

ПІДВИЩЕННЯ ПРОГНОЗУВАННЯ КУРСУ АКЦІЙ НА ОСНОВІ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ

Shcherbyna I.S., Sachuk S.V., Ishcheryakov S.M. Improving stock price forecasting based on recurrent neural networks. The article is devoted to the study of the application of machine learning for forecasting the stock price in the stock markets in the context of the growing capabilities of artificial intelligence. The article examines the issues and relevance of stock price forecasting, as well as describes current approaches to such forecasting as fundamental and technical analysis. Within technical analysis, widely used methods and algorithms of statistics, pattern recognition, sentiment analysis and machine learning are considered.

The article focuses on the advantages of using recurrent neural networks as one of the types of artificial neural networks that work with sequential data or time series. Traditional neural networks treat data inputs and outputs as independent of each other, while recurrent neural networks take into account previous elements in the sequence to generate the output, which allows for better prediction of price trends.

The article describes the main architectures of recurrent neural networks that exist at the moment and their features, including bidirectional recurrent neural networks, closed recurrent block, and long short-term memory. The last architecture - long short-term memory - was considered in the most detail. This architecture is one of the most difficult to learn and configure, at the same time, it is the most effective for complex tasks such as forecasting the dynamic course of shares. One of its key advantages is solving the problem of gradient damping, that is, the ability to store the context within which price data is analyzed, which makes it possible to detect long-term historical trends and predict future values with high accuracy.

Keywords: Forecasting, time series, machine learning, recurrent neural networks.

Щербина І.С., Сачук С.В., Іщеряков С.М. Підвищення прогнозування курсу акцій на основі рекурентних нейронних мереж. Стаття присвячена дослідженню застосування машинного навчання для прогнозування курсу акцій на фондових ринках у контексті зростаючих можливостей штучного інтелекту. У статті розглянуто проблематику та актуальність прогнозування курсу акцій, а також описано поточні підходи до такого прогнозування як фундаментальний та технічний аналіз. В межах технічного аналізу розглядаються широко застосовувані методи та алгоритми статистики, розпізнавання шаблонів, аналізу настроїв та машинного навчання.

Основну увагу у статті приділено перевагам використання рекурентних нейронних мереж як одному із типів штучних нейронних мереж, які працюють з послідовними даними або часовими рядами. Традиційні нейронні мережі розглядають входи та виходи даних як незалежні один від одного, в той час як рекурентні нейронні мережі враховують попередні елементи у послідовності для генерації виходу, що дає можливість кращого прогнозування цінних трендів.

Стаття описує, які основні архітектури рекурентних нейронних мереж існують на даних момент та їх особливості, зокрема двонаправлені рекурентні нейронні мережі, закритий рекурентний блок, а також довга короточасна пам'ять. Найбільш детально розглянуто останню архітектуру – довга короточасна пам'ять. Дана архітектура є однією з найбільш складних для навчання та налаштування, в той же час, найбільш ефективна для складних задач як прогнозування динамічного курсу акцій. Однією з її ключових переваг є вирішення проблеми затухання градієнту, тобто можливості зберігання контексту в межах якого відбувається аналіз цінних даних, що дає можливість виявляти довгострокові історичні тенденції та прогнозувати майбутні значення з високою точністю.

Ключові слова: Прогнозування, часові ряди, машинне навчання, рекурентні нейронні мережі.

Вступ

Щодня на фондових біржах торгуються акції компаній на сотні мільярди доларів США. Ціна акції є показником як ефективності бізнес моделі і менеджменту конкретної компанії, так і суб'єктивної оцінки цих показників інвесторами та станом макроекономіки загалом.

Акції як один із видів фінансових інструментів, що торгуються на фондових біржах, є і одними з найбільш прибуткових через свою волатильність – швидку зміну ціни в залежності від об'єктивних та суб'єктивних факторів. Відтак, прогнозування цін на акції є надзвичайно важливим завданням щоб досягти високої прибутковості та не втратити свої інвестиції.

В той же час, прогнозування цін на акції є досить складною задачею через широкий спектр факторів, які впливають на ціну. Для вирішення цієї задачі інвестори використовують два основних підходи: фундаментальний та технічний аналіз. Перший вид аналізу підходить для довгострокових інвесторів, оскільки прагне визначити внутрішню цінність акції, в той час як другий зосереджений виключно на статистичних даних тенденцій руху ціни, що відображають співвідношення попиту та пропозиції в певний момент часу.

