4%) и на 7.9 процентных пунктов выше, чем у Mistral 7B (26.3%).

- DROP F1 (Discrete Reasoning Over Paragraphs): В этом тесте Llama3 8B показала результат 58.4, немного опережая Gemma 7B с показателем 56.3 и Mistral 7B с показателем 54.4.

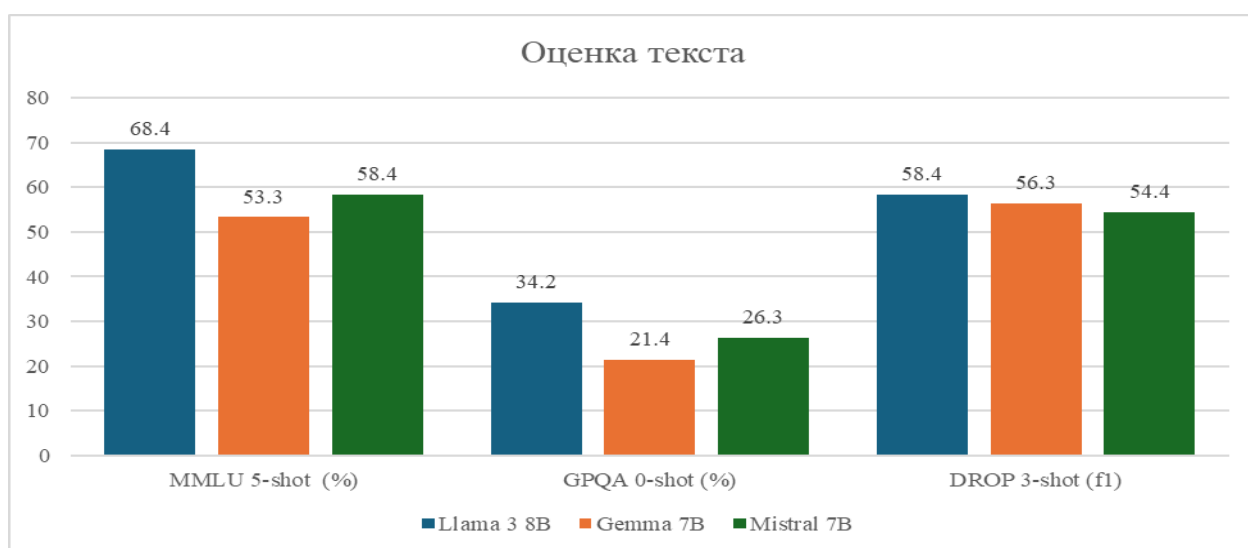


Рис. 1. – Сравнительный анализ эффективности больших языковых моделей по метрикам MMLU, GPQA и DROP

Эти результаты демонстрируют, что выбор модели для конкретных задач анализа данных должен основываться на специфических требованиях к точности и типу обработки текста. LLaма3 8B, например, показывает более высокую производительность в общих задачах понимания языка, в то время как другие модели могут быть предпочтительны для более специфических задач [10].

2. Сбор и подготовка данных:

Исходные данные для анализа собираются из MOSINT базы данных, содержащей комментарии, связанные с протестным движением в Монголии 4 декабря 2022 года под названием "Кража угля", организованным против ущерба государству в 44 триллиона тугриков. Данные подготавливаются через процесс очистки от несущественных символов и шума, а также аннотации, при которой комментарии классифицируются по степени провокационности и релевантности к теме протестов. Включается также этап предобработки данных, включающий лемматизацию и удаление стоп-слов, чтобы улучшить качество анализа.

3. Методы оценки моделей:

Для оценки эффективности моделей используются следующие метрики:

1. Точность (precision):

- Точность измеряет долю правильных положительных предсказаний от общего числа положительных предсказаний. Формула:

$$Precision = \frac{TP}{TP + FP}$$

где TP — число истинно положительных результатов, а FP — число ложно положительных результатов.

2. Полнота (recall):

- Полнота измеряет долю правильных положительных предсказаний от общего числа истинных положительных случаев. Формула:

$$Recall = \frac{TP}{TP + FN}$$

где TP — число истинно положительных результатов, а FN — число ложно отрицательных результатов.

