

Измерение инфляционных ожиданий на основе данных социальной сети Telegram

Научный руководитель – Соболева Екатерина Николаевна

Дернин Денис Андреевич

Студент (магистр)

Санкт-Петербургский государственный университет, Институт "Высшая школа менеджмента Санкт-Петербург, Россия

E-mail: dernin.des@yandex.ru

В современной экономике большинство центральных банков реализует режим инфляционного таргетирования, устанавливая конкретную количественную цель по инфляции и проводя монетарную политику для ее достижения. Ключевую роль в достижении этой цели играют инфляционные ожидания экономических агентов, влияющие на их потребительское и инвестиционное поведение. Инфляционные ожидания, привязанные к заявленной цели центрального банка, помогают поддерживать ценовую стабильность, а также повышают эффективность монетарной политики. Точная и своевременная оценка инфляционных ожиданий по-прежнему является большим вызовом для экономической науки. Традиционные модели, основанные на официальных опросах потребителей и фирм, обладают такими недостатками, как ограниченность выборки, запаздывание данных и редкая публикация результатов. В данной работе предлагается альтернативный способ измерения инфляционных ожиданий на основе данных из новостных каналов социальной сети Telegram.

Целью данной научно-исследовательской работы является разработка прогнозных моделей инфляционных ожиданий с использованием количественных индикаторов на основе данных из социальных сетей. Объект исследования – инфляционные ожидания экономических агентов, предмет – особенности влияния новостной информации на их формирование. Среди исследований, использующих методы машинного обучения для макроэкономического прогнозирования, выделяются работы С. Angelico и др [3], которые создали индекс инфляционных ожиданий, используя сообщений в Twitter, и исследование В. Щербакова и др. [4], в котором моделируются инфляционные тренды в Омской области на основании данных ВКонтакте.

В данной работе мы исходим из предпосылки, что негативный или позитивный сигнал, содержащиеся в новостных сообщениях экономической тематики, передается экономическим агентам и формирует их инфляционные ожидания. Для изучения этого эффекта был разработан Индекс настроений экономических новостей (ИНЭН), основанный на анализе сообщений из 28 новостных Telegram-каналов, включая такие издания, как «Медуза», «The Bell», «РБК», «Интерфакс», «Коммерсант», «Ведомости». За период с 1 января 2022 г. по 17 апреля 2024 г. было собрано 1 043 000 постов в этих каналах. Для сбора сообщений использовалась программа-парсер, написанная на языке Python, с использованием библиотеки telethon.

Полученная выборка была предварительно обработана: из текстов удалены стоп-слова, редкие слова и лишние символы, а также проведена лемматизация. Для отбора сообщений, относящихся к экономической тематике, был применен алгоритм семплирования Гиббса для Дирихле-мультиномиального распределения (GSDMM) [5], разработанный специально для коротких, разреженных текстов большой размерности. Особенностью данного алгоритма является автоматическое определение количества тематических кластеров. После

его применения к составленной выборке было сформировано 27 экономических кластеров и отобрано 89 278 сообщений экономической тематики, что составило 8,6% от общего объема данных. Для оценки эмоционального тона (сигнала) отобранных сообщений использовался Тональный словарь русского языка КартаСловСент [1], содержащий 28 197 слов и выражений русского языка с указанием их эмоциональной окраски и скалярного значения силы заряда в диапазон от -1 до 1. Индекс был рассчитан, как среднее значение эмоционального заряда слов в каждом сообщении, что позволяет оценить общий тон новости: низкие значения индекса указывают на преобладание негативного настроения в сообщении, а более высокие - на доминирование положительного эмоционального заряда.

Для оценки релевантности Индекса настроений экономических новостей (ИНЭН) и его потенциального применения в прогнозных моделях, мы провели сравнение его среднемесячных значений с существующими индикаторами экономических настроений и инфляционных ожиданий потребителей и фирм. Наибольшая корреляция (0,75) была выявлена с Индексом потребительских настроений, тогда как связь с ожидаемым уровнем инфляции [2], рассчитываемым ФОМом на основе опросов потребителей, оказалась менее выраженной (-0,46). При включении ИНЭН в модель линейной регрессией, где ожидаемый уровень инфляции выступает в роли зависимой переменной, коэффициент при ИНЭН оказался статистически значимым на уровне 5%, коэффициент детерминации составил 0,583. Диаграмма рассеяния представлена на рисунке 1, результат регрессии - на рисунке 2

Отрицательное значение коэффициента регрессии, а также коэффициента корреляции между ИНЭН и ожидаемой инфляцией указывают на то, что улучшение настроений экономических сообщений связано со снижением инфляционных ожиданий. Это позволяет рассматривать Индекс настроений экономических новостей как альтернативный высокочастотный индикатор инфляционных ожиданий, который может быть интегрирован в прогнозные модели различных макроэкономических показателей, в частности - уровня инфляции.

