

Технічні науки

УДК 004.942

Кухарев Сергій Олександрович

асистент кафедри математичних методів системного аналізу

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Кухарев Сергей Александрович

ассистент кафедры математических методов системного анализа

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского»

Kukhariev Serhii

assistant of Department of Mathematical Methods of System Analysis

National Technical University of Ukraine

"Igor Sikorsky Kyiv Polytechnic Institute"

Олексієнко Ганна Олегівна

студент

Національного технічного університету України

«Київський політехнічний інститут імені Ігоря Сікорського»

Олексиенко Анна Олеговна

студент

Национального технического университета Украины

«Киевский политехнический институт имени Игоря Сикорского»

Oleksiienko Hanna

Student of the

National Technical University of Ukraine

"Igor Sikorsky Kyiv Polytechnic Institute"

**МОДЕЛІ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ НА ПРИКЛАДІ
ВАРТОСТІ АКЦІЙ
МОДЕЛИ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ НА ПРИМЕРЕ
СТОИМОСТИ АКЦИЙ
TIME SERIES FORECASTING MODELS FOR SHARE PRICES**

***Анотація.** В даній роботі розглянуті методи для прогнозування часових рядів – методи згладжування, авторегресії та нейронних мереж.*

***Ключові слова:** часовий ряд, авторегресія, згладжування, нейронні мережі, згорткові нейронні мережі, рекурентні нейронні мережі.*

***Аннотация.** В данной работе рассмотрены методы для прогнозирования временных рядов - методы сглаживания, авторегрессии и нейронных сетей.*

***Ключевые слова:** временной ряд, авторегрессия, сглаживание, нейронные сети, сверточные нейронные сети, рекуррентные нейронные сети.*

***Summary.** This paper describes methods for forecasting time series - smoothing methods, auto regression and neural networks.*

***Key words:** time series, autoregression, smoothing, neural networks, convolutional neural networks, recurrent neural networks.*

Вступ. Фондовий ринок є однією з найважливіших сфер ринкової економіки, оскільки він надає компаніям доступ до капіталу, дозволяючи інвесторам купувати акції в компанії. Оскільки ціна акцій постійно коливається, передбачення того, як буде поводитись фінансовий ринок, є нагальною задачею для економістів. Задача є актуальною для всієї міжнародної економіки, оскільки можливість точного передбачення вартості акцій тісно пов'язано з отриманням

фінансового прибутку, а також зі зменшенням інвестиційного ризику та захисту інвестиційних прибутків від волатильності ринку.

Метою даної роботи є аналіз методів прогнозування часових рядів та виявлення на цій основі параметрів впливу на точність деяких моделей, що використовуються для прогнозування вартості акцій.

Для виконання роботи було взято дані про акції компанії Google з офіційного сайту американського біржового ринку NASDAQ за 5 років, починаючи з 1 квітня 2014 року. Датасет представляє з себе щоденну інформацію про ціну на акцій на початку та в кінці дня, максимальну та мінімальну вартість акції за день, та кількість проданих акцій.

В якості метрики для оцінки точності прогнозу для даної роботи було обрано метрики MAE (Mean Absolute Error) та MSE (Mean Squared Error), які можуть набувати значення від 0 до ∞ , і не враховують напрямок помилок. Чим менше значення приймає показник, тим точнішим є прогноз. MAE вимірює середню абсолютну величину помилок у наборі прогнозів для безперервних змінних. MAE є лінійною оцінкою, що означає, що всі індивідуальні відмінності зважуються однаково в середньому. MSE є середнім квадратичним відхиленням похибки прогнозу. Оскільки помилки підносяться до квадрату перед тим, як вони усереднюються, MSE надає відносно високу вагу великим похибкам. Це означає, що MSE є найбільш корисним, коли великі помилки особливо небажані, що відповідає цілям даної роботи.

Згладжування – це важливий і широко поширений метод прогнозування фінансових ринків. Як правило, різні методи згладжування базуються на концепції ковзних середніх. Це допомагає зменшити вплив випадкового компонента у часовому ряді. Загальна формула для зваженого середнього

$$\hat{y}_t = \frac{1}{N} \sum_{i=1}^N w_i y_{t-i},$$

де N - число попередніх моментів часу, що було взяти до уваги при побудові прогнозу,

u_{k-i} - реальні значення показника в момент часу t_{k-i} ,

w_i - ваговий коефіцієнт для i -того компонента ряду[1].

