

Применение нейронных сетей для оценки доверия населения к центральному банку

Матевосова Анастасия Михайловна (э401)

Научный руководитель: Картаев Филипп Сергеевич

10.10.2024

В прошлом году...

«Оценка влияния доверия к ЦБ на эффективность монетарной политики» —
название на этапе заявления темы курсовой

«Разработка метода построения индикатора доверия населения к
центральному банку на основе сентимент-анализа текстовых данных» —
название работы для конференции МФТИ в апреле 2024 г. (слайды далее)

«Моделирование роли доверия к центральному банку на основе сентимент-
анализа» — *название работы, представленной на конкурсе Банка России*

Постановка проблемы

- В настоящее время отсутствуют способы количественного измерения и оценки уровня доверия к ЦБ.
- В отличие от подходов измерения инфляционных ожиданий на основе опросов, социологические опросные методы не получили такого широкого распространения в контексте измерения уровня доверия к ЦБ.
- Несмотря на то, что различные виды ожиданий содержат в себе также и фактор доверия к ЦБ, его выделение из ожиданий представляется отдельной трудной задачей, требующей решения, в силу наличия многих факторов формирования ожиданий.

Актуальность

- Доверие к центральному банку - один из важнейших факторов, влияющих на эффективность монетарной политики и экономическую конъюнктуру в целом.
- Высокие значения показателя доверия к центральному банку помогают заякорить инфляционные ожидания населения на необходимом уровне, что особенно важно в условиях политики инфляционного таргетирования. Так как это позволяет центральному банку с меньшими потерями для выпуска и экономической активности достигать цели по инфляции.
- Доверие к центральному банку повышает уверенность населения и фирм в ценовой стабильности, тем самым формируя потребительские и инвестиционные настроения, способствующие экономическому росту.
- Для повышения эффективности монетарной политики центральному банку необходимо опираться на показатель доверия в процессе принятия решений, а также контролировать его динамику.

Новизна

- В настоящем исследовании предлагается авторская методология построения индикатора доверия к ЦБ, которая ставит перед собой цель построения первого в своём роде индикатора доверия к центральному банку.
- Данная методология опирается на современные методы сентимент-анализа текстовых данных социальных сетей.
- С помощью предложенного подхода впервые производится оценка и анализ динамики доверия к Банку России.

Цель и задачи (в работе прошлого года)

Цель: оценка влияния доверия к Банку России на эффективность проводимой им монетарной политики и формирование инфляционных ожиданий населения посредством построения индикатора доверия к ЦБ на основе сентимент-анализа данных сети Интернет.

Задачи:

1. Предложена авторская методология построения индикатора доверия к центральному банку на основе сентимент-анализа текстов комментариев сети Интернет.
2. С помощью предложенной автором методологии осуществлено построение индикатора уровня доверия к Банку России с использованием комментариев социальной сети ВКонтакте.
3. Проанализировано влияние доверия к ЦБ на инфляционные ожидания российского населения посредством построения ARMA-X-GARCH-X модели на основе рассчитанных индикаторов доверия к Банку России и Big Data индикатора инфляционных ожиданий.

Методы

- Парсинг постов и комментариев через VK API с помощью библиотеки request на python;
- Использование регулярных выражений для отбора тематических новостных постов;
- Применение методов sentiment-анализа, нейронных сетей для выявления эмоционального окраса комментариев.

Авторская методология построения индикатора доверия к ЦБ (Этапы 1-2)

Первый этап: формирование репрезентативной выборки СМИ, публикующих новости экономической тематики с открытой возможностью комментирования.

Второй этап:

2.1) На основе сформированного набора ключевых слов производится отбор новостных постов, соответствующих экономической тематике, в официальных группах рассматриваемых СМИ в социальной сети. Затем производится сбор комментариев к выделенной группе новостных постов экономической тематики.

2.2) На основе составленного набора ключевых выражений производится отбор новостных постов, содержащих упоминание центрального банка, в официальных группах рассматриваемых СМИ в социальной сети. После этого к выделенной группе новостных постов с упоминанием ЦБ собираются комментарии.

Этап 2: ключевые выражения

Набор ключевых выражений для отбора постов экономической тематики и постов с упоминанием ЦБ

Наличие экономической тематики	[«инфляци» «экономи» «кредит» «валют» «рубл» «цен»]
Упоминание центрального банка в тексте поста	[«банк россии» «банка россии» «банку россии» «банком россии» «банке россии» «цб» «центробанк» «центральный банк» «центрального банка» «центральному банку» «центральным банком» «центральном банке»]

Результаты этапов 1-2

Выборка из 10 СМИ: «Аргументы и факты» (aif), «Эксперт» (expert_ru), «Известия» (izvestia), «Коммерсант» (kommersant), «Московский Комсомолец» (mk_ru), РБК (rbc), РИА (ria), ТАСС (tassagency), Ведомости (vedomosti), Вести (vesti).



