



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ФАКУЛЬТЕТ КОМПЬЮТЕРНЫХ НАУК ДЕПАРТАМЕНТ ПРОГРАММНОЙ ИНЖЕНЕРИИ

ПОСТРОЕНИЕ ПРЕДСКАЗАТЕЛЬНЫХ МОДЕЛЕЙ ФИНАНСОВЫХ РЫНКОВ

Подготовили студенты 2 курса образовательной
программы «Программная инженерия»:

Д. Озмаден,
А.О. Калугин



АКТУАЛЬНОСТЬ РАБОТЫ



От точности прогнозирования цен финансовых активов
зависит прибыль, получаемая трейдерами



ПОСТАНОВКА ЗАДАЧИ

Разработать метод прогнозирования цен на фондовом рынке, которые будут наиболее эффективны при алгоритмической торговле.

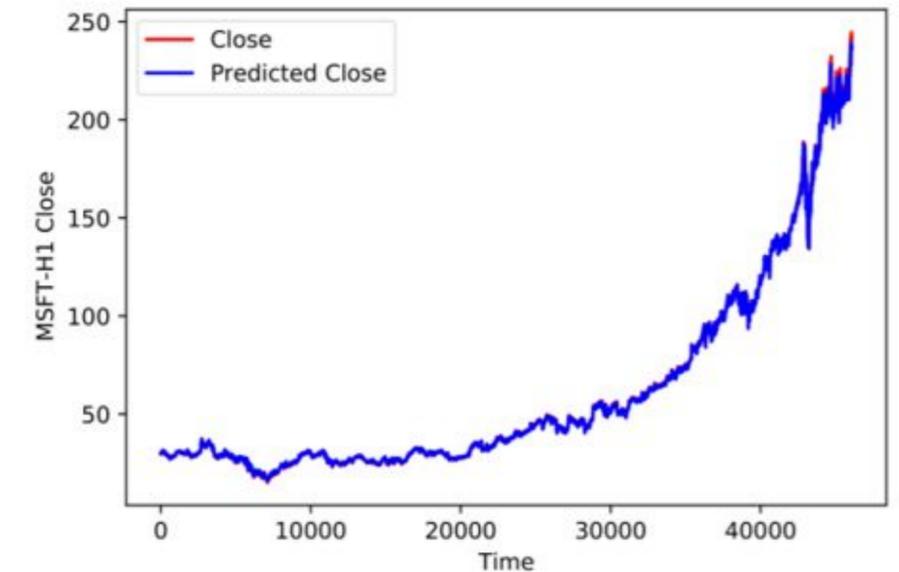




РЕАЛИЗОВАННЫЕ СТРАТЕГИИ



Предсказания ансамбля LSTM нейронных сетей



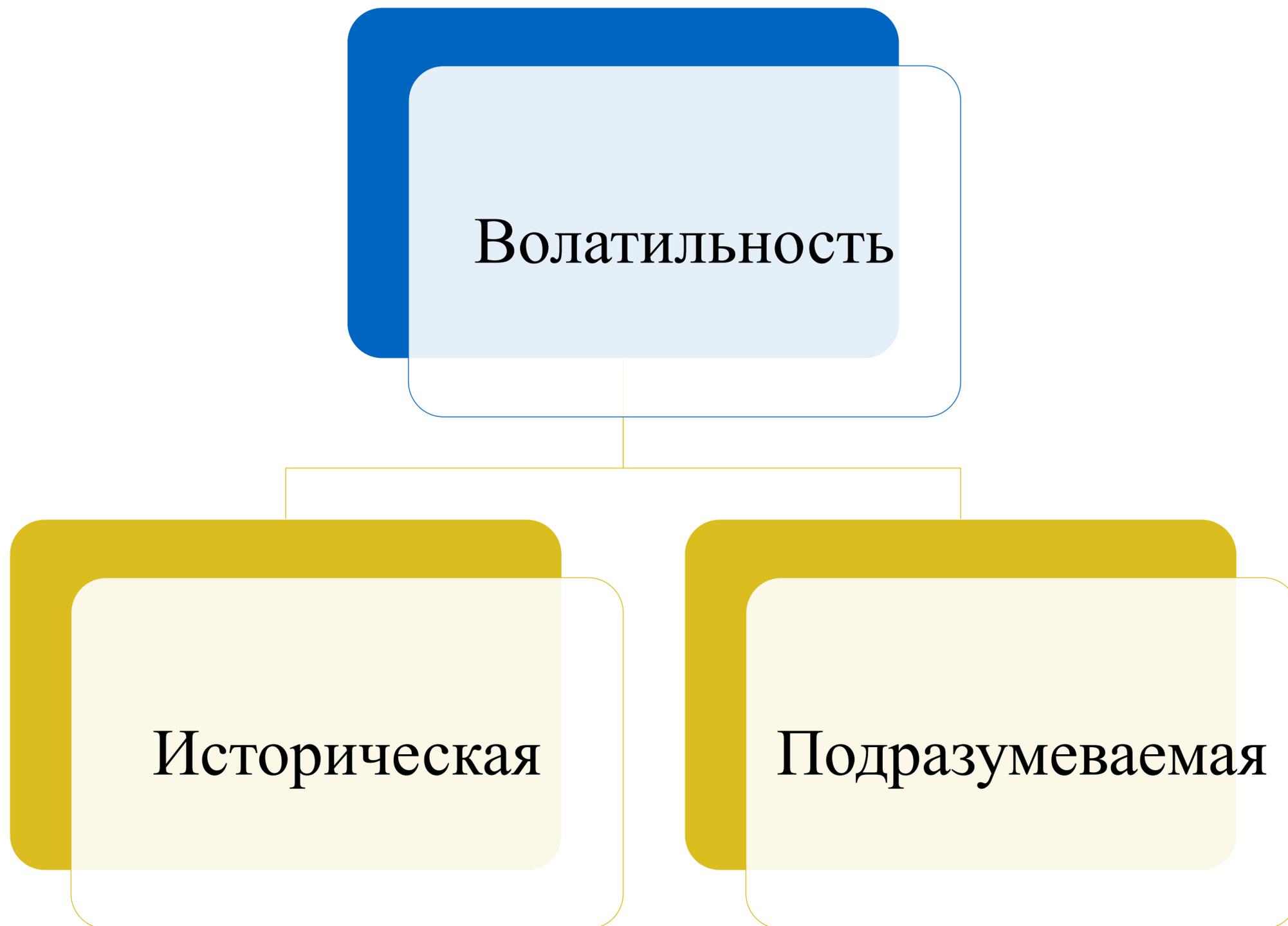
Анализ волатильности



Линии поддержки и сопротивления



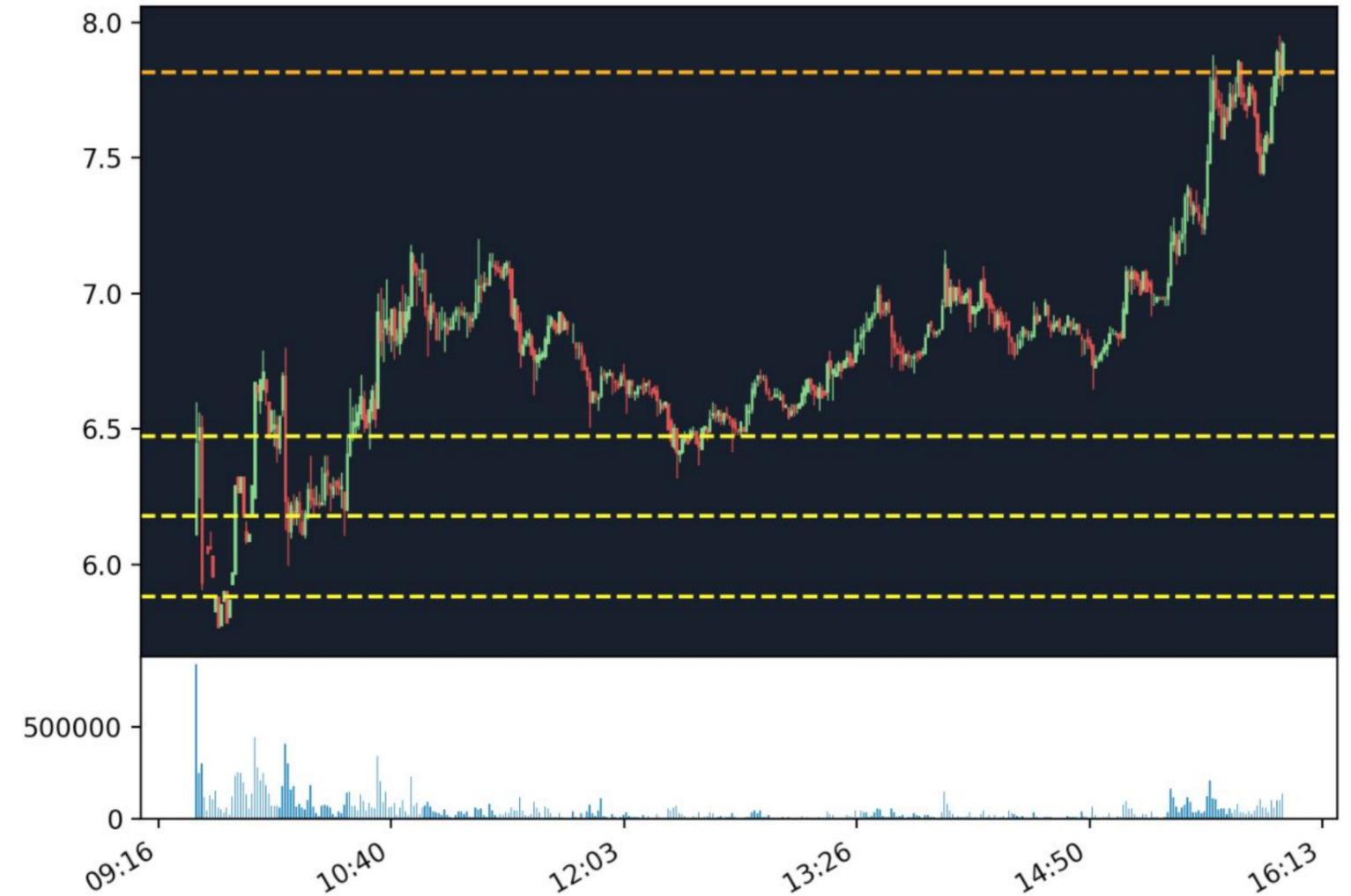
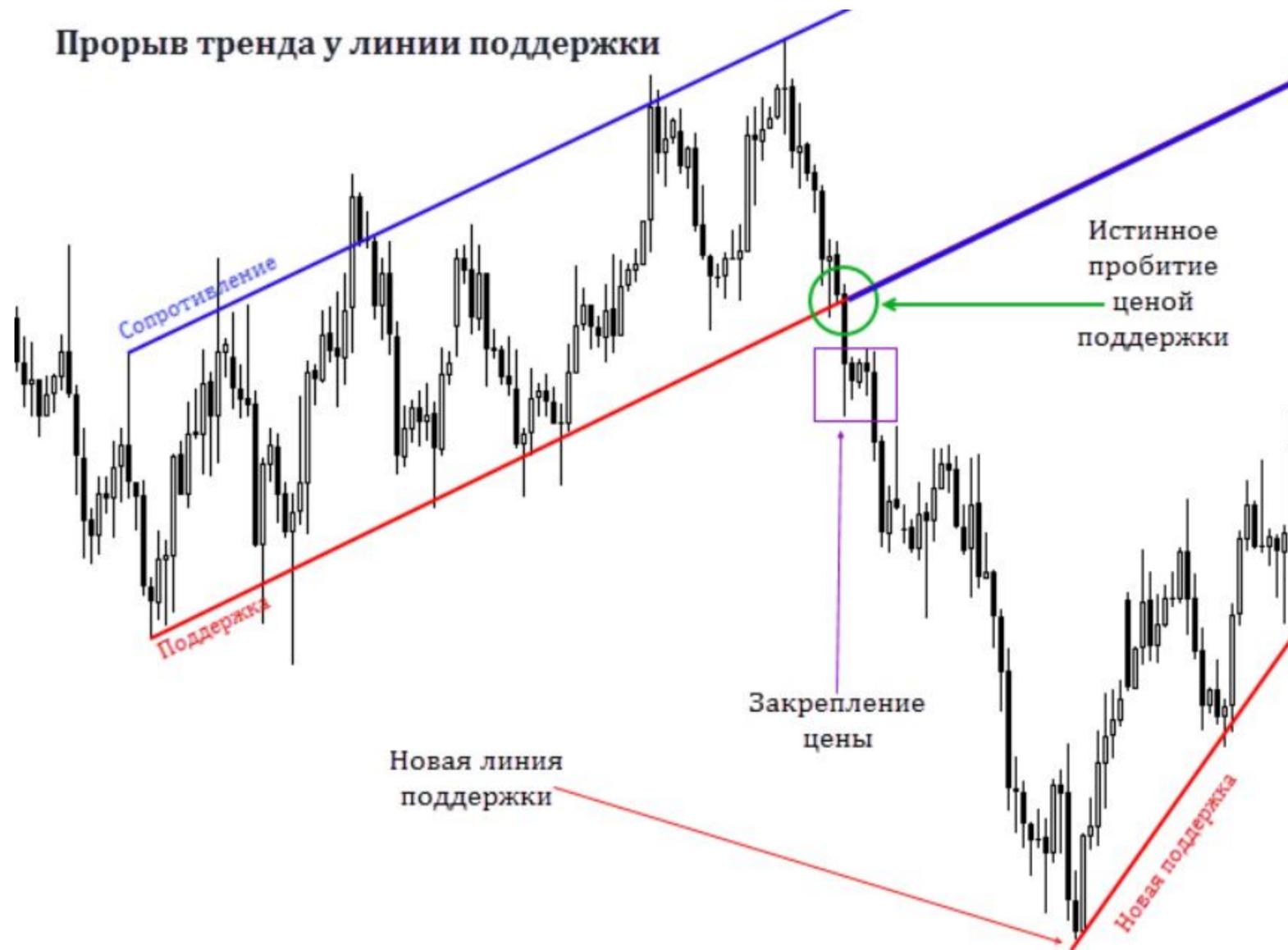
ДВА ТИПА ВОЛАТИЛЬНОСТИ