Інтенсивний розвиток машинного навчання дозволяє застосовувати його досягнення та потенціал для підвищення ефективності прогнозування курсу акцій. Машинне навчання, зокрема рекурентні нейронні мережі, показали хороші результати у такому прогнозуванні. Відтак, подальші дослідження в цій сфері несуть велику перспективу для практичного застосування.

Постановка завдання.

Аналіз руху фондового ринку та цінової поведінки є надзвичайно складним через динамічну, нелінійну, нестационарну, непараметричну, шумну та хаотичну природу ринків. На фондові ринки впливає багато взаємопов'язаних факторів, які включають економічні, політичні, психологічні та специфічні для компанії змінні [1].

Технічний і фундаментальний аналіз є двома основними підходами до аналізу фінансових ринків [2]. Щоб інвестувати в акції та отримати високий прибуток з низьким ризиком, інвестори використовували ці два основні підходи для прийняття рішень на фінансових ринках [3].

Фундаментальний аналіз в основному базується на трьох основних аспектах [4]:

(i) макроекономічному аналізі, такому як валовий внутрішній продукт та індекс споживчих цін, який аналізує вплив макроекономічного середовища на майбутній прибуток компанії,

(ii) галузевий аналіз, який оцінює вартість компанії на основі стану галузі та перспектив,

(iii) аналіз компанії, який аналізує поточну діяльність і фінансовий стан компанії для оцінки її внутрішньої вартості. Для фундаментального аналізу існують різні підходи до оцінки, наприклад техніка апроксимації середнього зростання та техніка апроксимації постійного зростання.

Щодо технічного аналізу, його сфери такого аналізу можна згрупувати в наступні групи [4]: настрої, потік коштів, необроблені дані, тренд, імпульс, обсяг, цикл і волатильність.

Аналіз останніх досліджень та публікацій.

Останні досягнення в аналізі цінних паперів і прогнозуванні діляться на чотири категорії: *статистика, розпізнавання образів, машинне навчання і аналіз настроїв* [5]. Ці категорії здебільшого належать до ширшої категорії технічного аналізу, однак існують деякі методи машинного навчання, які також поєднують ширші категорії технічного аналізу з підходами фундаментального аналізу для прогнозування фондових ринків.

До появи методів машинного навчання *статистичні методи*, які часто передбачають лінійність, стаціонарність і нормальність, забезпечували спосіб аналізу та прогнозування акцій. Часові ряди в аналізі фондового ринку – це хронологічна колекція спостережень, таких як загальні щоденні продажі та ціни акцій [6]. Однією групою статистичних підходів, які підпадають під категорію однофакторного аналізу, завдяки використанню часових рядів як вхідних змінних, є авторегресійне ковзне середнє (ARMA), авторегресійне інтегроване ковзне середнє (ARIMA), узагальнена авторегресійна умовна гетероскедастична (GARCH) волатильність і модель авторегресії з плавним переходом (STAR) [1]. Модель ARIMA є широко використовуваною технікою для аналізу фондового ринку [7]. ARMA поєднує моделі авторегресії (AR), які намагаються пояснити ефект зворотного імпульсу та середнього значення, які часто спостерігаються на торгових ринках, і моделі ковзного середнього (MA), які намагаються вловити ефекти шоку, що спостерігаються в часових рядах. Ключовим

обмеженням моделі ARMA є те, що вона не враховує кластеризацію волатильності, ключове емпіричне явище в багатьох фінансових часових рядах. ARIMA є природним розширенням класу моделей ARMA і може звести нестационарний ряд до стаціонарного. ARIMA адаптується до даних часових рядів для прогнозування майбутніх точок [8].

Розпізнавання шаблонів є синонімом машинного навчання, але щодо аналізу цін на акції ці два методи застосовуються дуже різними способами. Розпізнавання шаблонів зосереджується на виявленні шаблонів і тенденцій у даних. Шаблони на фондових ринках — це повторювані послідовності, знайдені на свічкових діаграмах (кожна свічка візуалізує ціну відкриття, закриття, високу та низьку ціну протягом визначеного інтервалу часу), які фондові аналітики історично використовували як сигнали купівлі та продажу [9]. Технічний аналіз ґрунтується на закономірностях, виявлених безпосередньо в біржових даних; це включає візуальний аналіз діаграм, побудованих з часом, щоб показати варіації ціни, обсягу або інших похідних показників, таких як імпульс. Побудова діаграм – це техніка технічного аналізу для порівняння ринкової ціни та історії обсягу з моделями діаграм для прогнозування майбутньої поведінки ціни на основі ступеня відповідності [10]. Знайомі шаблони діаграм, як правило, отримані з їхніх форм, це прогаліни, шипи, прапорці, вимпели, клини, блюдця, трикутники, голова та плечі, а також різні вершини та низи [2], [11].

