

Макроэкономическая политика

ОЦЕНКА ПРЕДСКАЗАТЕЛЬНОЙ СИЛЫ ИНВЕСТОРОВ В ОПРЕДЕЛЕНИИ КЛЮЧЕВОЙ СТАВКИ ЦБ РФ С ПОМОЩЬЮ LLM

Борисенко Георгий Александрович

аспирант,

МГУ имени М.В. Ломоносова, экономический факультет

(г. Москва, Россия)

Аннотация

Исследование направлено на прогнозирование решений ЦБ РФ по ключевой ставке через анализ мнений инвесторов в социальной сети «Пульс» (Т-Банк). Несмотря на важность ключевой ставки ЦБ, большинство исследователей ставят целью прогнозирование макроэкономических индикаторов, связанных со ставкой ЦБ, но не ее саму. Для восполнения этого пробела в работе предложен подход, сочетающий краудсорсинговые данные с обработкой большими языковыми моделями (LLM). На основе семантически размеченных постов и комментариев из социальных сетей с помощью YandexGPT-5-Lite-Instruct, построены прогнозы направления изменения ставки за 48 заседаний ЦБ РФ (2019–2025 гг.). Интеграция LLM позволила корректно интерпретировать контекст и эмоциональную окраску неформальных высказываний, избежав ошибок словарных методов. Точность прогнозов на основе данных «Пульса» составила 90%, превысив показатели аналитических агентств в условиях экстренных заседаний.

Ключевые слова: ставка ЦБ РФ, прогнозирование, большие языковые модели, анализ настроений.

JEL коды: C45, C53, E47, E52, G41.

Для цитирования: Борисенко Г.А. Оценка предсказательной силы инвесторов в определении ключевой ставки ЦБ РФ с помощью LLM // Научные исследования экономического факультета. Электронный журнал. 2026. Том 18. Выпуск 2. С. 43-57. DOI: 10.38050/2078-3809-2026-18-2-43-57.

Введение

Решения Центрального банка Российской Федерации (ЦБ РФ) по ключевой ставке оказывают фундаментальное влияние на макроэкономическую динамику, определяя стоимость заемных ресурсов, инфляционные ожидания и инвестиционную активность. Точность прогно-

зов по ставке ЦБ имеет ключевое значение для финансовых рынков, однако большинство существующих исследований сосредоточено на моделировании макроэкономических индикаторов, влияющих на решения регулятора, а не на прогнозировании самой ставки. Среди них инфляция (Абдурахманов, 2023), ВВП, процентные ставки денежных рынков, обменный курс (Шарафутдинов, 2023), безработица (Юревич, Ахмадеев, 2022) и др. В то же время бурное развитие цифровых коммуникаций и появление новых источников данных создают возможности для анализа коллективных ожиданий участников рынка.

В последние годы все больший интерес вызывают краудсорсинговые и поведенческие данные, отражающие настроения инвесторов в социальных сетях и онлайн-форумах. Подобные подходы активно развиваются в зарубежных исследованиях, где показатели сентимента пользователей «Twitter», «Reddit» и аналогичных платформ успешно применяются для прогнозирования рыночных индикаторов (Padmanayana et al., 2021). Однако в российской практике потенциал таких данных до сих пор используется ограниченно. Это особенно заметно в контексте прогнозирования ключевой ставки ЦБ РФ – показателя, который формирует основу денежно-кредитной политики и ожиданий участников финансового рынка.

Настоящее исследование направлено на оценку предсказательной силы мнений инвесторов и формирование подхода к прогнозированию решений ЦБ РФ по ключевой ставке на основе анализа высказываний в социальной сети «Пулс» (Т-Банк). В центре внимания находится вопрос о том, могут ли агрегированные ожидания участников платформы служить достоверным индикатором будущего изменения ставки.

Для анализа текстовых данных используются методы обработки естественного языка с применением больших языковых моделей (LLM), позволяющие учитывать контекст, модальность и эмоциональную окраску неформальных сообщений. Такой подход отличается от словарных методов большей гибкостью интерпретации и позволяет выявлять устойчивые паттерны ожиданий, не сводимые к набору заранее заданных ключевых слов.

Предлагаемый в работе подход расширяет существующие исследования ожиданий финансовых рынков за счет использования краудсорсинговых текстовых данных русскоязычной инвестиционной платформы и их интерпретации с применением больших языковых моделей. В отличие от работ, опирающихся на агрегированные макроэкономические показатели или формализованные экспертные оценки, анализируется неструктурированная информация, отражающая ожидания частных инвесторов в моменте. Тем самым работа дополняет эмпирические исследования в области анализа ожиданий и денежно-кредитной политики новым типом данных и способом их обработки.

Несмотря на относительно ограниченное число наблюдений, обусловленное частотой заседаний ЦБ РФ, анализ краудсорсинговых данных позволяет решать иную по своей природе задачу – фиксировать изменения ожиданий участников рынка в преддверии конкретных решений регулятора. В отличие от макроэкономических временных рядов, ориентированных на долгосрочную динамику, комментарии инвесторов концентрируются вокруг отдельных заседаний и отражают реакцию на текущие информационные сигналы. Это делает подобные данные особенно релевантными для краткосрочного анализа и сопоставления рыночных ожиданий с последующими решениями ЦБ РФ.

1. Обзор литературы

Несмотря на отсутствие статей по прогнозированию ключевой ставки, существует множество статей, связанных с прогнозированием процентной ставки в экономике с помощью различных текстовых данных, которые необходимо рассмотреть.