ЛИНИИ ПОДДЕРЖКИ И СОПРОТИВЛЕНИЯ

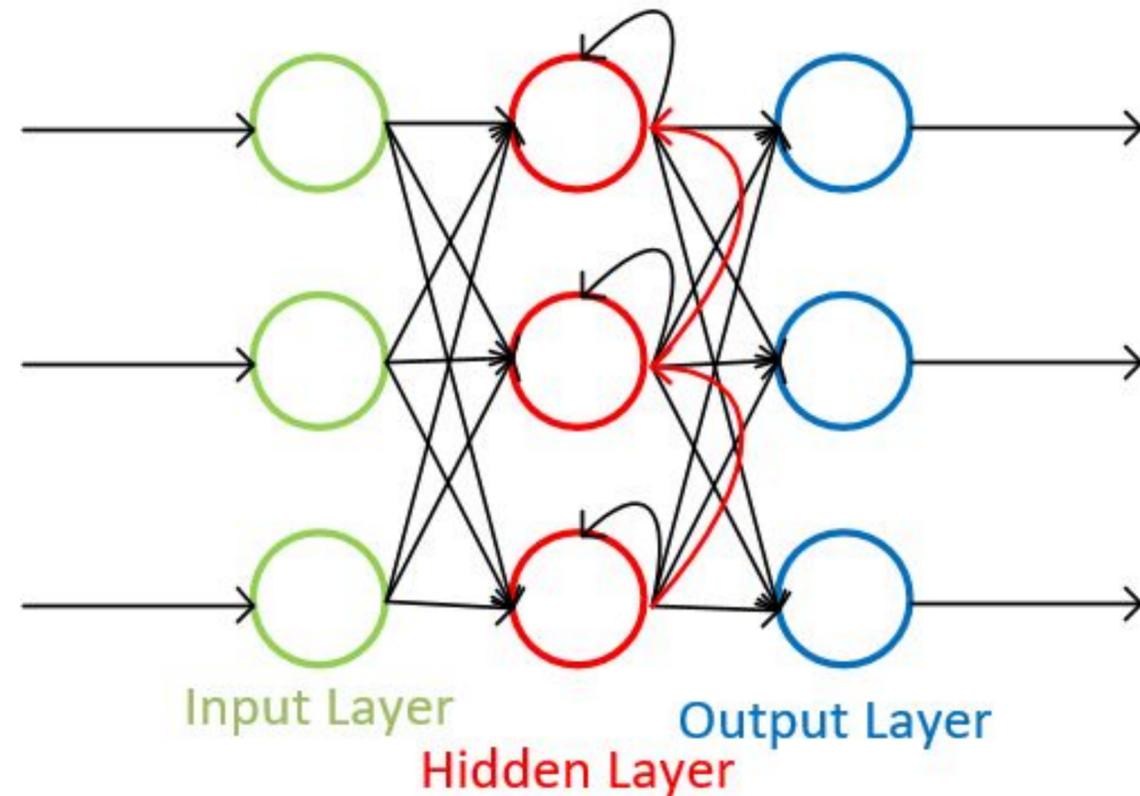
Прорыв тренда у линии поддержки



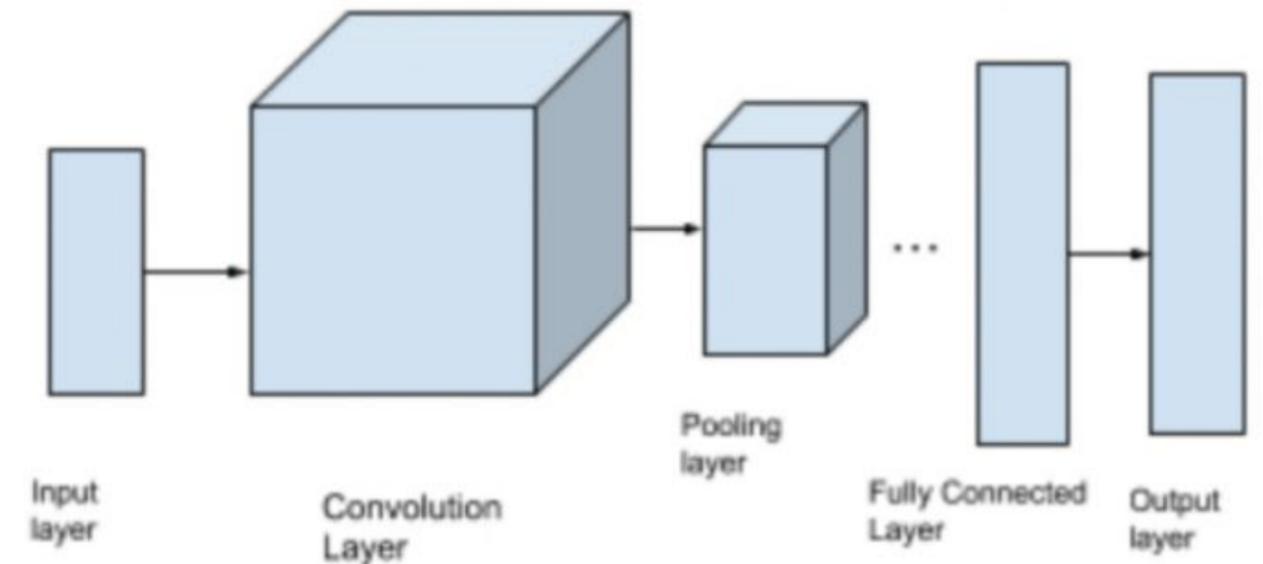


НЕЙРОННЫЕ СЕТИ

Recurrent Neural Network



Convolutional Neural Network



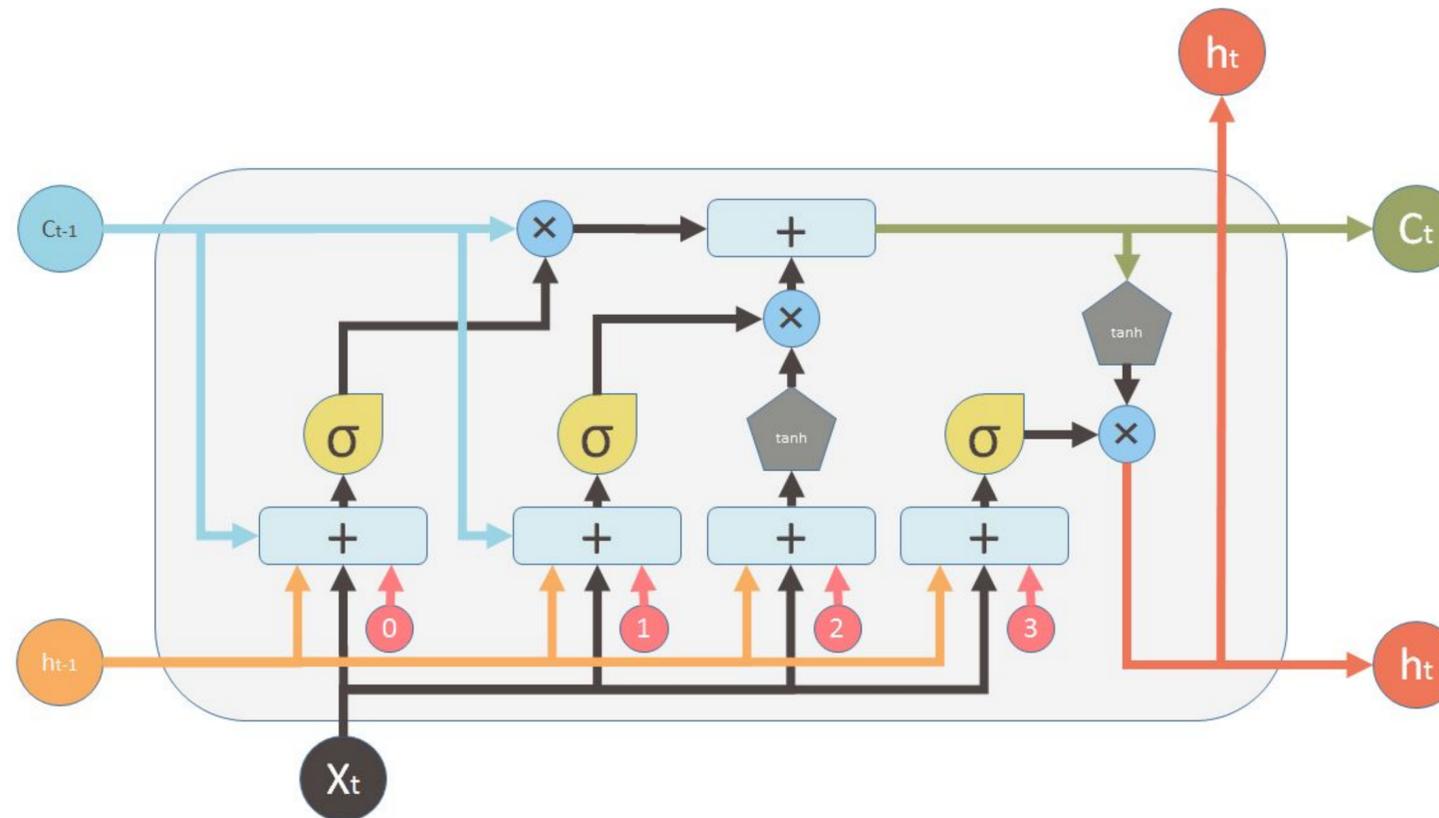
Свёрточная нейронная
сеть (CNN)