Аналіз настроїв є ще одним підходом, який останнім часом використовується для аналізу фондового ринку [12]. Це процес прогнозування фондових трендів за допомогою автоматичного аналізу текстових корпусів, таких як стрічки новин або твіти, специфічні для фондових ринків і публічних компаній. Методи класифікації настроїв в основному поділяються на підхід машинного навчання та підхід на основі лексики, який далі поділяється на підходи на основі словника або корпусу продемонстрували потенціал використання сигналів настрою з неструктурованого тексту для підвищення ефективності моделей для прогнозування тенденцій волатильності на фондовому ринку [13], [14].

Потенціал *машинного навчання* у прогнозуванні фінансових ринків широко вивчався [15]. Завдання машинного навчання в цілому поділяються на контрольоване та неконтрольоване навчання. У контрольованому навчанні доступний набір позначених вхідних даних для навчання алгоритму та спостережених вихідних даних. Однак у неконтрольованому навчанні доступні лише непомічені або спостережувані вихідні дані. Метою контрольованого навчання є навчання алгоритму для автоматичного зіставлення вхідних даних із заданими вихідними даними. Після навчання машина навчилася бачити точку вхідних даних і передбачати очікуваний результат. Мета неконтрольованого навчання — навчити алгоритм знаходити шаблон, кореляцію або кластер у заданому наборі даних. Він також може служити попередником контрольованих навчальних завдань [13]. Кілька алгоритмів були використані для прогнозування курсу акцій. Простіші методи, такі як єдине дерево рішень, дискримінантний аналіз і наївний метод Байєса, були замінені більш продуктивними алгоритмами, такими як випадковий ліс, логістична регресія та нейронні мережі [16].

Завдяки нелінійним характеристикам, керованим даними та легким для узагальнення, багатовимірний аналіз із використанням глибоких штучних нейронних мереж став домінуючим і популярним інструментом аналізу фінансового ринку [1].

Мета статті. Метою цієї статті є аналіз методів підвищення ефективності прогнозування курсу акцій на фондових ринках за допомогою та машинного навчання.

Основний матеріал. Як було зазначено вище, при прогнозуванні цін на акції застосовується широкий спектр підходів, серед яких статистичний підхід, розпізнавання шаблонів, застосування машинного навчання та аналіз настроїв.

Найбільш перспективними на даному етапі є методи прогнозування на основі машинного навчання, оскільки вони мають ряд переваг. Зокрема, моделі машинного навчання з мінімальною або без участі людини взагалі здатні самонавчатися та виявляти шаблони, які не завжди можуть помітити люди-експерти. Також, машини позбавлені ризиків людського фактору, зокрема емоцій жадібності та страху, які часто є ключовими причинами невдалого інвестування або завчасного виходу із потенційно прибуткових інвестицій. Врешті, машини

витрачають значно менше часу на аналіз даних та прийняття рішень, аніж це потрібно від людини-експерта для врахування великої сукупності різних факторів.

Машинне навчання – це галузь штучного інтелекту та інформатики, яка зосереджується на використанні даних і алгоритмів для імітації способу навчання людей, поступово покращуючи його точність [17].

Для прогнозування курсу акцій застосовується багато методів машинного навчання. Традиційними алгоритмами машинного навчання є опорні векторні машини, випадковий ліс, логістична регресія, наївний баєсів класифікатор, дерево класифікації та регресії і екстремальне посилення градієнта, тоді як архітектури глибоких нейронних мереж включають багатошаровий перцептрон, мережа глибокої віри, та рекурентні нейронні мережі [16].

Особливої уваги аналітики фондових ринків почали приділяти рекурентним нейронним мережам, які є одними з типів штучних нейронних мереж. Штучні нейронні мережі були створені як імітація роботи людського мозку для вирішення задач, для яких використання людей-програмістів для розробки відповідних алгоритмів була би не доцільною з точки зору витрат часу та ресурсів. Ці мережі складаються з послідовних шарів вузлів, включаючи вхідний, один або кілька прихованих та вихідний шари. Кожен вузол, або штучний нейрон, зв'язаний з іншими і має вагу та поріг. Якщо вихід конкретного вузла перевищує встановлене порогове значення, цей вузол активується, передаючи дані на наступний рівень мережі. Якщо ж ні, він не пересилає дані далі.

