

Дослідження Модифікації Гібридної Рекурентної Нейронної Мережі для Прогнозування Цін Акцій

Дмитро Качан
Інститут Прикладного Системного Аналізу
КПІ ім. Ігоря Сікорського
Київ, Україна
dskach4n@gmail.com

Надія Недашківська
Інститут Прикладного Системного Аналізу
КПІ ім. Ігоря Сікорського
Київ, Україна
nedashkovskaya.nadezhda@iik.kpi.ua

Research on Modification of a Hybrid Recurrent Neural Network For Stock Price Forecasting

Dmytro Kachan
Institute for Applied System Analysis
Kyiv Polytechnic Institute
Kyiv, Ukraine
dskach4n@gmail.com

Nadezhda Nedashkovskaya
Institute for Applied System Analysis
Kyiv Polytechnic Institute
Kyiv, Ukraine
nedashkovskaya.nadezhda@iik.kpi.ua

Анотація—Запропоновано гібридну модель нейронної мережі, яка поєднує рекурентні, згорткові шари та механізм уваги для прогнозування цін акцій. Експериментальним шляхом модель порівняно з LSTM та GRU моделями автокодувальників. Розроблено систему прогнозування цін акцій статистичними методами та методами машинного навчання.

Abstract—A hybrid neural model is proposed, which combines recurrent, convolutional layers and an attention mechanism for stock price prediction. Experimentally, the model is compared with LSTM and GRU models of autoencoders. A system for forecasting stock price using statistical and machine learning methods has been developed

Ключові слова—рекурентні і згорткові нейронні мережі; LSTM; GRU; модель кодувальник-декодувальник; модель послідовність у послідовність; механізм уваги; регуляризація

Keywords—recurrent and convolutional neural networks; LSTM; GRU; encoder-decoder; sequence-to-sequence models; attention mechanism; regularization

I. ВСТУП

Прогнозування цін акцій – важлива задача в фінансовій сфері, яка може допомогти інвесторам приймати обґрунтовані рішення. З розвитком штучного інтелекту, глибокі нейронні мережі стали потужним інструментом для прогнозування складних даних, таких як ціни акцій.

В роботі досліджено модифікацію гібридної нейронної мережі, яка складається з рекурентних [1, 2] та згорткових шарів, для прогнозування цін акцій компанії Amazon. Мета – знайти модель, яка може точно прогнозувати ціни акцій на основі історичних даних, враховуючи складні динамічні залежності, які існують на фондовому ринку.

Навчання рекурентної нейронної мережі включає її розгортання в часі і використання алгоритму зворотного розповсюдження помилки Backpropagation Through Time (BPTT) [3]. Велика глибина розгорнутої в часі мережі може призводити до проблеми вибухового зростання градієнтів, яка частково вирішується прийомами відсічення градієнтів або укороченого зворотного розповсюдження в часі (truncated BPTT). Спеціальні типи комірок, такі як LSTM [4] та GRU [2] краще вирішують як вказану проблему, так і додаткові проблеми навчання моделей на довгих послідовностях [5]: складність вивчити довготривалі образи та поступове зникнення пам'яті перших входів. Комірки LSTM і GRU – одна з головних причин успіху рекурентних нейронних мереж за останні роки.

II. СИСТЕМА ПРОГНОЗУВАННЯ ЦІН АКЦІЙ

Розроблено систему прогнозування цін акцій, яка включає наступні складові:

- Збір та підготовка даних: історичних у формі часових рядів, текстів новин, оцінок експертів.
- Побудова різних моделей прогнозування цін акцій: економетричних, рекурентних нейронних мереж та інших. Навчання моделей на підготовлених даних.
- Налаштування гіперпараметрів моделей.
- Тестування моделей на відкладеному наборі даних, оцінка їх продуктивності і точності. Побудова прогнозів. Оцінка якості прогнозів.
- Аналіз результатів моделювання на основі результатів валідації та тестування моделей. Визначення області їх можливого покращення.

III. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Для дослідження використано щоденні дані про ціни акцій компанії Amazon з 01 січня 2020 року по 29 листопада 2022 року, отримані за допомогою сервісу ufinance. Дані включають чотири характеристики: Open, High, Low та Close – ціна відкриття, найвища, найнижча та ціна закриття акцій протягом дня. Моделі будувалися для прогнозування кожної з цих характеристик (цін).

Для налаштування гіперпараметрів моделей використано бібліотеку Keras Tuner та метод Random Search для автоматичного пошуку найкращого набору значень гіперпараметрів, серед яких швидкість навчання, параметри регуляризації, розміри рекурентних, згорткових, повнозв'язних та інших шарів.