У випадку простого ковзного середнього вагові коефіцієнти дорівнюють одиниці, а у випадку експоненціально зваженого ковзного середнього коефіцієнти задаються як $\alpha(1 - \alpha)^i, 0 < \alpha < 1$. Таким чином більший акцент робиться на останні точки даних.

Метод ковзного середнього було реалізовано для різних значень параметру N - кількості попередніх моментів часу, що було взяти до уваги при побудові прогнозу (рис. 1). Було перевірено, що при зменшенні довжини вікна N модель показує більш точний результат на тестовій вибірці, що вказує на те, що найостанніші є дані найбільш впливові при прогнозуванні, тобто прогноз при врахуванні останніх 5 днів є більш точним, ніж прогноз, в якому враховуються останні 20 днів.

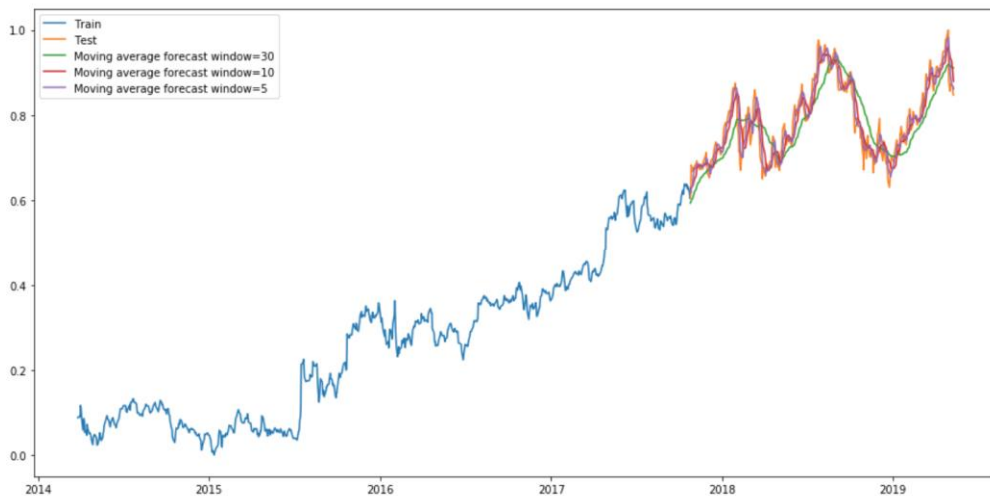


Рис. 1. Графік прогнозу простого ковзного середнього

Для експоненційного ковзного середнього (рис. 2) параметром є рівень згладжування - коефіцієнт α , що являє собою ступінь зменшення зважування від

0 до 1. Чим менший рівень згладжування, тим точніший є прогноз, оскільки кожне попереднє значення важить більше.