3. F1-мера (F1-score):

- F1-мера является гармоническим средним между точностью и полнотой. Она используется для баланса между ними, особенно когда важно учитывать как ложные положительные, так и ложные отрицательные результаты. Формула:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Эти метрики были первоначально введены в области информационного поиска. В классическом исследовании "The evaluation of relevance" Кэлвина Мура и Стивена Эмбла (C. W. Cleverdon and M. Keen) были определены и обсуждены основные концепции точности и полноты. Это исследование было выполнено в 1966 году и стало основой для последующих исследований в этой области [11]. Важность данных метрик обусловлена необходимостью достижения баланса между избеганием ложных срабатываний и упущения важных комментариев.

Эксперименты и результаты

1. Настройка моделей

Для проведения экспериментов были использованы три большие языковые модели: Llama3 8B, Mistral 7B и Gemma 7B. Эти модели были запущены во время эксперимента на компьютере с использованием инструмента Ollama, который предоставляет удобный интерфейс для работы

с большими языковыми моделями. Параметры настройки моделей включали следующее:

- LLaма3 8B: Модель LLaма 3 8B, разработанная Meta, представляет собой авторегрессивную большую языковую модель с 8 миллиардами параметров, оптимизированную для диалога и различных задач генерации естественного языка [12].

- Mistral 7B: Mistral 7B — это языковая модель с 7 миллиардами параметров, разработанная Mistral AI, известная своей превосходной производительностью и эффективностью, особенно по сравнению с более крупными моделями, такими, как Llama 2 13B и Llama 1 34B, в различных тестах, включая задачи рассуждения, математики и генерации кода [13].

- Gemma 7B: Gemma 7B — это современная модель с открытым языком, разработанная Google, которая включает 7 миллиардов параметров и демонстрирует превосходную производительность в таких тестах, как MMLU, HellaSwag и HumanEval, часто превосходя более крупные модели, такие как Llama 2 13B [14].

Технические требования для экспериментов:

- ОЗУ (RAM): Минимум 16 ГБ.
- Графический процессор (GPU): Высокопроизводительный GPU с минимум 8 ГБ видеопамяти (VRAM), предпочтительно NVIDIA с поддержкой CUDA.
- Место на диске: Около 4-6 ГБ для каждого модельного файла.

В этом исследовании использовалась следующая конфигурация:

- ОЗУ (RAM): 32 ГБ.
 - Графический процессор (GPU): NVIDIA GeForce RTX 3070 8GB.
 - Место на диске: SSD Samsung 100 гигабайт своводное пространство.
-

2. Применение моделей для выявления провокационных комментариев

4 декабря 2022 года на центральной площади Чингисхана в Улан-Баторе, столице Монголии, состоялся протест под названием "Кража угля", организованный против ущерба в размере 44 триллионов тугриков, нанесенного государству. В ходе этого протеста комментарии, опубликованные под статьями и прямыми эфирами на Facebook, связанные с данным событием, были проанализированы с использованием базы данных MOSINT. Для выявления провокационных комментариев, подстрекающих к незаконным действиям, был проведен анализ с использованием моделей с открытым исходным кодом, таких как Llama3 8B, Mistral 7B и Gemma 7B.

В базе данных MOSINT с 18 часов вечера 4 декабря 2022 года было зарегистрировано всего 19418 комментариев из них были отобраны комментарии, написанные исключительно кириллицей, и классифицировано 1448 комментариев. Эти комментарии представляли собой разнообразные мнения, включая провокационные (P), поддерживающие (S) и нейтральные (N) комментарии, что позволило провести всесторонний анализ.

Подготовка данных: Собранные комментарии были предварительно очищены от шума, такого как избыточные пробелы, специальные символы и несущественные для анализа части текста. После очистки данные были классифицированы вручную по степени провокационности. Классификация включала три категории:

- Провокационные комментарии (P): Комментарии, содержащие агрессивные или подстрекательские высказывания.
 - Поддерживающие комментарии (S): Комментарии, выражающие поддержку движению или позитивное мнение.
 - Нейтральные комментарии (N): Комментарии, не выражающие явного мнения или эмоциональной окраски.
-

Применение моделей: Для классификации комментариев использовались три модели:

1. Llama3 8B
2. Mistral 7B
3. Gemma 7B

Каждая из моделей была настроена на классификацию подготовленных данных. Процесс включал загрузку предварительно очищенных данных в каждую модель и выполнение классификации.

В ходе исследования для классификации комментариев использовался скрипт на Python. Скрипт выполняет следующие шаги:

1. Загружает комментарии из CSV-файла.
2. Генерирует запрос для модели, чтобы классифицировать каждый комментарий как нейтральный, поддерживающий или провокационный.
3. Использует выбранную модель (gemma, llama3, mistral) для классификации комментария.
4. Сохраняет результаты классификации в новый CSV-файл.