Благодаря удобному и быстрому API (application programming interface), для парсинга в зарубежной литературе очень популярно изучение сентимента постов из «Twitter» (нынешний «X», заблокирован в РФ). Например, в статье (Yasir et al., 2020) авторы используют модели машинного обучения для предсказания спреда доходности к погашению коротких и длинных государственных облигаций Великобритании, Турции, Китая, Гонконга и Мексики около некоторых значимых для этих стран событий (например, Brexit для Великобритании или протесты в Гонконге). Индексы сентимента строились на ежедневной основе по соответствующим событиям (с помощью фильтрации по времени) комментариям пользователей с помощью различных словарей, которые оценивают позитивность и негативность текстов по среднему сентименту всех слов. В результате авторы приходят к выводу, что комментарии пользователей «Твиттера» содержат полезную информацию для построения прогноза.

Анализом сентимента для оценки неожиданности решения ФРС по ставке занимаются в статье (Schmansk et al. 2023). Неожиданность изменения ставки оценивают по изменению доходностей государственных облигаций. В качестве источника текстовых данных используются сразу 3 для сравнения результатов: комментарии пользователя «Твиттера», новостные заголовки и опубликованные FED и FOMC материалы. Индексы сентимента строились, как и в прошлом исследовании, по словарям. В результате исследования авторы выясняют, что наибольшую корреляцию сентимента и изменения доходностей (неожиданностей) имеют индексы сентимента, построенные по заявлениям регулятора. Тем не менее в результате проведения углубленного исследования выясняется, что пользователи «Твиттера» также имеют высокую предсказательную силу, и индексы сентимента, построенные по комментариям пользователей, дополняют, а не заменяют индексы, построенные по новостям и официальным публикациям.

Задача оценки сентимента – очень популярная задача у исследователей, работающих с текстами. На английском языке существует множество словарей для данных различных модальностей: для финансовых новостей или обычных комментариев в соцсетях. Но под сентиментом текста в зависимости от задачи можно понимать немного разные вещи: сентимент комментариев к посту (Xu et al., 2019), сентимент отзывов на товары (Jagdale et al., 2019), сентимент новости для динамики акций (Darapaneni et al., 2022). Более того, источники, по которым создан словарь и на которых он применяется, могут отличаться, что делает словари не универсальными (Islam, Zibran, 2017). Таким образом, невозможно составить словарь, который подходил бы под каждую задачу в мире, нужен более универсальный способ.

Для решения проблемы универсального выделителя сентимента из текстов можно обратиться к передовым разработкам ИИ. Существующие языковые чат-модели могут быть использованы для выполнения любой задачи с текстом. Правда, чем идейно сложнее задача, тем больше потребуется модель (Wei et al., 2022). Но выделение сентимента текста – далеко не самая сложная задача, что показано в исследовании (Luo, Gong, 2024), в котором авторы с помощью большой языковой чат-модели (англ. LLM – Large Language Model) пытаются решить задачу классификации финансовых новостей. Так как данная модель является чат-моделью

лью для выполнения инструкций, авторы помимо самого текста новости подают промпт, который описывает, какую задачу модели предстоит решить и в каком формате необходимо дать ответ. В результате авторам удастся существенно улучшить точность классификации по сравнению с прошлыми исследованиями.

2. Данные и методология¹

2.1. Текстовые данные

Для того, чтобы определить мнение инвесторов по направлению изменения ключевой ставки перед заседанием ЦБ РФ, необходимо гарантировать, что комментарии оставлены именно инвесторами, поэтому был выбран соответствующий источник данных – социальная сеть для инвесторов от «Т-Банка» «Пультс». Для регистрации в данной социальной сети обязательно иметь брокерский счет, то есть по умолчанию гарантируется, что анализируется мнение именно инвесторов (по крайней мере с юридической стороны вопросы участники соцсети ими являются).

С помощью самостоятельно реализованной парсинговой программы на языке python с применением пакета requests были получены ссылки на все посты в социальной сети «Пультс» с первого дня ее существования (3 сентября 2019) по 13 апреля 2025 г. Среди собранных 6037922 ссылок не все посты были опубликованы в открытых каналах, поэтому по итогу было получено 5676832 постов. Сбор данных занял в совокупности 2 месяца. Скорость сбора ограничивается возможными временными блокировками от «Т-Банка».

Каждое полученное наблюдение представляет из себя следующий набор данных:

- 1) никнейм автора поста;
- 2) id поста;
- 3) дату и время выхода поста;
- 4) текст поста;
- 5) тикеры инструментов, которые упоминаются в посте;
- 6) цены инструментов из поста на момент выхода поста;
- 7) количество комментариев к посту;
- 8) реакции к посту (различные смайлы);
- 9) количество реакций к посту;
- 10) даты и времена выхода последних 5 комментариев к посту;
- 11) последние 5 комментариев к посту.

Не все перечисленные данные используются в исследовании, а только текст поста, комментарии и даты. Остальные данные указаны в целях возможной дальнейшей кооперации с читателем; получить полный набор данных можно, обратившись к автору исследования по контактам, указанных в метаданных статьи.

Также стоит упомянуть, почему получены только последние 5 комментариев к каждому посту. Большинство современных сайтов, в том числе и сайт «Пультс», имеют динамическую структуру, т. е. подгружают дополнительную информацию по мере взаимодействия пользователя со страницей сайта, а не возвращают сразу всю информацию при обращении по ссылке.