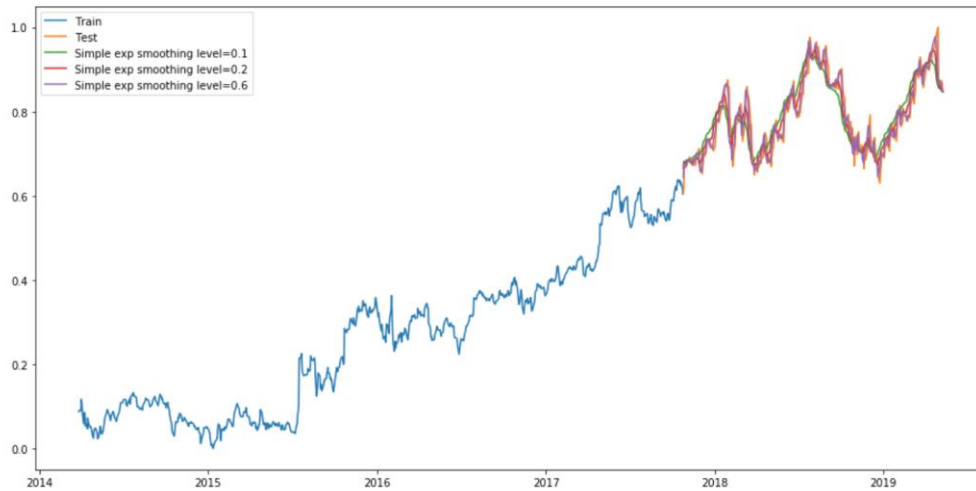


Рис. 2. Графік прогнозу методом експоненціального ковзного середнього

Для даних з чітко вираженим трендом, що відповідає вхідним даним роботи, метод подвійного експоненціального згладжування, що є рекурсивним застосуванням експоненційного фільтра двічі, дає кращий результат. Цей метод передбачав задання додатково параметру β , який відповідає за згладжування тренду (рис. 3). Комбінація пари α та β коригувало точність та якість прогнозу.

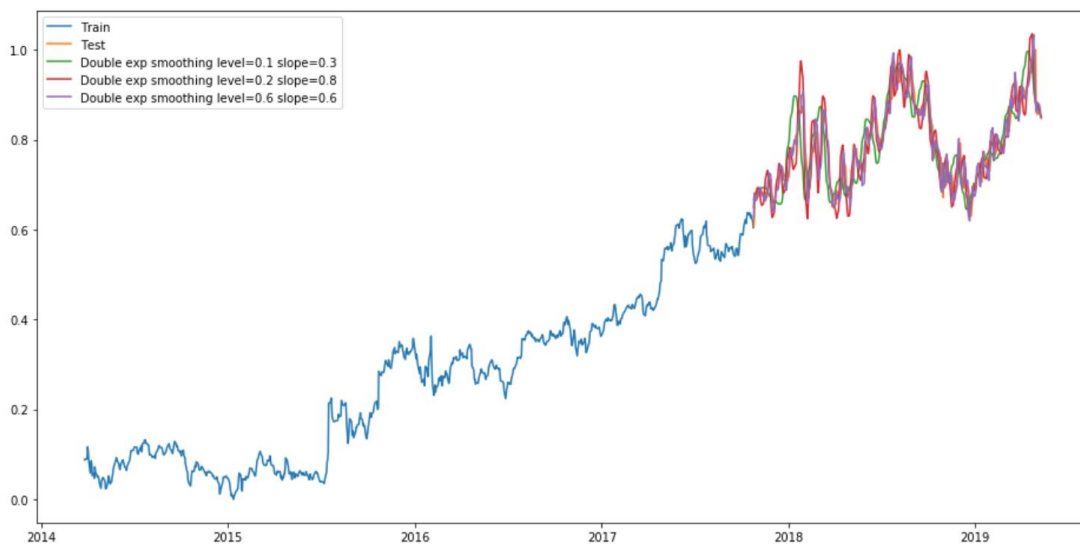


Рис. 3. Графік прогнозу методом подвійного експоненціального ковзного середнього

Результати реалізації усіх методів згладжування зведені до таблиці 1.

Таблиця 1

Результати реалізації методів згладжування

Метод	Характерні параметри	Test MSE	Test MAE
Ковзне середнє	Вікно $N = 30$	0.0029	0.0455
	Вікно $N = 10$	0.0012	0.0281
	Вікно $N = 5$	0.0005	0.0180
Експоненціальне ковзне середнє	Рівень згладжування $\alpha = 0.1$	0.0018	0.0340
	Рівень згладжування $\alpha = 0.2$	0.0011	0.0259
	Рівень згладжування $\alpha = 0.6$	0.00049	0.0165
Подвійне експоненціальне ковзне середнє	$\alpha = 0.1, \beta = 0.3$	0.00244	0.0388
	$\alpha = 0.2, \beta = 0.8$	0.00163	0.0322
	$\alpha = 0.6, \beta = 0.6$	0.00054	0.0177

Модель авторегресії є ефективним інструментом для розуміння і прогнозування майбутніх значень часового ряду, яка включає в себе регресування змінної по значеннях ряду у минулому. Авторегресивні частини цих моделей описують, як послідовні спостереження в часі впливають один на одного, тоді як частини ковзних середніх захоплюють деякі можливі неспостережувані потрясіння.