Этот скрипт эффективно демонстрирует процесс классификации комментариев с использованием больших языковых моделей и может быть адаптирован для различных целей анализа текста в социальных сетях.

Используемый промт для классификации комментариев был следующим:

Отнесите следующий комментарий к одной из трех категорий: нейтральный (N), поддерживающий (S), провокационный (P).

Комментарий: "{comment}"

Категории:

1. Нейтральный: Комментарий не выражает какого-либо сильного положительного или отрицательного мнения.
-



2. Поддерживающий: Комментарий выражает положительное или поддерживающее мнение.

3. Провокационный: Комментарий выражает отрицательное или провокационное мнение.

Укажите категорию классификации: ""

Примеры провокационных комментариев, классифицированных каждой моделью:

Llama3 8B:

1. "Эта полиция должна перестань защищать воров сейчас, пусть им умереть в стыде, как они не могут жить в этой стране или следовать за ворами." (P)

2. "Есть много профессий, на которых платят больше, чем зарплата полицейского, увольняйтесь с работы" (P)

Mistral 7B:

1. "Полицейским не платят достаточно, чтобы воевать" (P)

2. "Пришло время убить злобных мошенников, которые используют полицию в качестве защитных псов". (P)

Gemma 7B:

1. "Исполни свой долг. Воры ведут вас". (P)

2. "Разве эти полицейские не принесли присягу защищать 76?" (P)

3. Результаты и их анализ

Таблица № 1

Модель	Провокационные (P)	Поддерживающие (S)	Нейтральные (N)
Llama3 8B	482	769	196
Mistral 7B	638	342	467
Gemma 7B	532	424	491

График иллюстрирует распределение провокационных комментариев (P) по категориям (поддерживающие – S, нейтральные – N), как они были классифицированы тремя различными моделями: Llama3 8B, Mistral 7B и Gemma 7B. Контрольная выборка включала 554 провокационных комментария.

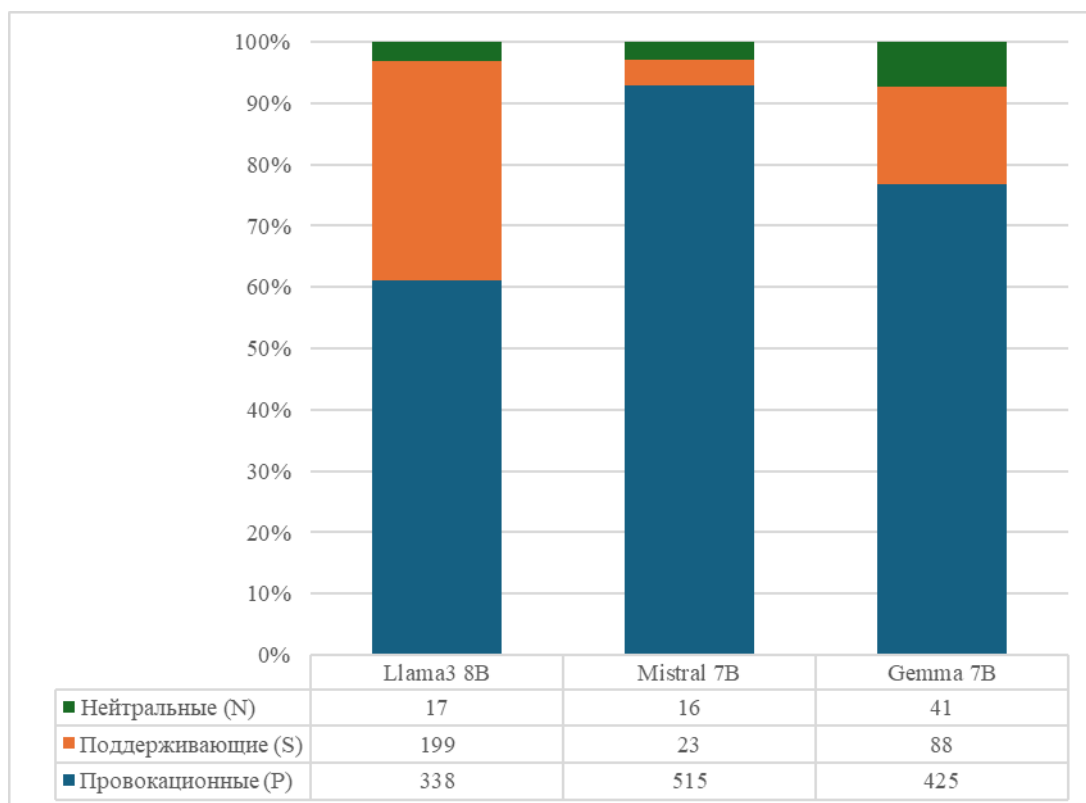


Рис. 2. – Распределение Провокационных комментариев (P) по категориям для различных моделей

Этот чарт показывает, как каждая модель перераспределила контрольные провокационные комментарии.