Рекуррентная нейронная сеть (RNN)

Нейронная сеть с долгой краткосрочной памятью (LSTM)



НЕЙРОННАЯ СЕТЬ С ДОЛГОЙ КРАТКОСРОЧНОЙ ПАМЯТЬЮ (LSTM)



Inputs:	outputs:	Nonlinearities:	Vector operations:
X_t Input vector	C_t Memory from current block	Sigmoid	Element-wise multiplication
C_{t-1} Memory from previous block	h_t Output of current block	Hyperbolic tangent	Element-wise Summation / Concatenation
h_{t-1} Output of previous block		Bias: 0	



СРЕДНЕКВАДРАТИЧНАЯ ОШИБКА

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}},$$

где \hat{y}_i – это реальная цена, y_i – это прогнозируемое значение, а n – количество всех наблюдений.



СРЕДНЯЯ АБСОЛЮТНАЯ ОШИБКА

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n},$$

где \hat{y}_i – это реальная цена, y_i – это прогнозируемое значение, а n – количество всех наблюдений.



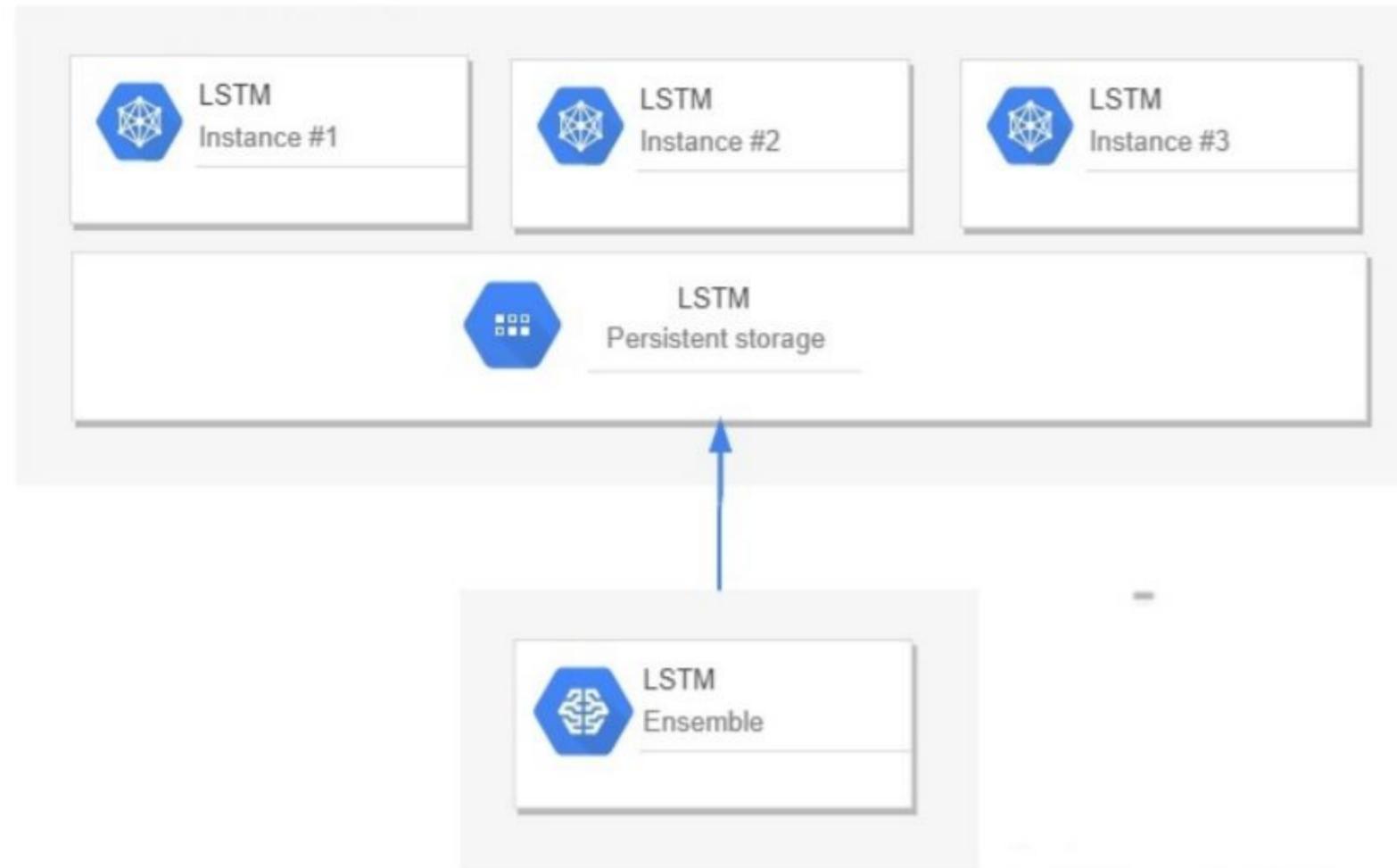
СРАВНЕНИЕ СУЩЕСТВУЮЩИХ РЕШЕНИЙ [1]

	RNN	LSTM	CNN
RMSE	43.957	41.003	42.967
MAE	29.916	28.285	30.138

1. Wenjie Lu, Jiazheng Li, Yifan Li, Aijun Sun, Jingyang Wang, "A CNN-LSTM-Based Model to Forecast Stock Prices", *Complexity*, vol. 2020, Article ID 6622927, 10 pages, 2020. <https://doi.org/10.1155/2020/6622927>



ПРЕДЛАГАЕМЫЙ МЕТОД



Ансамбль LSTM нейронных сетей (LSTM Ensemble)



ОЦЕНКА ПРЕДЛАГАЕМОГО МЕТОДА НА ДАННЫХ ПУБЛИКАЦИИ [1]

	RNN	LSTM	CNN	LSTM Ensemble
RMSE	43.957	41.003	42.967	31.01
MAE	29.916	28.285	30.138	22.463



ДАТАСЕТ ЯПОНСКИХ СВЕЧЕЙ [3]

Data Explorer

1.67 MB

- AAPL-D.csv
- AMZN-D.csv
- BAC-D.csv
- DIS-D.csv
- EA-D.csv
- FB-D.csv
- GOOGL-D.csv
- MSFT-D.csv
- NFLX-D.csv
- NKE-D.csv
- NVDA-D.csv
- TMO-D.csv
- TXN-D.csv
- UNP-D.csv
- VZ-D.csv
- WMT-D.csv
- XOM-D.csv

< AAPL-D.csv (287.92 KB)

Detail Compact Column 9 of 9 columns

asset_id	interval	timestamp	# open	# high	#
1 unique value	1 unique value				
AAPL	D	2000-02-02	25.1875	25.53125	24
AAPL	D	2000-02-03	25.078129999999998	26.0625	25
AAPL	D	2000-02-04	25.984379999999998	27.5	25
AAPL	D	2000-02-07	27.0	28.5625	26
AAPL	D	2000-02-08	28.5	29.03125	27
AAPL	D	2000-02-09	28.53125	29.28125	28
AAPL	D	2000-02-10	28.21875	28.46875	27