Модель ARMA характеризує стохастичний процес за допомогою двох компонентів - авторегресії (AR) та ковзного середнього (MA). Частина AR передбачає регресування змінної на власні минулі значення. Частина MA

включає моделювання похибки як лінійної комбінації похибок, що відбуваються в минулому. Позначення $ARMA(p, q)$ характеризує модель з p авторегресійними компонентами і q компонентами для ковзного середнього:

$$y_n = \sum_{i=1}^p a_i y_{n-i} + \sum_{i=1}^q b_i e_{n-i} + e_n,$$

де a_i, b_i - параметри моделі,

e_n - білий шум.

Для застосування моделі ARIMA використовуються формули моделі ARMA, проте на вхід замість y_t подається $\Delta y_t = y_t - y_{t-1}$, і задається додатковий параметр, який вказує на кількість разів, коли до вхідних спостережень було застосовано диференціювання. Для розглянутих методів авторегресії коефіцієнт Акайке дозволив програмно обрати модель, що дає найкращий прогноз. Отже, у випадку даного датасету це виявилась модель ARIMA(2,1,1). Результати виконання усіх трьох методів зведені в одну порівняльну таблицю 2.

Таблиця 2

Результати реалізації методів авторегресії

Метод	Характерні параметри	Test MSE	Test MAE
Проста модель авторегресії	$p = 2$	0.000347	0.013322
Модель авторегресії — ковзного середнього	$p = 3, q = 1$	0.000348	0.013346
Модель авторегресії — інтегрованого ковзного середнього	$p = 2, q = 1, d = 1$	0.000345	0.013335

Штучні нейронні мережі мають перевагу в прогнозуванні часових рядів, оскільки мають потенціал для вирішення складних проблем прогнозування. Важлива особливість ANN стосовно застосування до проблем прогнозування часових рядів полягає в здатності нейронних мереж до нелінійного моделювання,

без будь-якого припущення про статистичний розподіл часового ряду. Кожна модель адаптивно формується на основі даних. З цієї причини штучні нейронні мережі керуються даними та є самоадаптивними за своєю природою [2]. Загальна структура штучної нейронної мережі ґрунтується на сукупності з'єднаних вузлів - нейронів. Вихідне значення нейронної мережі математично задається так:

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g(\beta_{0j} + \sum_{i=1}^p \beta_{ij} y_{t-i}) + \varepsilon_t, \forall t,$$

де p – кількість вхідних змінних,
 q – кількість прихованих вузлів,
 α_j та β_{ij} – вагові коефіцієнти,
 ε_t – випадковий шум.

Найпростішим видом нейронної мережі є одношарова перцептронна мережа (в загальному випадку багатошаровий перцептрон Румельхарта), яка складається з одного шару вихідних вузлів, а входи подаються безпосередньо на виходи через ряд ваг. Універсальна теорема апроксимації для нейронних мереж стверджує, що кожен безперервну функцію, яка відображає інтервали дійсних чисел до деякого вихідного інтервалу дійсних чисел, можна апроксимувати довільно тісно багатошаровим перцептроном лише одним прихованим шаром [3]. Основною характеристикою багатошарового перцептрону є його архітектура, а саме кількість прихованих шарів, вузлів та активуючі функції. Також, оскільки результат навчання частково залежить від ініціалізації змінних, кожна модель тренувалась окремо 5 разів, та всі характеристики для порівняльної таблиці є усередненими значеннями. Для більш глибокого дослідження моделі в роботі було реалізовано 5 архітектур, точність кожної з яких наведено в таблиці 3.