Рекурентна нейронна мережа (РНМ) – це один з типів штучної нейронної мережі, яка працює з послідовними даними або часовими рядами. Ці алгоритми глибокого навчання переважно застосовуються для проблем, для яких важливий порядковий або часовий аспект, таких як розпізнавання мовлення та розпізнавання підписів на зображеннях, переклад мови обробка природної мови тощо. Вони використовуються в таких відомих програмах, Google Translate, голосовий пошук, голосовий асистент Siri. Що робить рекурентні нейронні мережі особливими у порівнянні з іншими штучними нейронними мережами, це їх здатність «пам'ятати» – вони використовують інформацію з попередніх вхідних даних, щоб впливати на поточні вхідні та вихідні дані. Традиційні глибокі нейронні мережі припускають, що входи та виходи є незалежними один від одного, але рекурентні нейронні мережі враховують попередні елементи у послідовності для генерації виходу. Однонаправлені рекурентні нейронні мережі не можуть враховувати майбутні події при прогнозуванні результатів даної послідовності.

Існують різні варіації або архітектури рекурентних нейронних мереж. Як приклад, можна навести такі:

Двонаправлені рекурентні нейронні мережі (BRNN). Це різновид мережевої архітектури РНМ. У порівнянні з однонаправленими РНМ, які базуються лише на попередніх вхідних даних для прогнозування поточного стану, двонаправлені РНМ використовують і майбутні дані для підвищення своєї точності.

Довга короткочасна пам'ять (LSTM). Це популярна архітектура РНМ, яку представили З. Гохрайтер та Ю. Шмідхубер як відповідь на проблему викликану затуханням градієнта. У своїй роботі [18], вони досліджують питання довгострокових залежностей. Якщо попередній стан, що впливає на поточний прогноз, не знаходиться у недавньому минулому, модель РНМ може виявити складнощі у точному прогнозуванні поточного стану. Щоб вирішити цю проблему, в мережах LSTM є «комірки» (нейрон) у прихованих шарах, що мають три вентиля: вхідний, вихідний і забуття. Ці механізми контролюють потік інформації, необхідний для прогнозування виходу мережі. Наприклад, якщо займенники статі, такі як «вона», повторювалися декілька разів у попередніх реченнях, це може бути виключено зі стану цієї комірки.

Закриті рекурентні блоки (GRU). Ця архітектура рекурентної нейронної мережі подібна до LSTM у тому, що вона також призначена для вирішення проблеми обмеження короткочасної пам'яті в моделях РНМ. Замість використання інформації для контролю «стану комірки», вона оперує прихованими станами та має два вентиля замість трьох: вентиль

скидання та вентиль оновлення. Так само, як вентиля у LSTM, ці два вентиля контролюють, скільки та які саме дані зберігаються.

Зазначена вище архітектура LSTM є однією з найбільш складних для навчання та налаштування, в той же час, найбільш ефективна для складних задач як прогнозування курсу акцій.

LSTM може фіксувати історичні тенденції та прогнозувати майбутні значення з високою точністю. Відтак, цей алгоритм є одним з найбільш досліджуваних в даній сфері. Для прикладу, в одному дослідженні було застосовано три різні моделі рекурентної нейронної мережі, а саме базову рекурентну нейронну мережу (RNN), довгу короткочасну пам'ять (LSTM) і закритий рекурентний блок (GRU) на ціні акцій Google, щоб оцінити, який варіант RNN працює краще [19]. Згідно із результатами, LSTM перевершив інші варіанти з точністю 72% на п'ятиденному горизонті, і автори також пояснили та відобразили приховану динаміку RNN.

Загалом, враховуючи колосальний прогрес у машинному навчанні останні роки, а також постійно зростаючі потужності комп'ютерної техніки, перспективи застосування машинного навчання у процесі прогнозування курсу акцій є дуже обнадійливими.

Висновки

Аналіз і прогнозування фондового ринку залишаються перспективною та складною проблемою. У міру того, як стає доступним більше даних, постають нові проблеми в отриманні та обробці даних, щоб отримати необхідні знання та проаналізувати вплив на ціни акцій. Ці виклики включають питання тестування в реальному часі, алгоритмічної торгівлі, довгострокових прогнозів і аналізу настроїв щодо документів компанії.

Дослідники, аналітики та професійні торговці цінними паперами здебільшого зосереджуються на короткостроковому прогнозуванні цін на акції порівняно з довгостроковими, тобто щотижневими чи місячними прогнозами на основі історичних даних. Деякі хороші підходи до довгострокового прогнозування ціни вже існують, такі як ARIMA. Фондові ринки, як правило, більш передбачувані в довгостроковій перспективі. Кілька новіших підходів як штучні нейронні мережі, такі як довга короткочасна пам'ять (LSTM), зараз досліджуються та порівнюються з існуючими підходами для прогнозування довгострокових залежностей у даних і курсах акцій, які є однаково цінними для інвесторів і спеціалістів із обробки даних.