Для оценки эффективности моделей использовались следующие метрики:

- Точность (Precision): Определяет долю правильно классифицированных провокационных комментариев среди всех комментариев, классифицированных как провокационные.

- Полнота (Recall): Определяет долю правильно классифицированных провокационных комментариев среди всех реальных провокационных комментариев.

- F1-мера: Гармоническое среднее между точностью и полнотой, предоставляющее общую оценку производительности модели.

Таблица № 2

Результаты оценки по моделям

Модель	Точность (Precision)	Полнота (Recall)	F1-мера
Llama3 8B	0.701	0.610	0.657
Mistral 7B	0.807	0.930	0.864
Gemma 7B	0.799	0.767	0.783

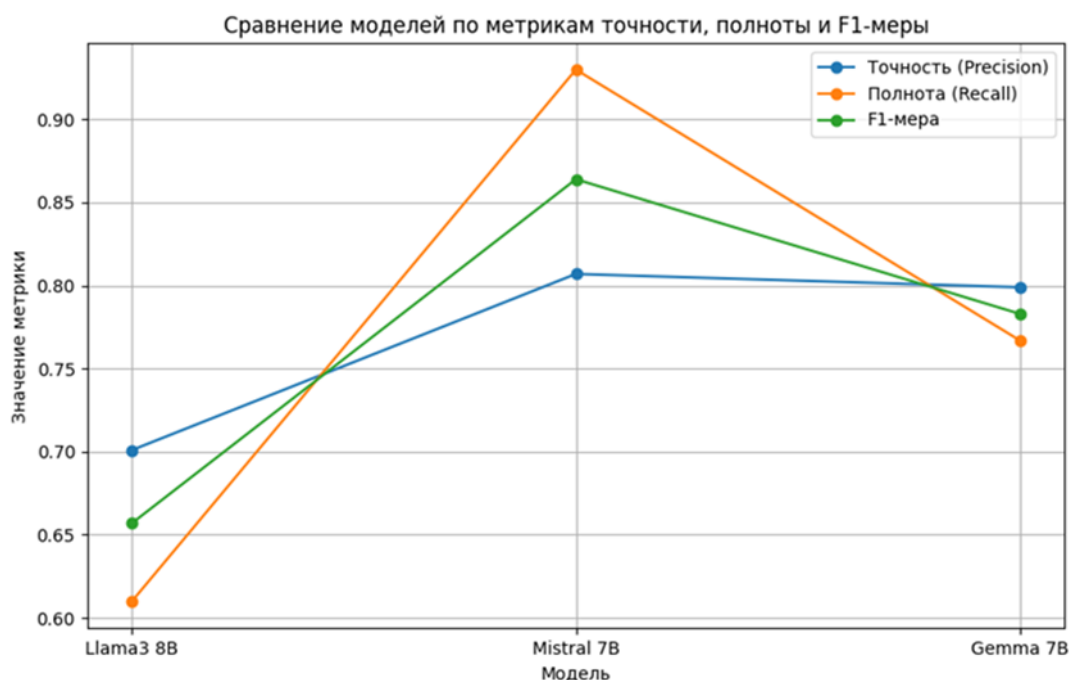


Рис. 3. – Сравнительный анализ эффективности моделей по выбранным метрикам

Результаты экспериментов показывают, что модель Mistral 7B является наиболее точной и эффективной для классификации провокационных комментариев среди рассмотренных моделей. Высокие значения метрик

точности, полноты и F1-меры делают её предпочтительным выбором для задач анализа и мониторинга социальных сетей.

Заключение

В ходе данного исследования была разработана и оценена методология использования больших языковых моделей для выявления провокационных комментариев в социальных сетях во время массовых протестов. Основные результаты и выводы исследования включают следующее:

1. Разработка и выбор моделей:

1.1. В исследовании были применены три современные большие языковые модели: LLaMA3 8B, Mistral 7B и Gemma 7B. Эти модели продемонстрировали высокую эффективность в анализе и классификации текстов из социальных сетей.

