



Предсказания инфляции при помощи ML: достижения последних лет

Олег Шибанов, PhD

Мнения в данной презентации не отражают точку зрения организаций, в которых работает автор

Тезисы про инфляцию и ML

1. Модели ML позволяли улучшать предсказания инфляции
 - I. Иногда лучше выглядят обычные «случайные леса», иногда – модели типа LASSO с ограничениями на коэффициенты
 - II. Почти для всех крупных экономик нелинейные модели улучшают результаты
 - III. Проблема: данные обычно заканчиваются в начале 2022
2. Предварительные результаты для России также положительные
 - I. Прогнозы машинного обучения и нейронных сетей лучше, чем разного вида авторегрессии
 - II. Аналитики также уступают ML
 - III. Минусы: результаты всё равно не идеальны

1. Результаты по нескольким странам

США (2023)

- [2023 - Deep learning models for inflation forecasting](#)
- Месячные данные 1978-2019, 134 макросерии
- Variational Autoencoders and Convolutional LSTM Networks (VAE-ConvLSTM) сравнивают с 24 другими моделями, включая LASSO, Random forest, VECM, Random walk и т.д.
- Результаты этой «глубокой модели» лучше

Модель	MSE	1	2	3	6	12
ConvLSTM	Медиана	0.17	0.42	0.45	0.46	0.46
SARIMA	Медиана	0.77	1.02	1.13	1.24	1.22



Еврозона (2023)

- [2023 - Forecasting euro area inflation with machine learning models](#)
- Quantile Regression Forest, данные за 2002-2022
- Сравнение с Eurosystem inflation forecasts (BMPE)
- Хотя не превосходит, но хорошо дополняет

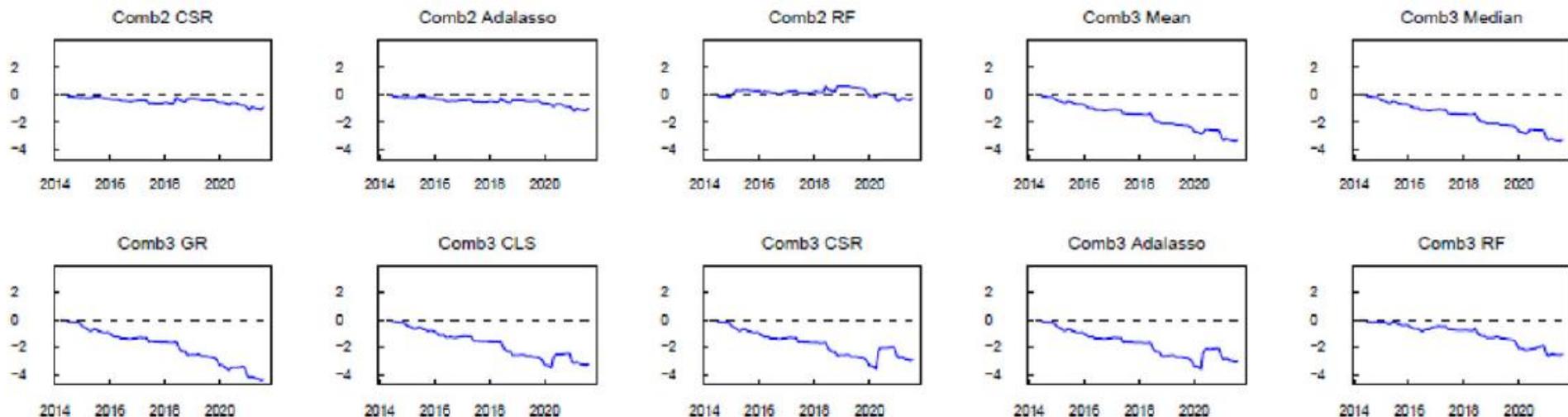
Figure 6: Headline Inflation, density forecasts of QRF and BMPE, h=6



Note: Black solid line: year-on-year growth rate of HICP (headline inflation); Red area: 16th to 84th quantiles of the QRF density forecasts for the horizon of six months ahead, year on year growth rate of HICP; Green line with circles: BMPE projections for the horizon of six months ahead, year on year growth rate of HICP.

Бразилия (2023)

- [2023 - Machine learning methods for inflation forecasting in Brazil: New contenders versus classical models](#)
- 167 макрофинансовых переменных, 2004-2021 (август), месячные данные, 50 моделей
- Сокращение ошибок по сравнению с ARMA в два раза
- Cumulative squared prediction error – чем ниже, тем лучше альтернативная модель
- Видно, что во времени ситуация улучшается



Индия (2024)

- [2024 - Inflation Forecasting in India: Are Machine Learning Techniques Useful?](#)
- Данные 1996Q2-2022Q1; сравнивают особенно с ARIMA и SARIMA, потому что это наиболее эффективные модели
- Результат: опять нейронки лучше, но качество не идеальное

RMSE of Rolling Sample Forecasts of Inflation (y-o-y) vis-à-vis Actual Inflation

Model Type		Model	One-quarter ahead (2019Q1-2021Q4)			Four-quarters ahead (2019Q4-2022Q1)		
			Full period	Pre- COVID	Post- COVID	Full period	Pre- COVID	Post- COVID
Univariate	Traditional	Random Walk	0.92	1.07	0.76	2.53	3.04	1.44
		ARIMA	0.80	0.71	0.87	1.25	1.43	0.91
		SARIMA	0.79	0.74	0.84	1.57	1.89	0.91
		LR	0.66	0.61	0.70	0.93	0.87	1.00
	ML	ANN	0.58	0.61	0.55	0.93	0.97	0.87
		RNN-LSTM	0.76	0.72	0.80	1.77	1.90	1.55