- К новостным постам экономической тематики было собрано 127681 комментариев пользователей
- К группе постов с упоминанием ЦБ – 57401 комментариев

Рисунок 1 Структура собранных комментариев к постам на экономические темы и к постам с упоминанием ЦБ за период с 2014 по 2023 год.

Авторская методология построения индикатора доверия к ЦБ (Этап 3)

Третий этап:

3.1) На основе методов текстовой обработки данных и машинного обучения рассчитываются показатели тональности собранных комментариев к постам на экономические темы. Их значения приводятся на отрезок от -1 до 1 и усредняются в рамках рассматриваемых временных единиц. Тем самым получается временной ряд $mood_{ec}$.

3.2) Аналогичным образом на основе методов текстовой обработки данных и машинного обучения производится расчёт временного ряда показателя тональности комментариев, собранных к постам с упоминанием ЦБ, $mood_{cb}$.

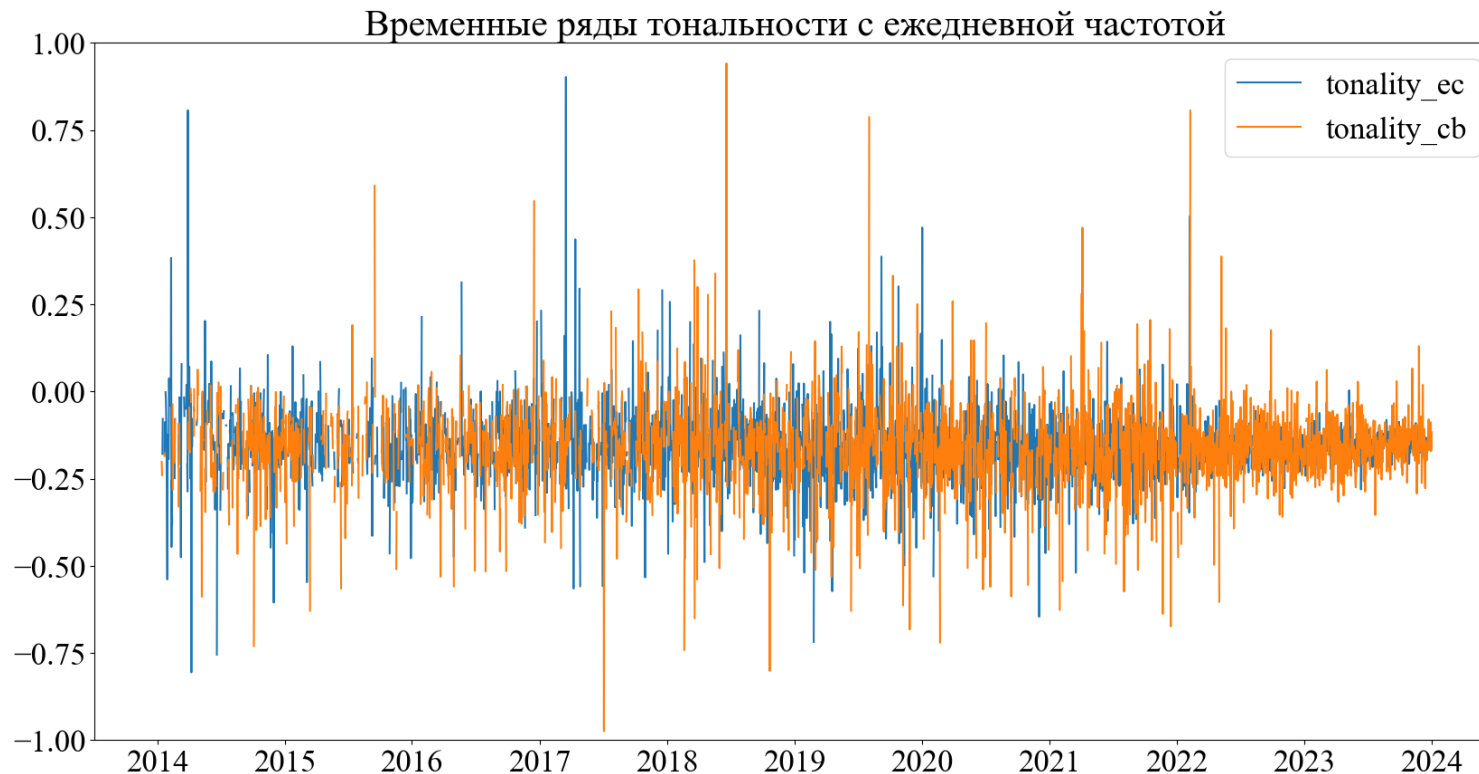
Этап 3: сентимент-анализ

Применяется модель из библиотеки Dostoevsky

```
text: ['Почему их не запретить? Всем же ясно они нацелены на выкачивание денег с людей']  
result: [{'negative': 0.607673168182373, 'neutral': 0.41490885615348816, 'skip': 0.15611489117145538, 'speech': 0.0216253343969  
5835, 'positive': 0.003955406602472067}]  
tonality: -0.603717761579901
```