В перспективе планируется повысить точность ИНЭН за счет использования больших языковых моделей для оценки тональности сообщений. Кроме того, в планах — переход к ежедневному расчету и публикации индекса, что позволит сделать его более актуальным и полезным для анализа текущих экономических тенденций.

Источники и литература

- 1) Кулагин Д.И. Открытый тональный словарь русского языка КартаСловСен // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конференции «Диалог». Москва: Издательство РГГУ. – 2021. – Выпуск 20. – С. 1106–1119. – URL: <https://github.com/dkulagin/kartaslov/blob/master/README.md> (дата обращения 01.04.2024)
- 2) Инфляционные ожидания и потребительские настроения / Банк России. [Б.м.], - 2024. - №4 (88). URL: https://www.cbr.ru/analytics/dkp/inflationary_expectations/infl_exp_24-04/ (дата обращения 01.05.2024)
- 3) Angelico C. Can We Measure Inflation Expectations Using Twitter? / C. Angelico, J. Marcucci, M. Miccoli, F. Quarta // Bank of Italy Temi di Discussione (Working Papers). – 2021. – N 1318
- 4) Karpov I. Regional inflation analysis using social network data / V. Shcherbakov, I. Karpov // Economy of Region. – 2024. – Vol. 20, № 4. – P. 1-23
- 5) Yin J., Wang J. A Dirichlet Multinomial Mixture Model-Based Approach for Short Text Clustering // Proceedings of the 20th ACM SIGKDD International Conference

on Knowledge Discovery and Data Mining (KDD'14). – Association for Computing Machinery, 2014. – P. 233–242

Иллюстрации

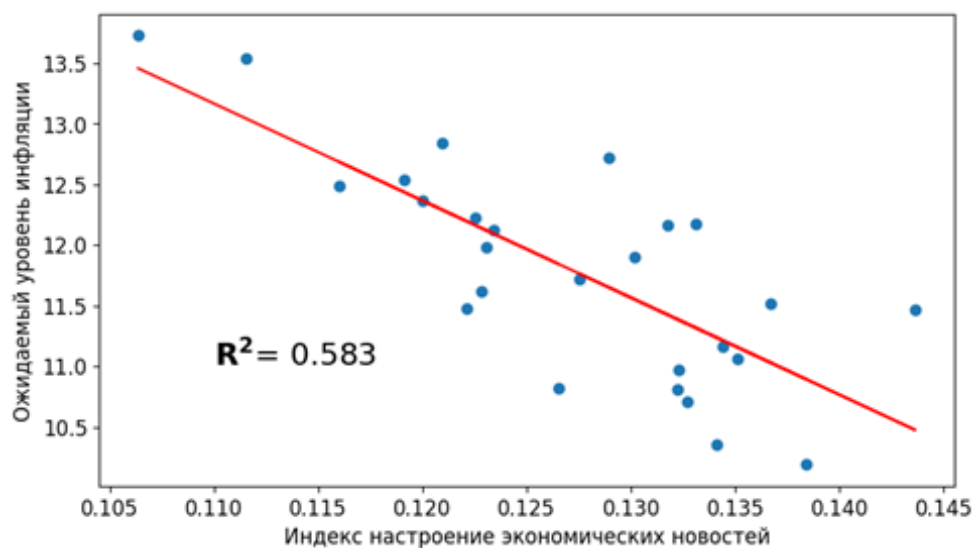


Рис. : Рисунок 1 - Диаграмма рассеяния для Индекса настроений экономических новостей и ожидаемого уровня инфляции

ИНЭН	R2	BIC	N
-5.793**	0,583	51,22	26

Рис. : Рисунок 2 - Результаты регрессии ожидаемого уровня инфляции на ИНЭН