1.2. Проведенный анализ показал, что Mistral 7B является наиболее точной и эффективной моделью для выявления провокационных комментариев среди рассмотренных моделей. Она достигла наивысших значений метрик точности (0.807), полноты (0.930) и F1-меры (0.864).

2. Методология и сбор данных:

2.1. Для экспериментов использовались данные из базы данных MOSINT, включающие комментарии, связанные с протестами в Монголии 4 декабря 2022 года. Данные были тщательно подготовлены, включая этапы очистки, аннотации и предобработки.

2.2. Модели LLaMA3 8B, Mistral 7B и Gemma 7B были применены для классификации комментариев, разделенных на три категории: провокационные, поддерживающие и нейтральные.

3. Эксперименты и анализ:

3.1. Эксперименты показали, что Mistral 7B продемонстрировала наиболее высокую эффективность в выявлении провокационных комментариев,

что делает её предпочтительным выбором для задач анализа и мониторинга социальных сетей.

3.2. Результаты классификации подтвердили высокую точность и полноту модели Mistral 7B, что важно для достижения баланса между избеганием ложных срабатываний и упущением важных комментариев.

Результаты нашего исследования подтверждают, что применение БЯМ, таких как LLama3 8B и Mistral 7B, значительно повышает точность анализа комментариев в социальных сетях. Однако, как подчеркивает Лягошина, внедрение таких моделей также сопровождается вызовами, связанными с контролем качества и этическими аспектами применения ИИ в публичном пространстве [15]. Это требует ответственного подхода к использованию БЯМ для обеспечения достоверности и объективности анализа.

Выводы

1. Эффективность больших языковых моделей:

1.1. Применение больших языковых моделей, таких как LLama3 8B, Mistral 7B и Gemma 7B, значительно улучшает точность и оперативность анализа текстов в социальных сетях. Эти модели могут эффективно классифицировать провокационные комментарии, что имеет важное значение для управления общественными настроениями и предотвращения возможных конфликтов.

2. Применение в реальных условиях:

2.1. Разработанная методология может быть применена в реальных условиях для мониторинга социальных сетей во время массовых протестов и других ситуаций политической нестабильности. Это позволит оперативно выявлять провокационные комментарии и принимать меры для снижения их влияния на общественное мнение.

3. Дальнейшие исследования:

3.1. Текущее исследование выявило ряд направлений для дальнейших исследований, таких как снижение вычислительной сложности моделей и адаптация их для работы с различными языками. Будущие исследования могут предложить более эффективные и менее ресурсоемкие методы анализа текстов в социальных сетях.

В заключение, применение больших языковых моделей для выявления провокационных комментариев в социальных сетях представляет собой перспективное направление исследований, способное существенно повысить точность и эффективность анализа контента в условиях политической нестабильности.

Литература

1. Ларионова А.В., Горчакова О.Ю. Деструктивная коммуникация молодёжи в сети Интернет: социально-политический контекст // Вестник РФФИ. Гуманитарные и общественные науки. 2021. №5. С. 141-150.

2. Асманкин С.А. Применение больших языковых моделей для анализа текстовой информации в управлении рисками // Взаимодействие науки и общества – путь к модернизации и инновационному развитию: сборник статей Международной научно-практической конференции (12 апреля 2024 г., г. Воронеж). Воронеж: ОМЕГА САЙНС, 2024. С. 260-263.

3. Nicholas, G., Bhatia, A. Lost in Translation: Large Language Models in Non-English Content Analysis // Center for Democracy & Technology. 2023. URL: cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis (дата обращения: 09.10.2024).

4. Иеромонах Тихон (Васильев). Манипуляция общественным мнением через виртуальные сообщества // Христианское чтение. 2016. №5. С. 207-222.

5. Брагин А.Д. Анализ образов будущего с помощью языковых моделей через исследование комментариев в социальных сетях // Взаимодействие науки и общества – путь к модернизации и инновационному развитию:



сборник статей IV Всероссийской (национальной) научной конференции с международным участием «Наука, технологии, общество: Экологический инжиниринг в интересах устойчивого развития территорий» (НТО-IV-2023). 2023. С. 175-180.

6. Белов С.Д., Зрелова Д.П., Зрелов П.В., Кореньков В.В. Обзор методов автоматической обработки текстов на естественном языке // Системный анализ в науке и образовании: сетевое научное издание. – 2020. – № 3. – С. 8–22. – URL: sanse.ru/download/401 (дата обращения: 22.10.2024).

