

Прогнозирование инфляции методами машинного обучения с использованием новостных индексов
(Сравнение предсказательной силы моделей машинного обучения с использованием инфляционных ожиданий ЦБ и новостного индекса)

Волгина Елизавета, э-401

Актуальность

1. Высокая инфляция по всему миру -> важно правильно прогнозировать
2. Сильные внешние шоки -> сложно учитывать (?) в стандартных моделях, нужны нелинейные модели (нейронка) + наиболее оперативным отражением всех шоков являются новости
3. Учет инфляционных ожиданий через опросы или изменения на финрынках имеет недостатки
->можно улучшить предсказательную силу прогноза инфляции, используя новостные индексы
4. Февраль 2023 – статья ЦБ про использование новостей для прогнозирования инфляционных ожиданий, строят новостные индексы (негативные новости)
-> **никто не успел пока использовать эти индексы для прогнозирования самой инфляции**
->* есть база данных новостей, но построить индексы на основе word2vec (у них tf-idf) (+про это пишут в статье из Annual review of economic, 2023)
5. Предыдущие работы (2020 и 2018) выявили, что методы ML превосходят по качеству прогнозов стандартные модели AR, RW -> хочу использовать новостные индексы именно в ML
6. Использование результатов моделей ML в обычных эконометрических моделях спорно -> новостные индексы строятся с ML -> хочу использовать новостные индексы именно в ML

- Объект: новостные индексы для целей прогнозирования инфляции
- Предмет: сравнение точности моделей инфляционного прогнозирования методами ML при использовании новостных индексов
- Цель: определить, улучшают ли новостные индексы точность моделей ML при прогнозировании инфляции
- Исследовательский вопрос: помогают ли новостные индексы точнее прогнозировать инфляцию, используя модели ML?

Задачи:

- 1) Анализ литературы про прогнозирование инфляции методами ML и построение новостных индексов (текстовый анализ)
 - 2) Построить новостные индексы на основе статьи (ЦБ 2023)
 - 3) Использовать новостные индексы для учета инфляционных ожиданий населения в моделях ML (выбрать на основе литературы (?нейронка) или построить всё что могу)
 - 4) Сравнить с аналогичными моделями, где в качестве инфляционных ожиданий используем данные ЦБ
 - 5) Интерпретировать результаты с помощью вектора Шепли
 - 6*) Построить новостные индексы по имеющемуся датасету новостей через word2vec
 - 7*) Изучить и использовать для сравнения моделей более подходящую метрику, чем RMSE
- *** Никто не сравнивал ML с более сложными моделями (BVAR)

Данные

- Новостные индексы (ЦБ 2023): Репозиторий GitHub с данными и кодом для парсинга новостей и построения новостных временных рядов (https://github.com/evstalina/russian_news_database)
- Регрессоры для инфляции (ЦБ 2020):
 - индекс физического объема ВВП (реальный ВВП), производительность труда (отношение реального ВВП к числу занятых), агрегат денежной массы M2, объем выданных кредитов в реальном выражении, норма безработицы, экспорт в реальном выражении, цена на нефть в долларах США, реальный располагаемый доход, процентная ставка денежного рынка, ИПЦ
 - январь 2002 г. по август 2018 г. (200 наблюдений) -> поварьировать с включением «неспокойных» последний лет (?)
 - ежемесячные процентные изменения

! В статье ряды очищают от сезонности, приводят к стационарному виду и стандартизируют, но есть работы, которые говорят, что для моделей ML не нужны стационарные ряды – попробовать и так, и так

Литература

Про новостные индексы:

- 1) Евстигнеева А., Карпов Д. (2023) Влияние негативных новостей на восприятие инфляции населением // Банк России: серия докладов об экономических исследованиях. № 111.
- 2) Andres-Escayola, E., Ghirelli, C., Molina, L. *et al.* (2023) Using Newspapers for Textual Indicators: Guidance Based on Spanish- and Portuguese-Speaking Countries. *Comput Econ.* <https://doi.org/10.1007/s10614-023-10433-z>
- 3) van Dijk, Dorinth and de Winter, Jasper, (February 17, 2023). Nowcasting GDP using tone-adjusted time varying news topics: Evidence from the financial press .De Nederlandsche Bank Working Paper No. 766
- 4) Dooruj Rambaccussing, Andrzej Kwiatkowski (2020), Forecasting with news sentiment: Evidence with UK newspapers, *International Journal of Forecasting* , 36 (4), pp. 1501-1516, ISSN 0169-2070, <https://doi.org/10.1016/j.ijforecast.2020.04.002>.
- 5) Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S.(2022). Making text count: Economic forecasting using newspaper text. *Journal of Applied Econometrics*, 37(5), 896–919. <https://doi.org/10.1002/jae.2907>
- 6) Pratap, Bhanu and Ranjan, Abhishek (June 1, 2022), Forecasting Food Inflation using News-based Sentiment Indicators . RBI-Occasional Papers-Vol. 42, No.2, 2021, Available at SSRN: <https://ssrn.com/abstract=4125337> or <http://dx.doi.org/10.2139/ssrn.4125337>

Про модели ML для прогнозирования:

- 1) Coulombe, P. G., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5), 920–964. <https://doi.org/10.1002/jae.2910>
- 2) Pavlov, E. (2020). Forecasting Inflation in Russia Using Neural Networks. *Russian Journal of Money and Finance*, 79(1), pp. 57–73.
- 3) Baybuza, I. (2018). Inflation Forecasting Using Machine Learning Methods. *Russian Journal of Money and Finance*, 77(4), pp. 42–59.
- 4) Ash, Elliott and Hansen, Stephen, Text Algorithms in Economics (September 1, 2023). *Annual Review of Economics*, Vol. 15, pp. 659-688, 2023
- 5) Masini, R. P., Medeiros, M. C., & Mendes, E. F. (2023). Machine learning advances for time series forecasting. *J Econ Surv*, 37, 76–111. <https://doi.org/10.1111/joes.12429>
- 6) Chakraborty, Chiranjit and Joseph, Andreas, Machine Learning at Central Banks (September 1, 2017). Bank of England Working Paper No. 674
- 7) Ahmed, Nesreen & Atiya, Amir & Gayar, Neamat & El-Shishiny, Hisham. (2010). An Empirical Comparison of Machine Learning Models for Time Series Forecasting. *Econometric Reviews*. 29. 594-621. 10.1080/07474938.2010.481556.
- 8) Kondratyev, Alexei, (April 11, 2018) Learning Curve Dynamics with Artificial Neural Networks
- 9) Petelin, G., Cenikj, G. and Eftimov, T., 2022. Towards understanding the importance of time-series features in automated algorithm performance prediction. *Expert Systems with Applications*, 119023.