3. Ozmaden, D. OHLC Financial data, Version 1. Available from: <https://www.kaggle.com/dozmaden/ohlcfinancial-data> (Дата обращения: 10.03.2021, режим доступа: свободный)



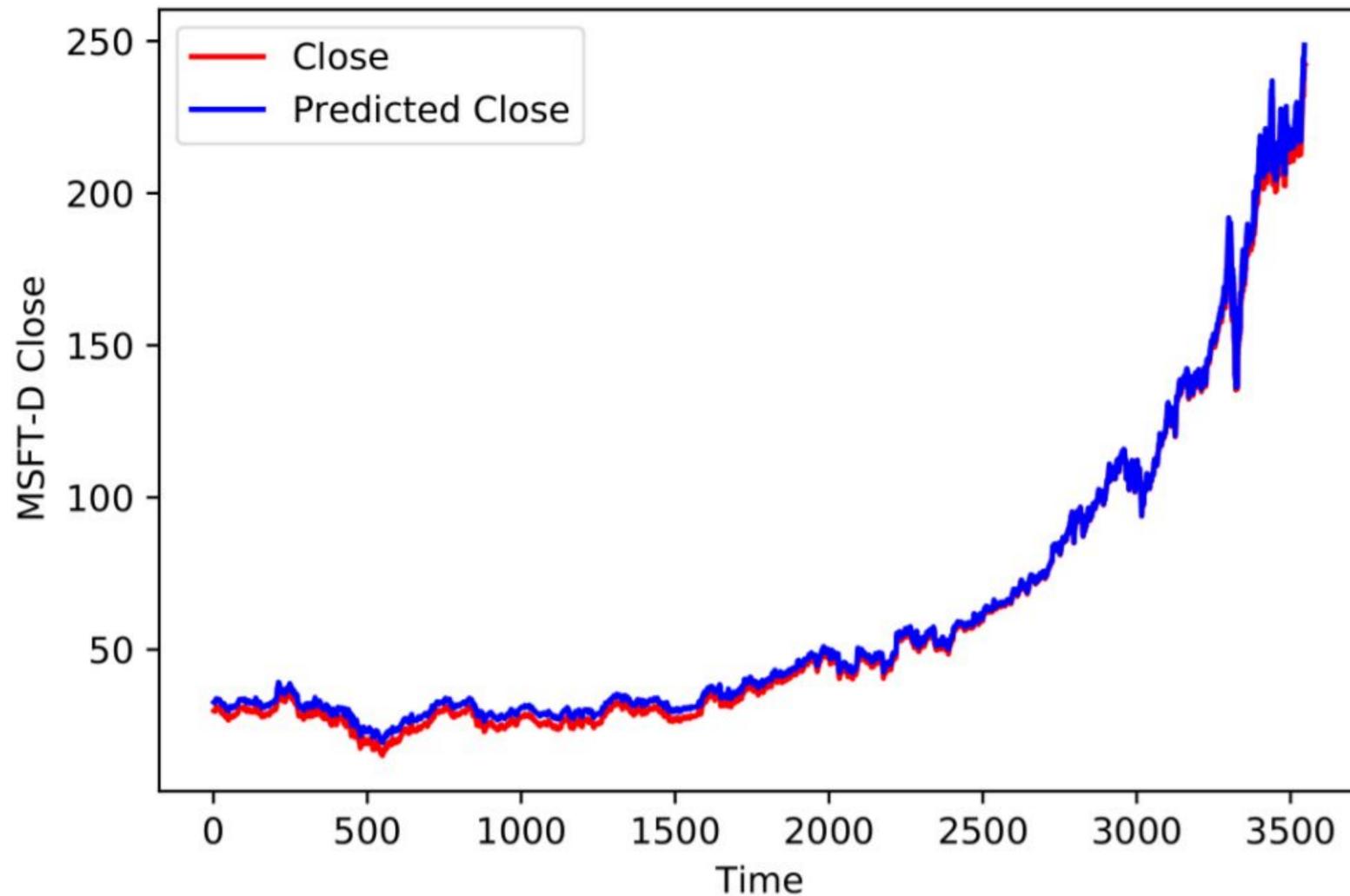
СРЕДНЯЯ ПРОЦЕНТНАЯ ОШИБКА

$$\text{MPE} = \frac{100\%}{n} \sum_{t=1}^n \frac{\hat{y}_t - y_t}{\hat{y}_t},$$

где \hat{y}_t – это реальная цена, y_t – это прогнозируемое значение, а n – количество всех наблюдений.



РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА НА ЕЖЕДНЕВНЫХ ДАННЫХ



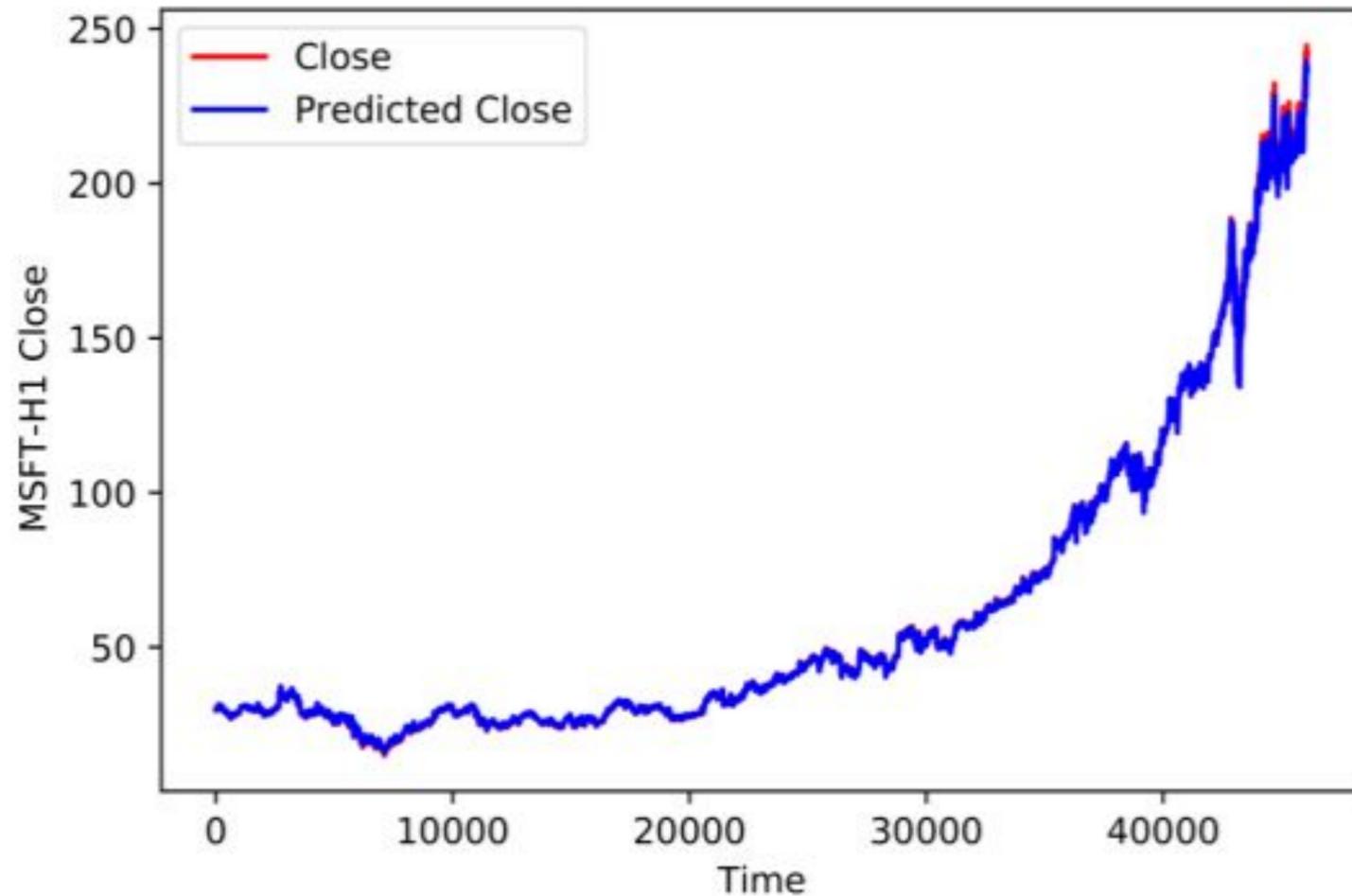
Ежедневный график исходных и предсказанных цен акций Microsoft.