¹ Весь код для сбора и обработки данных лежит в Github репозитории по ссылке: https://github.com/BorisenkoGeorgy/Pulse_key_rate_forecast.

Чтобы получить все комментарии, пришлось бы взаимодействовать со страницей каждого поста. Это возможно через пакет `selenium`, но требует открытия браузера в явном виде, что существенно вычислительно тяжелее (а следовательно, и дольше). Примерно четверть всех постов в пульсе имеет более 5 комментариев к ним.

Брать в анализ комментарии к постам также необходимо, так как в них зачастую происходит дискуссия по поводу основного поста. Соответственно, если речь в посте о ставке ЦБ, то скорее всего комментарии также будут посвящены ставке.

2.2. Исторические данные по ставке ЦБ

Данные по историческим значениям изменения ставки ЦБ РФ были получены с официального сайта². За время публикации ставки ЦБ РФ взято время выхода официального пресс-релиза. В большинстве случаев это 13:30 (за исключением внеочередных заседаний). Время изменения ставки для них взято из первого появления информации в медиа. (Автор обладает базой данных новостей из телеграмм-каналов, посвященных экономике.) С 3 сентября 2019 по 14 марта 2025 г. было проведено 48 заседаний ЦБ РФ, столько величин ставки и было получено.

2.3. Исторические данные по прогнозируемой аналитиками ставки ЦБ

Прогнозные значения аналитиков – диапазон значений, но для сравнения с предсказаниями инвесторов из «Пульса» необходимо только одно число. Начиная с середины 2020 г. медиана прогнозных значений ведущих аналитических агентств публикуется в популярном телеграмм-канале «MarketTwits»³. Для заседаний, проведенных до середины 2020 г., прогнозные ставки были получены из СМИ, где указывалось, что прогноз выполнен аналитиками «Сбербанка»⁴.

2.4 Получение семантической разметки

Разумеется, использовались не все посты из «Пульса» для выяснения способности инвесторов предсказывать направление движения ставки ЦБ РФ, так как очевидно, что далеко не каждый пост и комментарий упоминают ставку. Поэтому для начала необходимо отфильтровать данные, которые подразумевают ближайшее заседание ЦБ и содержат мнение по ключевой ставке.

Фильтрация комментариев проводилась в три этапа:

- 1) по времени;
- 2) по ключевым словам;
- 3) по смыслу.

Так как даты заседаний и время публикаций пресс-релизов известны, автор рассматривал только посты и комментарии за неделю до каждого заседания. Например, относительно пресс-релиза, который был опубликован в 13:30 21 марта 2025 г., учитывались только данные в недельном окне с 13:30 14 марта 2025 г. по 13:29 21 марта 2025 г. Такой подход позволяет гарантировать, что подавляющее большинство постов и комментариев упоминают именно ближайшее заседание ЦБ РФ.

² https://www.cbr.ru/hd_base/keyrate/ (дата обращения 13.05.2025).

³ <https://t.me/markettwits> (дата обращения 13.05.2025).

⁴ <https://www.rbc.ru/finances/22/10/2019/5daee9689a7947bb17344689> (дата обращения 13.05.2025).

Затем была произведена фильтрация по ключевым словам, чтобы убрать однозначно нерелевантные для исследования комментарии. Для этого был использован ряд ключевых слов, которые учитывали, что ставки бывают не только ключевые, но и ставки по вкладам, кредитам, свопам маржинальной торговле, оплаты труда и в казино (довольно частое упоминание у инвесторов слово «ставка» подразумевает ставку на рост или падение). Более того, ставки есть и в других странах помимо России. Различные ключевые слова смогли учесть часть возможных паттернов использования слова «ставка», когда явно подразумевается не ключевая ставка ЦБ РФ. Полный список ключевых слов можно найти в Приложении 1.

Большинство процентных ставок в экономике тесно связано со ставкой ЦБ. Автор понимает, что исключая комментарии по поводу этих ставок, скорее всего исключится часть комментариев, которые все-таки связаны с будущей ставкой ЦБ. Тем не менее учесть косвенные связи на данный момент не представляется возможным.

Более того, анализировать их связь с ключевой ставкой ЦБ несколько сложнее. Например, изменение доходностей государственных облигаций не обязательно говорит о скором изменении ставки ЦБ. При этом изменение ставки ЦБ почти всегда приводит к изменению доходностей (Абрамов et al., 2022). А например, провести анализ комментариев в стиле «Жду роста ставки ЦБ РФ» существенно проще, так как направление движения ставки указано в явном виде (но далеко не все посты и комментарии в датасете настолько простые).

Несмотря на широкий круг ключевых слов, которые использовались для фильтрации, не для всех возможных упоминаний слова «ставка» в подходящем под исследование контексте удалось отфильтровать посты и комментарии с помощью ключевых слов.

Чтобы провести дальнейшую очистку от нерелевантных наблюдений, автор воспользовался большой языковой чат-моделью (LLM) YandexGTP5-Lite-Instuct, веса которой находятся в открытом доступе. Подробное описание принципов работы чат-моделей не приложено, так как для качественного объяснения потребуется существенно увеличить объем статьи, но можно ознакомиться со статьей (Naveed et al., 2023), где относительно кратко и качественно все описано.