7. Тымчук Д.А., Свечкарев В.П. Многоагентное моделирование критических социальных поведений // Инженерный вестник Дона, 2010, №1 URL: ivdon.ru/ru/magazine/archive/n1e2010/175.

8. 1. Тымчук Д.А. Многоагентная модель социальных взаимодействий на основе групповой идентичности // Инженерный вестник Дона. 2011. № 1. URL: ivdon.ru/ru/magazine/archive/n1y2011/402.

9. Ванюков С. NLP vs LLM: A Comprehensive Guide to Understanding Key Differences // Medium. — URL: medium.com/@vaniukov.s/nlp-vs-llm-a-comprehensive-guide-to-understanding-key-differences-0358f6571910.

10. Llama3: 8B // Ollama. — URL: ollama.com/library/llama3:8b.

11. Akaike H. A new look at the statistical model identification // IEEE Transactions on Automatic Control. 1974. Т. 19, № 6. С. 716-723. DOI: 10.1109/TAC.1974.1100705.

12. meta-llama/llama3 // GitHub. — URL: github.com/meta-llama/llama3.

13. Announcing Mistral 7B // Mistral. — URL: mistral.ai/news/announcing-mistral-7b/.

14. Gemma // Google AI. — URL: ai.google.dev/gemma.

15. Лягошина Т.В. Большие языковые модели: влияние на публичный дискурс и общество в целом // Вестник Томского



государственного университета. Философия. Социология. Политология. 2024. № 79. С. 111–124. doi: 10.17223/1998863X/79/11.

References

1. Larionova A.V., Gorchakova O.Yu. Vestnik RFFI. Gumanitarny`e i obshhestvenny`e nauki. 2021. №5. pp. 141-150.
2. Asmankin S.A. Vzaimodejstvie nauki i obshhestva – put` k modernizacii i innovacionnomu razvitiyu: sbornik statej Mezhdunarodnoj nauchno-prakticheskoy konferencii (12 aprelya 2024 g., g. Voronezh). Voronezh: OMEGA SAJNS, 2024. pp. 260-263.
3. Nicholas, G., Bhatia, A. Center for Democracy & Technology. 2023. URL: cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis (date assessed: 09.10.2024).
4. Ieromonax Tixon (Vasil`ev). Xristianskoe chtenie. 2016. №5. pp. 207-222.
5. Bragin A.D. Vzaimodejstvie nauki i obshhestva – put` k modernizacii i innovacionnomu razvitiyu: sbornik statej IV Vserossijskoj (nacional`noj) nauchnoj konferencii s mezhdunarodny`m uchastiem «Nauka, texnologii, obshhestvo: E`kologicheskij inzhiniring v interesax ustojchivogo razvitiya territorij» (NTO-IV-2023). 2023. pp. 175-180.
6. Belov S.D., Zrelova D.P., Zrelov P.V., Koren`kov V.V. Sistemny`j analiz v nauke i obrazovanii: setevoe nauchnoe izdanie. 2020. № 3. pp. 8–22. URL: sanske.ru/download/401 (date assessed: 22.10.2024).
7. Ty`mchuk D.A., Svechkarev V.P. Inzhenerny`j vestnik Dona, 2010, №1 URL: ivdon.ru/ru/magazine/archive/n1e2010/175.
8. Ty`mchuk D.A. Inzhenerny`j vestnik Dona. 2011. № 1. URL: ivdon.ru/ru/magazine/archive/n1y2011/402.
9. Vanyukov S. NLP vs LLM: A Comprehensive Guide to Understanding Key Differences // Medium. URL: medium.com/@vaniukov.s/nlp-vs-llm-a-comprehensive-guide-to-understanding-key-differences-0358f6571910.



10. Llama3: 8B. Ollama. URL: ollama.com/library/llama3:8b.
11. Akaike H. IEEE Transactions on Automatic Control. 1974. Т. 19, № 6. pp. 716-723. DOI: 10.1109/TAC.1974.1100705.
12. meta-llama/llama3. GitHub. URL: github.com/meta-llama/llama3.
13. Announcing Mistral 7B. Mistral. URL: mistral.ai/news/announcing-mistral-7b/.
14. Gemma. Google AI. URL: ai.google.dev/gemma.
15. Lyagoshina T.V. Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sociologiya. Politologiya. 2024. № 79. pp. 111–124. doi: 10.17223/1998863X/79/11.

Дата поступления: 14.09.2024

Дата публикации: 29.10.2024