Результати навчання MLP

#	Модель	Кількість параметрів	Training MAE	Training MSE	Test MAE	Test MSE	Алгоритм	Активність
1	MLP 20-10-1	221	0.0645	0.0074	0.090	0.014	adam	relu
2	MLP 20-20-1	441	0.0462	0.0038	0.052	0.005	adam	relu
3	MLP 20-100-1	2201	0.0438	0.0038	0.037	0.002	adam	relu
4	MLP 20-100-1	2201	0.0035	0.0428	0.032	0.002	adam	tanh
5	MLP 20-100-1	2201	0.0668	0.0078	0.093	0.014	sgd	relu

За результатами цієї порівняльної таблиці цілком простежується залежність якості прогнозу від кількості вузлів у прихованому шарі, методу оптимізації та функції активації. Найкращий результат показала модель MLP 20-100-1 з методом оптимізації Adam, та гіперболічним тангенсом в ролі активуючої функції. Наведемо графік похибки прогнозу вартості акції на тестовій вибірці (рис. 4).

Згорткова нейронна мережа – це тип штучної нейронної мережі, в якій картина зв'язності між її нейронами натхненна організацією зорової кори тварини, окремі нейрони якої розташовані таким чином, що вони реагують на перекриваються області поля. Головною перевагою згорткових нейронних мереж є те, що ми використовуємо згорткові шари, щоб виявити ознаки мережі, що дозволяє тренувати нейронну мережу без складної попередньої обробки, оскільки корисні функції будуть вивчені під час навчання [3]. На відміну від багат шарового перцептрону, згорткові нейронні мережі мають складнішу

структуру, оскільки вимагають калібрувати кількість та тип шарів, кількість вузлів в кожному шарі, метод оптимізації, функцію активації та кількість фільтрів. Отже позначення CNN-20-200-3 буде позначати 20 вузлів на вхід як перший шар, 200 фільтрів розміром 3x3, MaxPooling шар та 2 fully-connected шарів. В даній роботі було протестовано 5 різних архітектур CNN, які показали наступні результати таблиці 4:

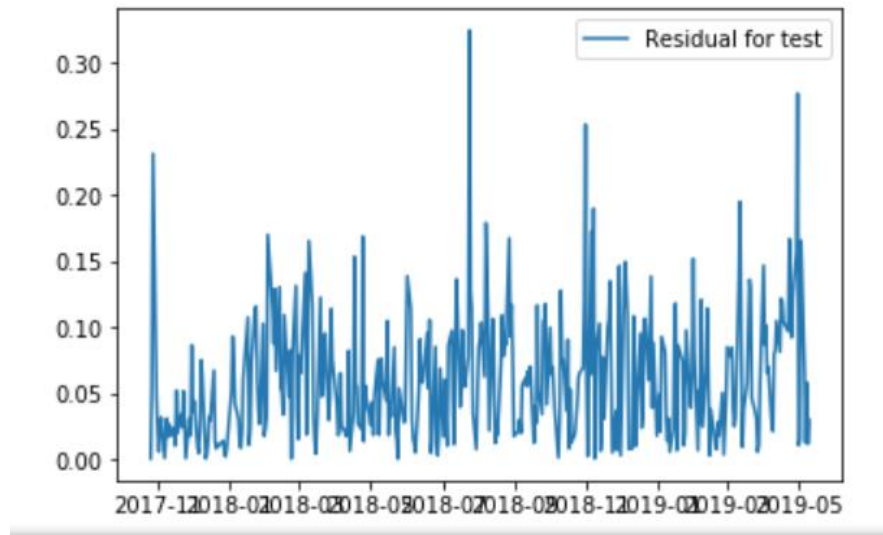


Рис. 4. Графік похибки MLP-20-100-1

Таблиця 4

Результати навчання CNN

#	Модель	Кількість параметрів	Training MAE	Training MSE	Test MAE	Test MSE	Алгоритм	Функція активації
1	CNN 20-256-3	199,937	0.0212	9.4573e-04	0.1243	0.0216	adam	relu
2	CNN 20-256-5	201,367	0.0182	7.4266e-04	0.174	0.0388	adam	relu
3	CNN 10-256-3	198,657	0.0238	0.0011	0.096	0.0137	adam	relu

4	CNN 20-256-3	199,937	0.0513	0.0050	0.29	0.110	sgd	relu
5	CNN 20-500-3	756,501	0.0515	0.0050	0.2592	0.0890	sgd	relu

Найкращий результат показала модель CNN 20-256-3 с методом оптимізації Adam, та активуючий функцією ReLU.