Як видно з таблиці 1, запропонована гібридна модель CNN-LSTM з увагою надає більш точний (за показником MAPE) прогноз наступного значення в часовому ряді [6], порівняно з LSTM та GRU моделями автокодувальників.

Крім того, запропонована гібридна модель CNN-LSTM набагато швидше навчається порівняно з іншими дослідженими моделями. Це робить її більш ефективною для обробки великих обсягів даних.

ТАБЛИЦЯ I. Порівняння моделей на однокроковому прогнозі

Модель	Значення MAPE, %			
	Open	High	Low	Close
ARIMA	0.63	-	-	-
LSTM Autoencoder	2.82	3.46	4.85	1.51
GRU Autoencoder	3.32	3.34	5.71	2.00
Пропонована модель CNN-LSTM з механізмом уваги	0.97	0.58	0.85	3.52

ТАБЛИЦЯ II. Порівняння моделей на багатокроковому прогнозі

Модель	Значення MAPE, %			
	Open	High	Low	Close
ARIMA	0.35	-	-	-
LSTM Autoencoder	3.46	4.34	3.81	4.26
GRU Autoencoder	1.82	2.15	2.38	2.8
Пропонована модель CNN-LSTM з механізмом уваги	11.77	15.86	12.63	13.67

В задачі багатокрокового прогнозування на 25 кроків вперед модель CNN-LSTM з увагою мала вищі значення середньої абсолютної відносної похибки MAPE, ніж інші побудовані моделі автокодувальників (таблиця 2) [6].

Обмеження дослідження в тому, що воно зосереджено лише на прогнозуванні цін акцій однієї компанії (Amazon). Результати можуть відрізнятися для інших компаній чи індексів. По-друге, ефективність моделі не гарантує її успішності у майбутньому, оскільки фондовий ринок постійно змінюється. Нарешті, важливо пам'ятати, що жодна модель не може з абсолютною точністю прогнозувати ціни акцій.

Поєднання прогнозів на основі різних моделей пропонується робити методами підтримки прийняття рішень [7, 8] з урахуванням множини критеріїв рішень, оцінок експертів та особи, яка приймає рішення.

IV. ВИСНОВКИ

Дослідження показало, що використання гібридної нейронної мережі CNN-LSTM з механізмом уваги – перспективний підхід для однокрокового прогнозування цін акцій.

Для багатокрокового прогнозу, на обраному наборі запропонована модель продемонструвала більш високі значення похибки MAPE, ніж інші побудовані моделі LSTM та GRU автокодувальників. Проте запропонована модель має переваги в швидкості навчання.

Розроблено модулі системи підтримки прийняття рішень у сфері фінансів. Така система допомагає інвесторам аналізувати дані та робити більш обізнані інвестиційні рішення. Пропоновану модель слід використовувати разом з фундаментальним аналізом та враховувати інші фактори, які впливають на ціни акцій.

ЛІТЕРАТУРА REFERENCES

- [1] I. Sutskever, O. Vinyals and Q. V. Le, "Sequence to sequence learning with neural networks", arXiv preprint arXiv:14093215, 2014.
- [2] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning Phrase Representations using RNN EncoderDecoder for Statistical Machine Translation", arXiv preprint arXiv:1406.1078, 2014.
- [3] P. J. Werbos, "Generalization of backpropagation with application to a recurrent gas market model", Neural Networks, vol. 1, no. 4, pp. 339-356, 1988.
- [4] S. Hochreiter, J. Schmidhuber, "Long short-term memory", Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [5] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber. "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies" in S.C. Kremer and J. F. Kolen, eds., A Field Guide to Dynamical Recurrent Neural Networks. IEEE press, 2001.
- [6] Д. С. Качан. (2023). "Прогнозування цін акцій методами глибоких нейронних мереж". Available: <https://ela.kpi.ua/items/d9b2e9c4-0bff-442c-81b5-1a1b412bbbb6>.
- [7] N.I. Nedashkovskaya. (2022). "Estimation of the Accuracy of Methods for Calculating Interval Weight Vectors based on Interval Multiplicative Preference Relations", IEEE 3rd Int. Conf. on Syst. An. & Int. Comp., Available: <https://ieeexplore.ieee.org/document/9922977>.
- [8] N. I. Nedashkovskaya, "Method for weights calculation based on interval multiplicative pairwise comparison matrix in decision-making models", Radio Electronics, Computer Science, Control, no. 3, pp. 155-167, 2022.