YandexGTP5-Lite-Instuct – современная генеративная большая языковая модель для русского языка на основе нейросетевого блока трансформера, разработанная компанией «Яндекс» и опубликованная в открытый доступ. «Lite» в ее названии обозначает, что это маленькая модель, которая состоит из 8 млрд параметров. По тестам, проведенным разработчиками модели, данная модель является лучшей моделью для русского языка среди моделей сопоставимого размера от других мировых разработчиков⁵. «Instuct» в названии обозначает, что модель может выполнять инструкции, которые ей задает пользователь.

Несмотря на то, что данная модель считается «маленькой» по меркам больших языковых моделей, число параметров в которых может достигать 2 трлн⁶, для несложных задач с небольшим количеством текста, она показывает неплохие результаты, это видно в отчете разработчиков. Что не менее важно, имея графический ускоритель (GPU – Graphic Processing Unit) с ней можно работать на локальном компьютере или в удаленной среде совершенно бесплатно и с относительно быстрой скоростью генерации ответа.

⁵ <https://habr.com/ru/companies/yandex/articles/885218/> (дата обращения: 13.05.2025).

⁶ <https://ai.meta.com/blog/llama-4-multimodal-intelligence/> (компания «Meta» является экстремистской организацией и заблокирована на территории РФ) (дата обращения: 13.05.2025).

Для большей скорости работы модели была выбрана ее квантизованная версия, где каждый параметр модели занимает 4 бита, вместо стандартных 16. Это позволяет увеличить скорость работы модели в 4 раза при ограничениях в памяти графического процессора с минимальными потерями в качестве. «Яндекс» также предоставляет официальную квантизованную версию своей модели⁷.

Итак, чтобы выделить посты и комментарии про ключевую ставку ЦБ, они были переданы модели с промптом, который описывал, что конкретно модели подается на вход и какого вида должен быть ответ. Также для того, чтобы сэкономить вычислительные ресурсы, промпт был расширен второй задачей. Если пост или комментарий относится к ставке ЦБ, то модели необходимо было выделить, а что же автор думает о направлении изменения ключевой ставки ЦБ РФ: ставка вырастет, не изменится, снизится или из текста нельзя сказать, что произойдет со ставкой. Полный промпт можно найти в Приложении 2.

Зачем давать дополнительную задачу в большую языковую модель, если в теории посты и комментарии можно разметить по ключевым словам? Если в тексте про ставку встречается слово «поднимет», «вырастет» и аналогичные, то, значит, ЦБ поднимет ставку. Данный подход также может работать, но он совершенно точно не будет справляться с сложными текстами, например: «Акции с потенциалом роста 50%! Европлан. Цену сильно укатали на ожиданиях плохого отчета, но за 15 лет своей истории финпоказатели компании растут ударными темпами. Учитывая то, что ранее прибыль росла (в условиях высокой ключевой ставки), то и сейчас ничего не изменится...».

Данный пост упоминает ключевую ставку ЦБ РФ, но из его текста совершенно невозможно понять направление изменения ставки. Если использовать ключевые слова, то данный текст получит разметку «ставка вырастет» за счет наличия слова «рост», хотя подразумевается рост акций, а не ставки. В случае обработки через LLM для данного комментария было получено, что он относится к ключевой ставке ЦБ, но по нему нельзя в явном виде сказать, в каком направлении автор ожидает изменение ставки.

Обработка данных с помощью большой языковой модели проводилась с помощью пакета llama_cpp, реализованного для языка python на соревновательной платформе Kaggle⁸, которая бесплатно предоставляет вычислительные ресурсы для пользователей платформы как для участия в соревнованиях, так и для личных целей.

После 15 часов работы GPU P100 была получена финальная разметка датасета. Итоговый датасет состоит из 8342 постов и 6974 комментариев, где участники соцсети выразили свое мнение по поводу траектории ставки ЦБ. То есть каждый текст в датасете относится к ключевой ставке и выражает явное мнение по поводу ее повышения, понижения или неизменности.

Далее, по полученной разметке были подсчитано общее количество комментариев и постов, где высказывалось мнение за повышение, сохранение или снижение ключевой ставки ЦБ РФ. Прогнозное значение изменения было выбрано как наиболее частотный ответ, например, если всего было обработано 100 текстов, среди которых в 60 текстах высказываются за повышение, в 30 – за сохранение, а в 10 – за снижение, то прогнозным значением считалось повышение ключевой ставки ЦБ РФ.

⁷ <https://huggingface.co/yandex/YandexGPT-5-Lite-8B-instruct-GGUF> (дата обращения: 13.05.2025).

⁸ <https://www.kaggle.com/> (дата обращения: 13.05.2025).

3. Результаты

В результате было получено, что инвесторы (участники соцсети «Пульс») очень плохо справляются с прогнозированием направления изменения ключевой ставки ЦБ РФ на ближайшем заседании. Из 48 заседаний, которые были проведены в промежутке между 10 октября 2019 г. и 21 марта 2025 г. инвесторы угадали направление ДКП в 43 случаях, т. е. точность предсказания предложенным методом составила практически 90%. Инвесторы ошиблись:

- 1) 7 февраля 2020 г., предсказав сохранение ставки с вероятностью 67% (снижение 33%), в то время как ее снизили;
- 2) 20 марта 2020 г., предсказав повышение ставки с вероятностью 46% (сохранение 41%), в то время как ее сохранили;
- 3) 19 июня 2020 г., предсказав сохранение ставки с вероятностью 46% (снижение 39%), в то время как ее снизили;
- 4) 18 марта 2022 г., предсказав повышение ставки с вероятностью 56%, (сохранение 36%), в то время как ее сохранили;
- 5) 20 декабря 2024 г., предсказав повышение ставки с вероятностью 44% (сохранение 40%).