Рекурентна нейронна мережа - це будь-яка штучна нейронна мережа, нейрони якої передають сигнали зворотного зв'язку один одному. Ідея RNN полягає у використанні послідовної інформації. У традиційній нейронній мережі ми припускаємо, що всі входи (і виходи) незалежні один від одного. Але для багатьох завдань це не найкраща ідея. RNN називаються рекурентними нейронними мережами, тому що вони виконують одне і те ж завдання для кожного елемента послідовності, при цьому вихідні дані залежать від попередніх обчислень [4]. Цей тип нейронних мереж цілком підходить для прогнозування вартості акцій, оскільки майбутні кроки можуть залежати від минулих.

Для реалізації прогнозування рекурентних нейронних мереж було обрано модель LSTM (Long Short-Term Memory Units). LSTM допомагають зберегти помилку, яку можна розповсюджувати через час і шари. Підтримуючи більш постійну помилку, вони дозволяють повторним мережам продовжувати вивчати протягом багатьох кроків часу[4]. Для аналізу параметрів впливу на прогноз вартості акцій було реалізовано 5 варіантів архітектуру LSTM, які було зведено в наведену нижче таблицю 4. Позначення LSTM-20-30-30-30-1 використовувалось для визначення нейронної мережі LSTM, з початковими даними за 20 днів назад, кількістю вузлів 30 на першому, другому та третьому шарі LSTM та одним fully-connected шаром з отриманим виходом з одного елемента.

Результати навчання LSTM

#	Модель	Кількість параметрів	Training MAE	Training MSE	Test MAE	Test MSE	Алгоритм	Активність
1	LSTM 20-30-30-30-1	18,631	0.0074	1.0564e-04	0.0198	0.0005	adam	relu
2	LSTM 20-30-30-40-1	22,681	0.0079	1.8716e-04	0.0188	0.0005	adam	relu
3	LSTM 30-30-30-30-1	18,631	0.0106	1.9467e-04	0.0159	0.0004	adam	relu
4	LSTM 10-30-30-30-1	18,631	0.0099	1.7086e-04	0.0150	0.0003	adam	relu
5	LSTM 10-30-30-30-1	18,631	0.0082	1.2895e-04	0.0152	0.0002 36	sgd	relu

Найкращий результат показала модель LSTM 10-30-30-30-1 с методом оптимізації Adam, та активуючою функцією ReLU та з найменшим вікном - лагом.

Висновки. У цьому дослідженні порівнювалися показники прогнозування між нейромережею та методом прогнозування класичних часових рядів, а саме вартістю акцій компанії на фондовій біржі. Проведений аналіз показав, що моделі нейронних мереж, наведені в даному дослідженні, показали дуже набагато кращу здатність точного прогнозування, і отже підтвердив перспективність та доцільність використання штучних нейронних мереж для подальшого дослідження їх застосування на фінансових ринках. На точність прогнозу моделі впливали різні фактори, залежно від її архітектури та вхідних даних.

Загалом аналіз виконання класичних алгоритмів та алгоритмів машинного навчання можна звести у порівняльну таблицю 6.

Порівняльна таблиця реалізованих методів

Моделі	Переваги	Недоліки
Згладжування	Здатність обробляти тенденції змінних рівнів і компоненти сезонності	Вразливі до екстремальних значень
Авторегресія	Можна легко автоматизувати	Сильні обмеження в припущеннях
Штучні нейронні мережі - ANN	Можливість обробки складних нелінійних шаблонів. Висока точність прогнозу	Потребує велику кількість даних.

Література

1. Бідюк П.І. Аналіз часових рядів: навчальний посібник. К: Політехніка, 2010. 317 с.
2. Robert J. Van Eyden The Application of Neural Networks in the Forecasting of Share Prices. New York: Finance and Technology Publishing, 1996. 326 p.
3. Di Persio L. Artificial Neural Networks architectures for stock price prediction: comparisons and applications. *INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING*. 2016. Vol. 31, No. 10. P. 404–405.
4. K. Kamijo, T. Tanigawa. Stock price pattern recognition: A recurrent neural network approach. In *Neural Networks in Finance and Investing*. New Delhi: Probus Publishing Company, 1993. 370 с.