Китай (2024)

- [2024 - China's inflation forecasting in a data-rich environment: based on machine learning algorithms](#)
- 1997-2021, 96 макропеременных
- Особенно на горизонт 12 месяцев выигрывает ML

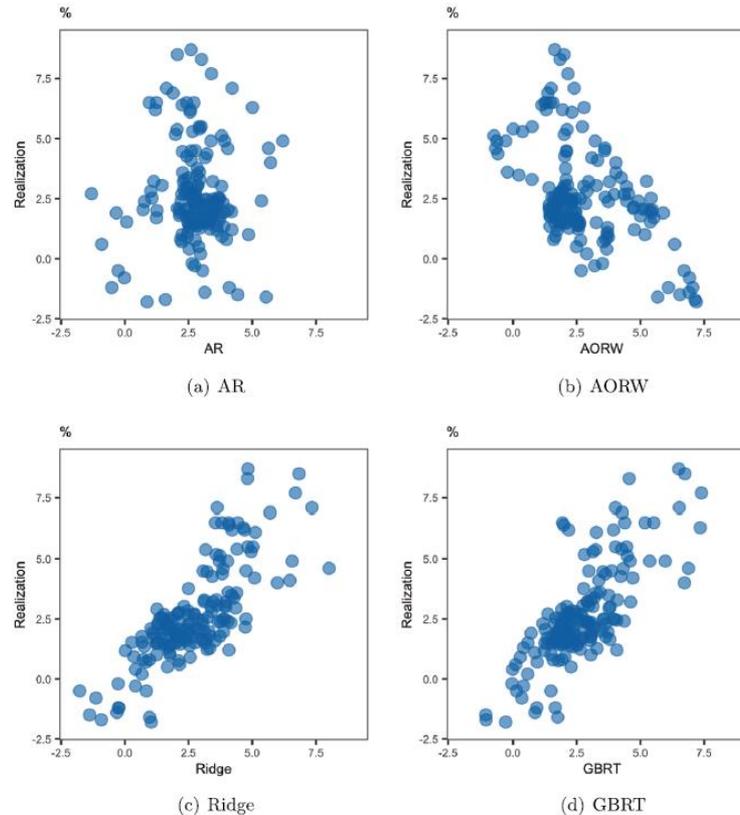


Figure 4. Scatter plots - forecasts and realizations of CPI inflation: $h=12$. These scatter plots show the forecasts (on the x axis) and realizations (on the y axis) for $h = 12$.

2. Результаты по России

Россия

- В ML-моделях лаги и курс важны для предсказания инфляции на короткие горизонты ([Павлов](#) (Деньги и кредит,2020), [Ермак](#) (ВКР,2023))
- Павлов: 2002-2018, 10 макропеременных
- Результат: нейронная сетка выглядит достаточно успешной на горизонтах 2-24 месяца, авторегрессия хуже

Table 1. Forecasting errors

Model	Forecast month				
	1	2	6	12	24
	RMSE				
AR (1)	0.2998	0.4500	0.4278	0.4286	0.4369
NN	0.3460	0.3880	0.3980	0.3990	0.4010
SVM	0.3100	0.3910	0.3980	0.3940	0.4070
Ridge	0.3180	0.3910	0.4180	0.4310	0.4230

Note: the best models for each horizon are highlighted in green, and the worst models – in yellow.

Россия (моё упражнение)

- Очень простые модели Random forest, Gradient boosting и линейная регрессия с 6 макропеременными (курс, зарплаты и т.п.)

$$\pi_{t+12} = a + b_1\pi_t + b_2FX_t + b_3Pension_{t-3} + b_4Wage_{t-3} + b_5M2_{t-2} + b_6Cash_{t-2} + \varepsilon_t$$

- Они не полностью ловят внутригодовую динамику, но удивительно хороши по итогам года
- В таблице: годовая инфляция; и абсолютные ошибки предсказания по данным 12-15 месяцев из прошлого по трём моделям – даже 2020-е отлично предсказаны

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
RF	0,50%	3,32%	1,17%	0,45%	0,75%	1,00%	1,62%	0,07%	0,07%	3,46%
GB	0,41%	3,29%	1,03%	0,06%	0,66%	0,90%	1,21%	0,08%	0,08%	3,31%
LR	0,37%	1,80%	2,49%	1,45%	1,22%	3,55%	2,48%	0,46%	0,44%	4,43%
Inflation	10,92%	9,00%	11,87%	13,28%	8,80%	8,78%	6,10%	6,57%	6,47%	11,35%
	2015	2016	2017	2018	2019	2020	2021	2022	2023	
RF	1,01%	1,71%	0,12%	0,61%	1,90%	0,83%	0,56%	0,33%	0,48%	
GB	1,92%	1,24%	1,05%	0,73%	2,01%	1,16%	1,05%	1,45%	0,25%	
LR	3,81%	2,24%	0,21%	0,88%	3,88%	0,08%	1,07%	5,67%	1,56%	
Inflation	12,91%	5,39%	2,51%	4,26%	3,04%	4,91%	8,39%	11,94%	7,42%	

Выводы

1. Кажется, машинное обучение достаточно хорошо работает в прогнозировании инфляции во многих крупных экономиках
2. Shapley decomposition позволяет выявить наиболее полезные в ML переменные
3. Качество модели по отношению к линейным довольно стабильно, даже в волатильные периоды
4. Надо больше результатов по 2021-23
5. В России качество «годовой» модели достаточно высокое