	MSFT	DIS	FB
RMSE	2.91	1.92	4.10
MPE	-6.473	-2.478	-1.3048224
Max. PE	11.9	12.474	17.348574
Min. PE	-35.71	-24.307	-21.688463
Correct direction	0.533	0.583	0.514
MPE (of correct predictions)	-5.064	-1.275	-0.5703
Max. PE (of correct predictions)	11.891	12.47	17.34857
Min. PE (of correct predictions)	-25.701	-17.91	-21.69
Wrong direction	0.467	0.417	0.486
MPE (of wrong predictions)	7.995	3.773	2.093
Max. PE (of wrong predictions)	35.711	24.307	21.61
Min. PE (of wrong predictions)	0.158	0.187	-11.32

Метрики ежедневных предсказаний
цен акций компаний:
Microsoft, Disney, Facebook



РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА НА ПОЧАСОВЫХ ДАННЫХ



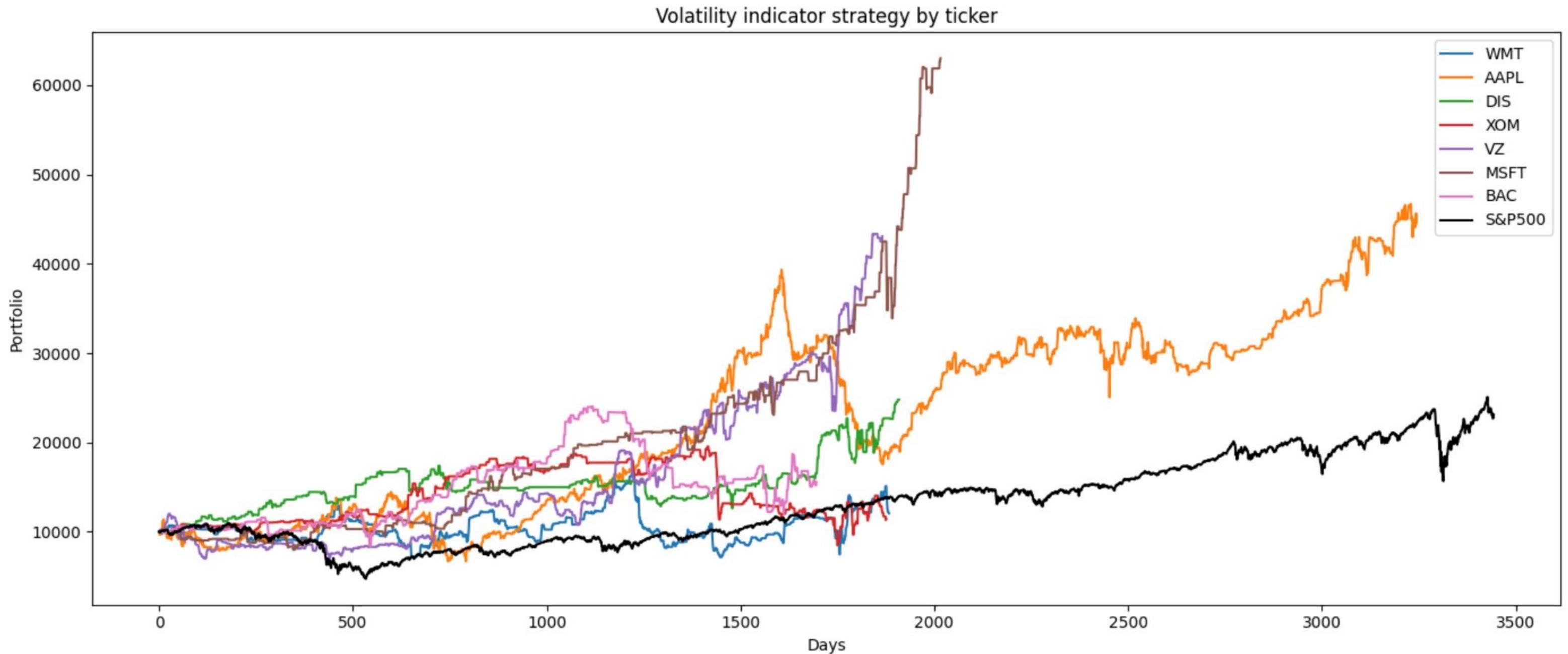
	MSFT	DIS	FB
RMSE	0.71	0.54	1.16
MPE	-0.174	0.154	-0.0727
Max. PE	10.443	11.016	18.809
Min. PE	-12.516	-13.377	-22.817
Correct direction	0.536	0.571	0.509
MPE (of correct predictions)	-0.1754	-0.0327	-0.0476
Max. PE (of correct predictions)	9.18	8.66	18.808
Min. PE (of correct predictions)	-11.34	-13.377	-22.818
Wrong direction	0.464	0.429	0.491
MPE (of wrong predictions)	0.172	-0.3305	0.0969197
Max. PE (of wrong predictions)	12.51	8.1851	18.755
Min. PE (of wrong predictions)	-10.44	-11.0167	-12.762719