Из полученных результатов можно видеть, что если инвесторы и ошибались, то в случаях 2, 3 и 5 разница между предсказанной вероятностью и вероятностью для правильного направления ДКП оказалась не очень большой. Более того, для случаев 1, 2, 3 получено очень мало наблюдений – 12, 37 и 59 соответственно. Возможно, имея больший набор данных, вероятности предсказаний изменились бы. Малое число наблюдений связано с тем, что «Пульс» на тот момент существовал от 5 до 9 месяцев (в зависимости от заседания) и был еще не очень популярен среди инвесторов, поэтому и комментариев про ставку ЦБ РФ для данного интервала получено немного.

Можно вычислить аналогичную величину для прогнозов аналитиков. Но важно отметить, что прогнозы предоставлялись только для заседаний ЦБ РФ, о проведении которых было объявлено заранее (таких 46). То есть из анализа исключаются заседания 28 февраля 2022 г. и 8 апреля 2022 г., так как в то время экономика РФ переживала сильный экономический шок, и у ЦБ РФ не было времени заранее объявлять о проведении заседаний, следовательно, аналитики не имели возможности к ним подготовить свои прогнозы. Итого, точность прогнозов по направлению движения ставки ЦБ РФ у аналитиков составила примерно 93,5%.

Они ошиблись:

- 1) 19 марта 2021 г., предсказав сохранение, в то время как ставку повысили (хотя инвесторы в «Пульсе» справились успешно, предсказав повышение с вероятностью 52% (45% за сохранение); всего на данную дату было обработано 124 комментария);
- 2) 18 марта 2022 г., предсказав повышение, в то время как ставку сохранили (инвесторы здесь тоже ошиблись, что описано выше);
- 3) 20 декабря 2024 г., предсказав повышение, в то время как ставку сохранили (инвесторы здесь тоже ошиблись, что описано выше).

Таким образом, получено, что частота верно угаданных решений по ставке у профессиональных аналитиков выше. Однако данные цифры нельзя сравнивать между собой напрямую ввиду ряда ограничений:

- 1) на начальном этапе жизни «Пульса» в нем было зарегистрировано мало пользователей, что делает результаты для заседаний в 2019 и 2020 гг. не репрезентативными (как описано выше, 3 из 5 ошибок инвесторы допустили именно в 2020 г.);
- 2) аналитики не делали предсказания на экстренные внеплановые заседания, в то время как предложенным методом через анализ комментариев в «Пульсе» для таких заседаний можно строить прогноз, и он оказался правильным.

Заключение

Проведенное исследование показало, что коллективные ожидания инвесторов, выраженные в публичных комментариях, содержат значимую информацию о будущих решениях Центрального банка Российской Федерации по ключевой ставке. Применение больших языковых моделей (LLM) для семантического анализа текстов позволило выявить направленность ожиданий и продемонстрировало высокую предсказательную силу краудсорсинговых данных.

Предложенный подход может рассматриваться как альтернативный инструмент мониторинга рыночных настроений. Для инвесторов и трейдеров он способен дополнять фундаментальный анализ, прежде всего в краткосрочном горизонте. Для банков и управляющих компаний данный подход может быть использован при оценке изменений рыночных ожиданий в рамках задач риск-менеджмента. Анализ публичных ожиданий также позволяет оценивать, насколько решения регулятора интерпретируются и понимаются участниками рынка.

Научный вклад исследования заключается в демонстрации того, что совмещение методов искусственного интеллекта и краудсорсинговых данных открывает новое направление в прогнозировании денежно-кредитной политики. Практическая значимость заключается в создании инструмента, способного служить индикатором рыночных ожиданий в режиме реального времени.

Ограничения работы связаны с малым количеством заседаний ЦБ РФ (48) за период с начала существования социальной сети «Пульс» по дате отсечения исследования. Данное ограничение никак нельзя преодолеть в связи с тем, что заседания ЦБ как правило проводятся ограниченное число раз в год. Тем не менее весь код исследования выложен в открытый доступ, это позволит в случае необходимости дополнительно проверить результаты исследования в будущем. Также небольшим ограничением является зависимость точности анализа сентимента от выбора модели. Предложенная модель является лишь одной из многих на выбор.

Предложенный подход может быть использован в качестве прикладного инструмента анализа ожиданий в периоды подготовки и проведения заседаний ЦБ РФ. В частности, агрегированные оценки сентимента инвесторов могут применяться для оперативного сопоставления ожидаемого направления изменения ключевой ставки с консенсус-прогнозами аналитических агентств за несколько дней или недель до заседания регулятора. Такое сопоставление позволяет выявлять расхождения между формализованными экспертными оценками и децентрализованными ожиданиями рынка, а также фиксировать изменения ожиданий в динамике по мере поступления новой информации, что может быть полезно при управлении портфелями ценных бумаг и риск-менеджменте.

Также полученные индикаторы могут использоваться в качестве вспомогательной переменной при построении краткосрочных прогнозных моделей или сценарных оценок, дополняя традиционные макроэкономические факторы информацией о текущих ожиданиях рынка. Такой инструментарий может быть востребован в исследовательских и аналитических задачах, связанных с анализом ожиданий и оценкой восприятия решений денежно-кредитной политики.