Метод перевода набора вероятностей в единый показатель тональности:

$$tonality = P_{positive} - P_{negative}$$



Авторская методология построения индикатора доверия к ЦБ (Этап 4)

Четвёртый этап:

Производится расчёт искомого индикатора доверия к ЦБ, как разность между тональностью комментариев к постам с упоминанием ЦБ и тональностью комментариев к постам на экономические темы, приведённая к отрезку [0,1]:

$$trust = \frac{mood_{cb} - mood_{ec} + 2}{4}$$

Индикатор доверия к Банку России с ежедневной частотой



Индикатор доверия к Банку России с ежегодной частотой



Заключение по работе прошлого года

- Таким образом, в рамках рассматриваемой работы была предложена методика построения индикатора доверия к ЦБ.
- Эта методика была имплементирована при построении индикатора доверия к Банку России с различной частотой.
- Рассмотрена применимость индикатора доверия с различной частотой.
- На основе индикатора с ежегодной частотой проанализирована динамика доверия населения к Банку России.
- Индикаторы доверия и индикатор инфляционных ожиданий [Матевосова, 2023] использовались при построении ARIMA-X-GARCH-X модели

Планы

Применение нейронных сетей для оценки доверия к центральному банку

Цель: Моделирование доверия населения к центральному банку посредством построения индикатора на основе сентимент-анализа комментариев, оставляемых пользователями в социальной сети.

Планы (в основном относятся к этапу 3 методологии)

- В прошлом году для сентимент-анализа использовала библиотеку Dostoevsky
- Теперь планирую сконцентрироваться на выборе наилучшей модели для классификации по тональности конкретно моих данных
- Для этого:
 - 1) Разметить данные
 - 2) Выбрать метрику качества
 - 3) Рассмотреть разные подходы (нейросетевые, LLM, классический ML):
 - 4) Сравнить подходы с точки зрения качества классификации комментариев по тональности. (Посмотреть как это будет отражаться на итоговом индикаторе?)

Планируемые к рассмотрению подходы

- ❑ **Dostoevsky** (применена, но нужно будет померить качество)
- ❑ **Большие языковые модели (LLM): RuBERT, YGPT** (планирую применять готовые, а также делать **fine-tuning** и сравнивать качество)
- ❑ **Другие LLM:** пока думаю, какие ещё. Может быть модель Т-Банка и ещё какие-то LLM. Зависит от располагаемых мощностей.
- ❑ Возможно попробовать какие-то из базовых архитектур нейронных сетей
- ❑ Можно попробовать какие-то подходы на основе классического ML

Планируемая структура ВКР

Глава 1 Доверие в экономике и доверие к центральному банку

1. Роль доверия в экономике в целом (новое)
2. Роль доверия к ЦБ (переформулировать и дополнить)
3. Существующие способы измерения доверия к ЦБ (дополнить)

Глава 2 Возможности применения нейронных сетей и больших языковых моделей

1. Технологии на основе Big Data в экономике (также привести 2 свои статьи)
2. Развитие LLM и нейронных сетей и что такое LLM (из истории LLM)
3. Применение LLM в экономических исследованиях, в частности для кластеризации (почти сделан)
4. Задаче-ориентированный подход (про бэнчмарки, большое число существующих LLM и обоснование необходимости тщательного выбора модели для сентимент-анализа)

Глава 3 Измерение доверия

1. Общая методология
2. Структура собранных данных, разметка
3. Какие LLM буду тестировать, дообучать
4. Тестирование LLM, дообучение
5. Сравнение, выводы

Название диплома (варианты):

- Применение нейронных сетей для оценки доверия к центральному банку
- Применение нейронных сетей для построения индикатора доверия к центральному банку

Или переход к смежной теме:

- Применение нейронных сетей для оценки восприятия экономической ситуации населением

(Тогда беру один из промежуточных индикаторов, построенный по постам экономической тематики; но нужно менять набор ключевых выражений и тогда все заново парсить?)

Спасибо за внимание!