Метрики почасовых предсказаний цен
акций компаний:
Microsoft, Disney, Facebook

Почасовой график исходных и
предсказанных цен акций Microsoft



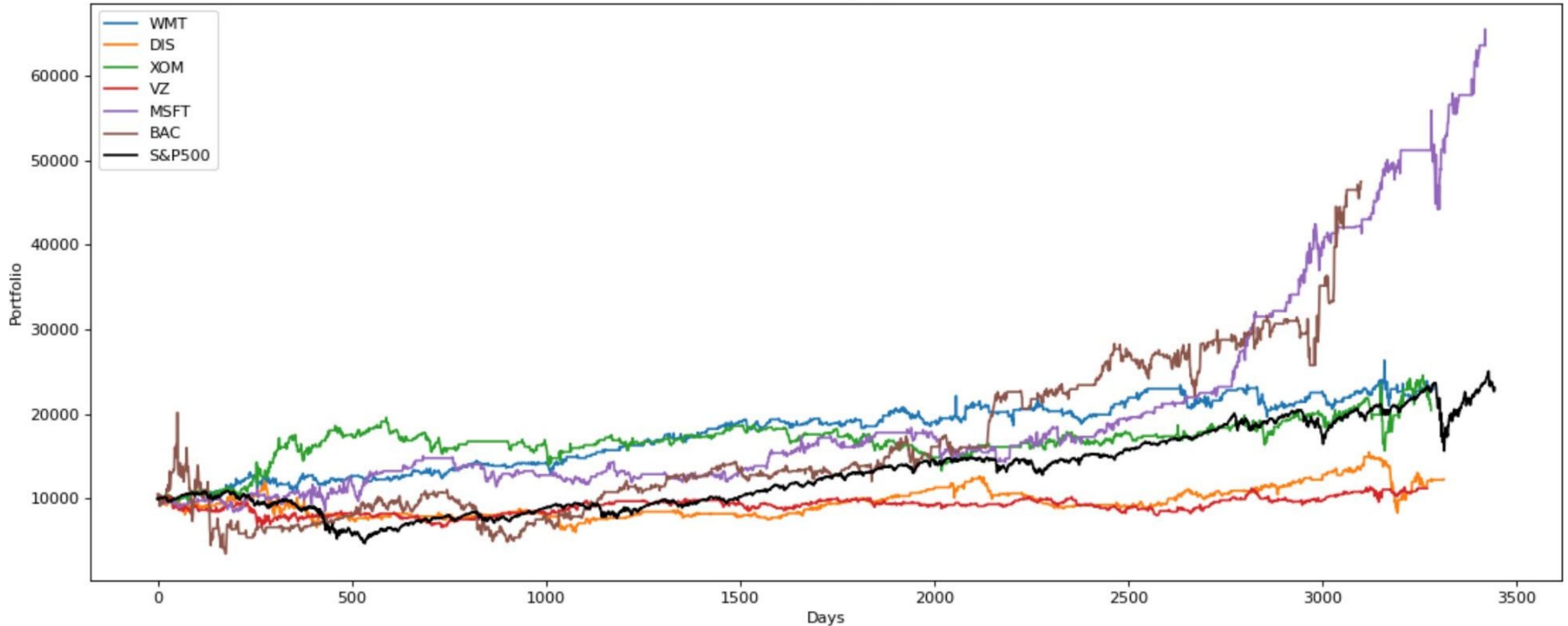
РЕЗУЛЬТАТЫ АЛГОРИТМИЧЕСКОЙ ТОРГОВЛИ НА ОСНОВЕ ИНДИКАТОРОВ ВОЛАТИЛЬНОСТИ





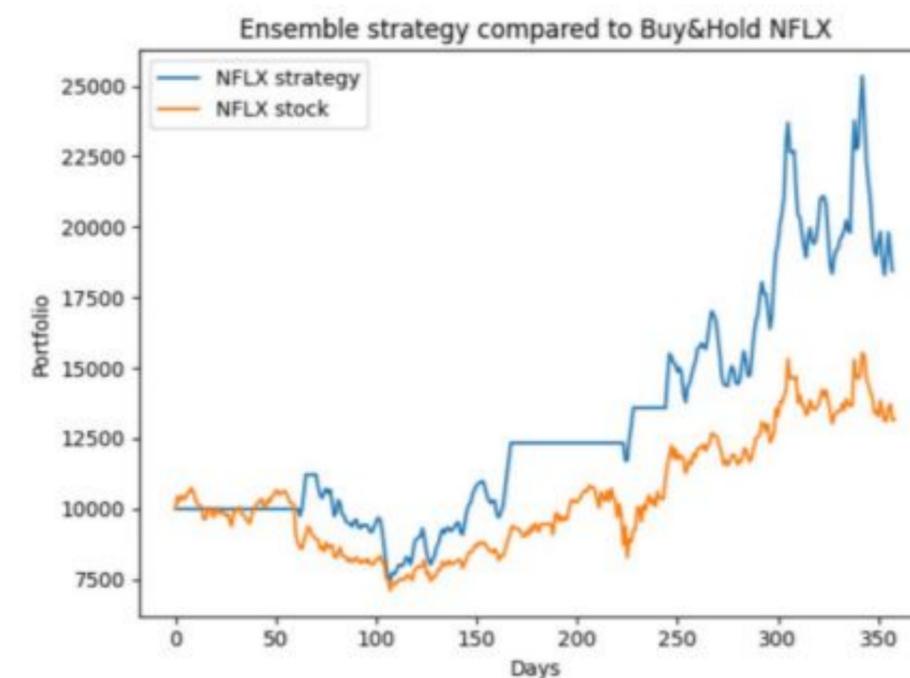
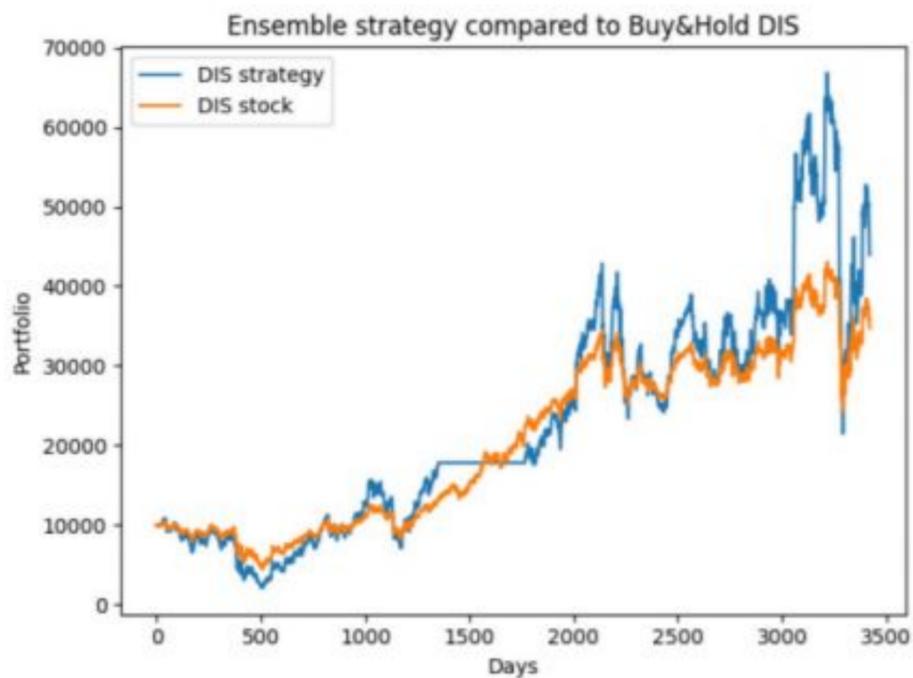
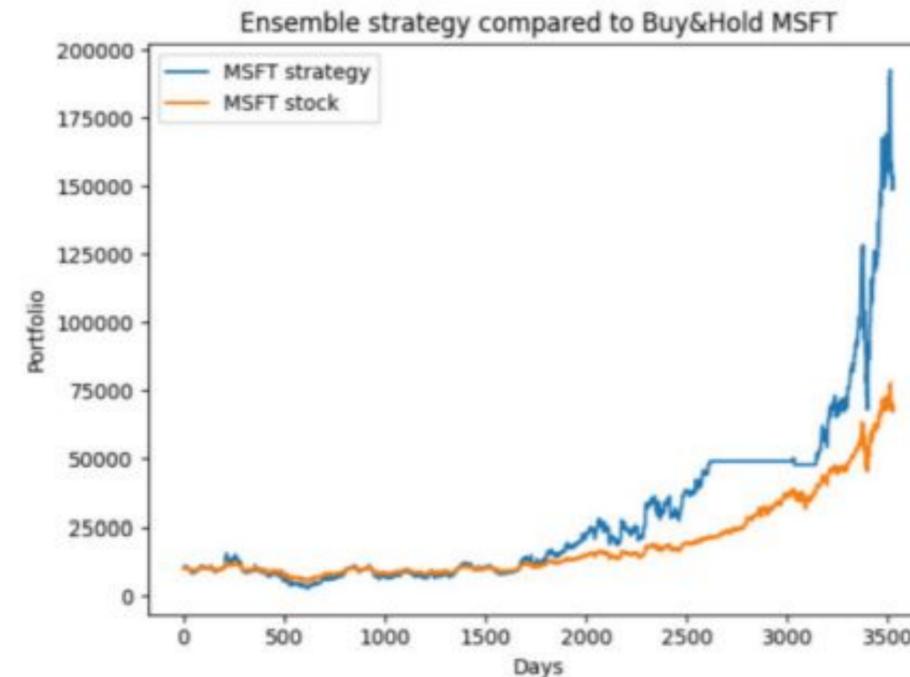
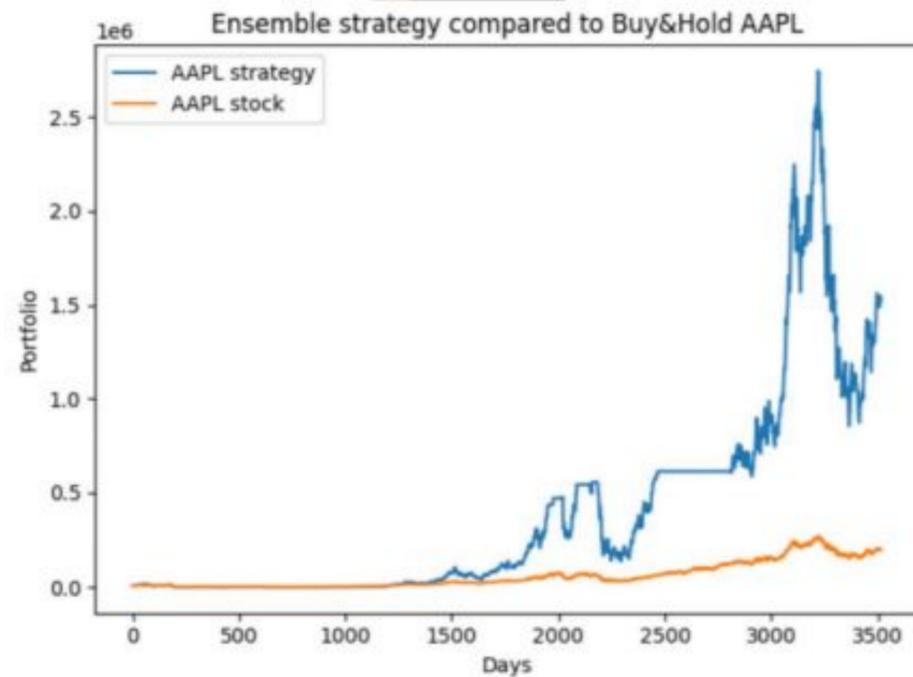
РЕЗУЛЬТАТЫ АЛГОРИТМИЧЕСКОЙ ТОРГОВЛИ НА ОСНОВЕ МЕТОДА К-СРЕДНИХ