Хотя в рамках данного исследования метод был апробирован на примере решений ЦБ РФ по ключевой ставке, его применение не ограничивается исключительно данным типом событий. Аналогичный подход может быть использован для анализа ожиданий участников рынка в отношении других регулярно публикуемых макроэкономических показателей, таких как недельные данные по инфляции, ежемесячные оценки динамики ВВП или показатели рынка труда. В таких случаях краудсорсинговые данные позволяют фиксировать реакцию и ожидания инвесторов в привязке к конкретным датам публикации статистики, что расширяет возможности краткосрочного анализа и сопоставления рыночных ожиданий с фактическими значениями показателей.

Таким образом, перспективными направлениями дальнейших исследований являются расширение анализа на другие макроэкономические показатели, а также использование данного подхода как инструмента для построения индикаторных переменных для уточнения макроэкономических показателей.

Список литературы

Абдурахманов М.А. Моделирование влияния прогнозов ключевой ставки Банка России на ожидания участников рынка // Деньги и кредит. 2023. Т. 82. № 2. С. 3–20.

Абрамов В., Тишин А.В., Стырин К.А. Денежно-кредитная политика и кривая доходности // Серия докладов об экономических исследованиях. 2022. № 95В. С. 49.

Шарафутдинов А.Р. Прогнозирование российских ВВП, инфляции, ставки процента и обменного курса с помощью модели DSGE-VAR // Деньги и кредит. 2023. Т. 82. № 3. С. 32–86.

Юревич М.А., Ахмадеев Д.Р. Возможности прогнозирования уровня безработицы на основе анализа статистики запросов (в поисковых системах) // Terra Economicus. 2021. Т. 19. № 3. С. 53–64. <https://doi.org/10.18522/2073-6606-2021-19-3-53-64>.

Darapaneni N., Paduri A.R., Sharma H. et al. Stock price prediction using sentiment analysis and deep learning for Indian markets // arXiv preprint arXiv:2204.05783. 2022. <https://doi.org/10.48550/arXiv.2204.05783>.

Islam M.R., Zibran M.F. A comparison of dictionary building methods for sentiment analysis in software engineering text // 2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). IEEE, 2017. P. 478–479. <https://doi.org/10.1109/ESEM.2017.67>.

Jagdale R.S., Shirsat V.S., Deshmukh S.N. Sentiment analysis on product reviews using machine learning techniques // Cognitive Informatics and Soft Computing: Proceeding of CISC 2017. Springer Singapore, 2019. P. 639–647. https://doi.org/10.1007/978-981-13-0617-4_61.

Luo W., Gong D. Pre-trained large language models for financial sentiment analysis // arXiv preprint arXiv:2401.05215. 2024. P. 6. <https://doi.org/10.48550/arXiv.2401.05215>.

Naveed H., Khan A.U., Qui S. et al. A comprehensive overview of large language models // arXiv preprint arXiv:2307.06435. 2023. P. 47. <https://doi.org/10.48550/arXiv.2307.06435>.

Padmanayana V., Bhavya K. Stock market prediction using Twitter sentiment analysis // Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. 2021. Vol. 7. No. 4. P. 265–270. <https://doi.org/10.32628/CSEIT217475>.

Schmanski B., Scotti C., Vega C., Benamar H. Fed Communication, News, Twitter, and Echo Chambers. 2023. P. 61. <https://doi.org/10.17016/FEDS.2023.036>.

Wei J., Bommasani R., Raffel C. et al. Emergent abilities of large language models // arXiv preprint arXiv:2206.07682. 2022. P. 30. <https://doi.org/arXiv.2206.07682>.

Xu G., Meng Y., Qiu X. et al. Sentiment analysis of comment texts based on BiLSTM // Ieee Access. 2019. Vol. 7. P. 51522–51532. <https://doi.org/10.1109/ACCESS.2019.2909919>.

Yasir M., Afzal S., Latif K. et al. An efficient deep learning based model to predict interest rate using twitter sentiment // Sustainability. 2020. Vol. 12. No. 4. P. 1660–1677. <https://doi.org/10.3390/su12041660>.

Githubb. Код, который использовался для написания статьи: URL: https://github.com/BorisenkoGeorgy/Pulse_key_rate_forecast (дата обращения: 13.05.2025).

Раздел официального сайта ЦБ РФ, где выложена информация про ключевую ставку ЦБ: URL: https://www.cbr.ru/hd_base/keyrate/ (дата обращения: 13.05.2025).

Крупнейший Telegram канал с оперативными новостями: URL: <https://t.me/markettwits> (дата обращения: 13.05.2025).

Новость с официального сайта РБК про прогноз ставки ЦБ РФ аналитиками Сбербанка: URL: <https://www.rbc.ru/finances/22/10/2019/5daee9689a7947bb17344689> (дата обращения: 13.05.2025).

Описание модели YandexGPT5: URL: <https://habr.com/ru/companies/yandex/articles/885218/> (дата обращения: 13.05.2025).

Информация о самой большой языковой модели с официального сайта Meta: URL: <https://ai.meta.com/blog/llama-4-multimodal-intelligence/> (дата обращения: 13.05.2025).

Карточка модели YandexGPT5 Lite Instruct: URL: <https://huggingface.co/yandex/YandexGPT-5-Lite-8B-instruct-GGUF> (дата обращения: 13.05.2025).