Наразі, є ряд наукових та практичних досліджень, які на противагу класичним методами прогнозування, що здійснюється на основі суб'єктивного аналізу фундаментальних та технічних показників експертами та потребує професійних знань та і значний час, підтверджують ефективності моделей машинного навчання. Відтак, ця галузь є надзвичайно актуальною та перспективною для подальших досліджень.

Список використаної літератури:

1. Zhong, Xiao, and David Enke. 2017. Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications* 67: 126–39.
2. Park, Cheol-Ho, and Scott H. Irwin. 2007. What do we know about the profitability of technical analysis? *Journal of Economic Surveys* 21: 786–826.
3. Arévalo, Rubén, Jorge García, Francisco Guijarro, and Alfred Peris. 2017. A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. *Expert Systems with Applications* 81: 177–92.
4. Hu, Yong, Kang Liu, Xiangzhou Zhang, Lijun Su, E. W. T. Ngai, and Mei Liu. 2015. Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review. *Applied Soft Computing* 36: 534–51.
5. Dev Shah, Haruna Isah, Farhana Zulkernine, Stock Market Analysis: A Review and Taxonomy of Prediction Techniques, *Int. J. Financial Stud.* 2019, 7(2), 26

6. Fu, Tak-chung, Fu-lai Chung, Robert Luk, and Chak-man Ng. 2005. Preventing Meaningless Stock Time Series Pattern Discovery by Changing Perceptually Important Point Detection. Paper presented at the International Conference on Fuzzy Systems and Knowledge Discovery, Changsha, China, August 27–29.
7. Hiransha, M., E. A. Gopalakrishnan, Vijay Krishna Menon, and Soman Kp. 2018. NSE stock market prediction using deep-learning models. *Procedia Computer Science* 132: 1351–62.
8. Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. Hoboken: John Wiley & Sons.
9. Velay, Marc, and Fabrice Daniel. 2018. Stock Chart Pattern recognition with Deep Learning. *arXiv* arXiv:1808.00418.
10. Leigh, William, Naval Modani, Russell Purvis, and Tom Roberts. 2002. Stock market trading rule discovery using technical charting heuristics. *Expert Systems with Applications* 23: 155–59.
11. Leigh, William, Cheryl J. Frohlich, Steven Hornik, Russell L. Purvis, and Tom L. Roberts. 2008. Trading with a Stock Chart Heuristic. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38: 93–104.
12. Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2: 1–8.
13. Bhardwaj, Aditya, Yogendra Narayan, and Maitreyee Dutta. 2015. Sentiment analysis for Indian stock market prediction using Sensex and nifty. *Procedia Computer Science* 70: 85–91.
14. Seng, Jia-Lang, and Hsiao-Fang Yang. 2017. The association between stock price volatility and financial news—A sentiment analysis approach. *Kybernetes* 46: 1341–65.
15. Shen, Shunrong, Haomiao Jiang, and Tongda Zhang. 2012. *Stock Market Forecasting Using Machine Learning Algorithms*. Stanford: Department of Electrical Engineering, Stanford University, pp. 1–5.
16. Ballings, Michel, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. 2015. Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications* 42: 7046–56.
17. Machine learning. URL: <https://www.ibm.com/topics/machine-learning>.
18. Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, 9(8), pp.1735-1780.
19. Di Persio, Luca, and Oleksandr Honchar. 2017. Recurrent Neural Networks Approach to the Financial Forecast of Google Assets. *International Journal of Mathematics and Computers in simulation* 11: 7–13.

Автори статті

Щербина Ірина – кандидат технічних наук, доцент, Державний університет інформаційно-комунікаційних технологій, Київ, Україна.

Сачук Сергій – студент, Державний університет інформаційно-комунікаційних технологій, Київ, Україна.

Іщеряков Сергій – кандидат технічних наук, доцент, Державний університет інформаційно-комунікаційних технологій, Київ, Україна.

Authors of the article

Shcherbyna Iryna - Candidate of Science (technic), associate professor, State University of Information and Communication Technologies, Kyiv, Ukraine.

Sachuk Serhii – student, State university of Information and Communication Technologies, Kyiv, Ukraine.

Ishcheryakov Serhiy - Candidate of Science (technic), associate professor, State University of Information and Communication Technologies, Kyiv, Ukraine.