K-means strategy by ticker





РЕЗУЛЬТАТЫ АЛГОРИТМИЧЕСКОЙ ТОРГОВЛИ НА ОСНОВЕ ПРЕДСКАЗАНИЙ АНСАМБЛЯ LSTM



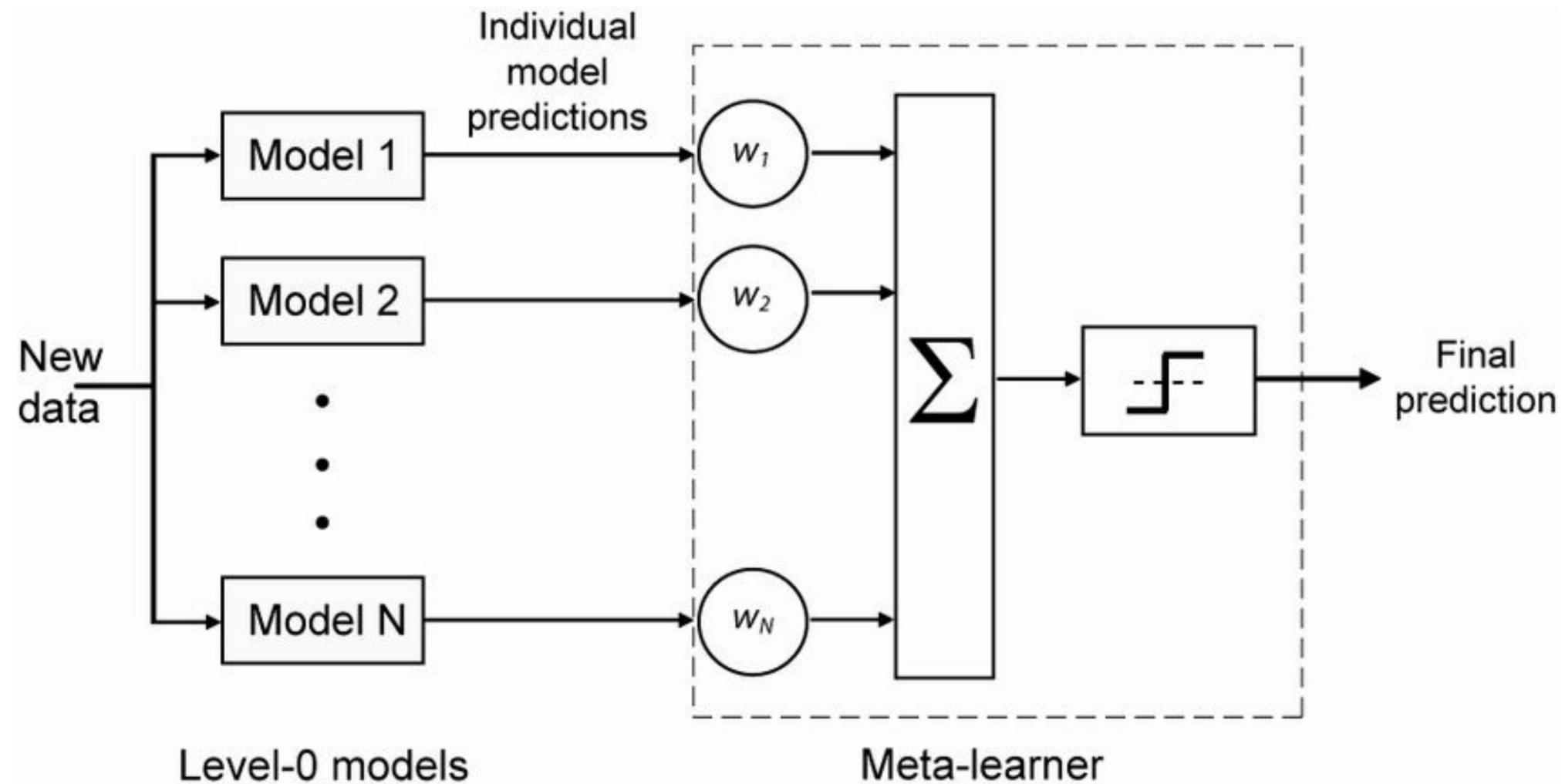


ВЫВОДЫ

- Средний прирост за сделку больше снижения
- Модель предсказывает правильно будущий тренд более чем в 50% случаев
- Стратегии торговли, заложенная в моделях, наиболее эффективны на минутных интервалах, где необходима быстрота принятия решений;
- Возможна параллельная торговля на множестве активов



ПЕРСПЕКТИВЫ РАЗВИТИЯ



Использование стэкинга (Stacked generalization)



СПИСОК ИСТОЧНИКОВ

1. Wenjie Lu, Jiazheng Li, Yifan Li, Aijun Sun, Jingyang Wang, "A CNN-LSTM-Based Model to Forecast Stock Prices", *Complexity*, vol. 2020, Article ID 6622927, 10 pages, 2020. <https://doi.org/10.1155/2020/6622927>
2. M. Tan, S. Yuan, S. Li, Y. Su, H. Li and F. He, "Ultra-Short-Term Industrial Power Demand Forecasting Using LSTM Based Hybrid Ensemble Learning," in *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2937-2948, July 2020, doi: 10.1109/TPWRS.2019.2963109.
3. Ozmaden, D. OHLC Financial data, Version 1. Available from: <https://www.kaggle.com/dozmaden/ohlcfinancial-data> (Дата обращения: 10.03.2021, режим доступа: свободный)

НАШИ КОНТАКТЫ

Озмаден Дениз
dozmaden@edu.hse.ru

Артём Калугин
aokalugin@edu.hse.ru