Официальный сайт kaggle: URL: <https://www.kaggle.com/> (дата обращения: 13.05.2025).

Приложение 1

Список ключевых слов

Входят слова и регулярные выражения: ставк.

Не входят слова и регулярные выражения (обратная фильтрация): ставка на, ставки на, итоги, youtube, rutube, поставк, доставк, приставк, приставк, заставк, выставк, отставк, загруз, маржа, кред, депоз, вклад, накопитель, счет, купон, компенс, льгот, ипотек, выпуск, япон, америк, фрс, ецб, сша, армен, аргентин, кита, герман, франц, европ, австрал, канад, мексик, украин, белорус, казахс, зеланд, евроз, великобрит, англ, бразил, инди, вьетна, таджикис, азербайдж, киргиз, узбекист, грузи, франц, германи, турц, польш, облиг, своп, ваши ставки, ставок, реклам, (?i)(ставк.*делать|делать.*ставк), (?i)(ставк.*делаю|делаю.*ставк), (?i)(ставк.*сделаны|сделаны.*ставк), (?i)(ставк.*сделал|сделал.*ставк), (?i)(ставк.*делал|делал.*ставк), (?i)(ставк.*делаем|делаем.*ставк), (?i)(ставк.*делайте|делайте.*ставк), (?i)(ставк.*принима|принима.*ставк), (?i)(ставк.*стави|стави.*ставк), (?i)(ставк.*ставл|ставл.*ставк), (?i)(ставк.*ставь|ставь.*ставк), (?i)(выигры.*ставк|выигры.*ставк), повысил, увеличил, понизил, снизил, уменьшил, изменил, оставил, повышена, понижена, изменена, увеличена, уменьшена, сыграл, сработал, ставки риска, спред, заявил, совпало, займ, повлиял, 37%, ставкам, налог, s&p 500, моя ставка, мои ставки.

Приложение 2

Промт, использованный для разметки датасета после фильтрации

«Тебе на вход будет дан комментарий из соцсети для инвесторов “Пульс”. Твоя задача указать с помощью индексов, относится ли данный комментарий к ключевой ставке ЦБ РФ. В комментарии не обязательно упоминается ЦБ РФ, но под ставкой может все равно подразумеваться ключевая ставка.

Также могут быть комментарии, где упоминаются другие ставки, например, ставка оплаты, налоговая ставка, ставки иностранных компаний (в таких постах тогда часто встречается тикер иностранной компании), ставка как в казино. Такие ставки мне неинтересны, надо выделять комментарии именно, где подразумевается ставка ЦБ РФ.

Помимо того, что тебе необходимо выделить комментарии про ставку, тебе также необходимо сказать, в какую сторону изменится ставка ЦБ РФ, если об этом сказано в комментарии. Ставка может вырасти, не измениться или снизиться. Однако непосредственно данных слов может не быть в комментарии, возможны варианты.

Если по тексту не ясно, в каком направлении изменится ставка, так и напиши.

Какие индексы возвращать (то, что должно быть в ответе):

- если текст не относится к ставке ЦБ РФ, верни число 0, и больше ничего не надо (то есть никакого дополнительного текста);
- если текст содержит информацию про ставку ЦБ РФ, в таком случае тебе еще необходимо определить направление изменения ставки, тогда итоговый ответ необходимо представить в формате двух чисел через запятую между ними:

1) первое число – верни 1, так как данный текст про ставку ЦБ РФ;

2) второе число – если говорится, что ставку снизят, то -1 , если не изменят, то 0, если повысят, то 1; если же по тексту нельзя сказать, как изменится ставка, то верни 2.

То есть данный комментарий про ставку ЦБ РФ и по нему можно сказать, что автор ожидает повышение ставки ЦБ. Для этого случая от тебя требуется два числа через запятую и больше ничего.

Важно: больше ничего, кроме одного или двух чисел через запятую не возвращай. Не надо пояснять свой ответ, не надо давать дополнительные комментарии, не надо дублировать слова из исходного комментария, нужны просто число или числа и больше ничего.

Далее приводился текст комментария».

ASSESSING THE PREDICTIVE POWER OF INVESTORS IN DETERMINING THE KEY RATE OF THE CENTRAL BANK OF THE RUSSIAN FEDERATION USING LLM

Georgii A. Borisenko

Postgraduate student,

*Lomonosov Moscow State University, Faculty of Economics
(Moscow, Russia)*

Abstract

The main aim of this study is to forecast key rate decisions of the Central Bank of Russian Federation by analyzing investor opinions in the Pulse social network (T-Bank). Despite the importance of the Central Bank's key rate, most studies aim to forecast macroeconomic indicators related to the Central Bank rate, but not the rate itself. To fill this gap, the paper proposes an approach that combines crowdsourcing data with processing by large language models (LLM). Based on semantically tagged posts and comments from the social network using YandexGPT-5-Lite-Instruct, forecasts directions of rate changes for 48 meetings of the Central Bank of the Russian Federation (2019–2025). The integration of LLM made it possible to correctly interpret the context and emotional coloring of informal statements, avoiding the errors of dictionary methods. The accuracy of forecasts based on Pulse data was 90%, exceeding the indicators of analytical agencies in the context of emergency meetings.

Keywords: Central Bank of the Russian Federation rate, forecasting, large language models, sentiment analysis.

JEL: C45, C53, E47, E52, G41.

For citation: Borisenko, G.A. (2026) Assessing the Predictive Power of Investors in Determining the Key Rate of the Central Bank of the Russian Federation Using LLM. Scientific Research of Faculty of Economics. Electronic Journal, vol. 18, no. 2, pp. 43-57. DOI: 10.38050/2078-3809-2026-18-2-43-57.

References

- Abdurahmanov M.A. Modelirovanie vlijanija prognozov kljuchevoj stavki Banka Rossii na ozhidanija uchastnikov rynka. Den'gi i kredit. 2023. Vol. 82. No. 2. P. 3–20. (In Russ.).
- Abramov V., Tishin A.V., Styurin K.A. Denezhno-kreditnaja politika i krivaja dohodnosti. Serija dokladov ob jekonomicheskikh issledovanijah. 2022. No.95V. P. 49. (In Russ.).
- Sharafutdinov A.R. Prognozirovanie rossijskikh VVP, infljacji, stavki procenta i obmennogo kursa s pomoshh'ju modeli DSGE-VAR. Den'gi i kredit. 2023. Vol. 82. No. 3. P. 32–86. (In Russ.).

Jurevich M.A., Ahmadeev D.R. Vozможnosti prognozirovaniya urovnja bezraboticy na os-nove analiza statistiki zaprosov (v poiskovyh sistemah). Terra Economicus. 2021. Vol. 19. No. 3. P. 53–64. <https://doi.org/10.18522/2073-6606-2021-19-3-53-64>. (In Russ.).

Darapaneni N., Paduri A.R., Sharma H. et al. Stock price prediction using sentiment analysis and deep learning for Indian markets. arXiv preprint arXiv:2204.05783. 2022. <https://doi.org/10.48550/arXiv.2204.05783>.

Islam M.R., Zibran M.F. A comparison of dictionary building methods for sentiment analysis in software engineering text. 2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). IEEE, 2017. P. 478–479. <https://doi.org/10.1109/ESEM.2017.67>.

Jagdale R.S., Shirsat V.S., Deshmukh S.N. Sentiment analysis on product reviews using machine learning techniques. Cognitive Informatics and Soft Computing: Proceeding of CISC 2017. Springer Singapore, 2019. P. 639–647. https://doi.org/10.1007/978-981-13-0617-4_61.

Luo W., Gong D. Pre-trained large language models for financial sentiment analysis. arXiv preprint arXiv:2401.05215. 2024. P. 6. <https://doi.org/10.48550/arXiv.2401.05215>.

Naveed H., Khan A.U., Qui S. et al. A comprehensive overview of large language models. arXiv preprint arXiv:2307.06435. 2023. P. 47. <https://doi.org/10.48550/arXiv.2307.06435>.

Padmanayana V., Bhavya K. Stock market prediction using Twitter sentiment analysis // Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. 2021. Vol. 7. No. 4. P. 265–270. <https://doi.org/10.32628/CSEIT217475>.

Schmanski B., Scotti C., Vega C., Benamar H. Fed Communication, News, Twitter, and Echo Chambers. 2023. P. 61. <https://doi.org/10.17016/FEDS.2023.036>.

Wei J., Bommasani R., Raffel C. et al. Emergent abilities of large language models. arXiv preprint arXiv:2206.07682. 2022. P. 30. <https://doi.org/arXiv.2206.07682>.

Xu G., Meng Y., Qiu X. et al. Sentiment analysis of comment texts based on BiLSTM. Ieee Access. 2019. Vol. 7. P. 51522–51532. <https://doi.org/10.1109/ACCESS.2019.2909919>.

Yasir M., Afzal S., Latif K. et al. An efficient deep learning based model to predict interest rate using twitter sentiment. Sustainability. 2020. Vol. 12. No. 4. P. 1660–1677. <https://doi.org/10.3390/su12041660>.

Githunb. Kod, kotoryj ispol'zovalsja dlja napisaniya stat'i: Available at: https://github.com/BorisenkoGeorgy/Pulse_key_rate_forecast (accessed: 13.05.2025).

Razdel oficial'nogo sajta CB RF, gde vylozhena informacija pro kljuchevuju stavku CB: Available at: https://www.cbr.ru/hd_base/keyrate/ (accessed: 13.05.2025). (In Russ.).

Krupnejshij Telegram kanal s operativnymi novostjami: Available at: <https://t.me/markettwits> (accessed: 13.05.2025). (In Russ.).

Novost' s oficial'nogo sajta RBK pro prognoz stavki CB RF analitikami Sberbanka: Available at: <https://www.rbc.ru/finances/22/10/2019/5daee9689a7947bb17344689> (accessed: 13.05.2025). (In Russ.).

Opisanie modeli YandexGPT5: Available at: <https://habr.com/ru/companies/yandex/articles/885218/> (accessed: 13.05.2025). (In Russ.).

Informacija o samoj bol'shoj jazykovoj modeli s oficial'nogo sajta Meta: Available at: <https://ai.meta.com/blog/llama-4-multimodal-intelligence/> (accessed: 13.05.2025).

Kartochka modeli YandexGPT5 Lite Instruct: Available at: <https://huggingface.co/yandex/YandexGPT-5-Lite-8B-instruct-GGUF> (accessed: 13.05.2025). (In Russ.).

Oficial'nyj sayt kaggle: Available at: <https://www.kaggle.com/> (accessed: 13.05.2025)