

ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ З ВИКОРИСТАННЯМ МАСИВІВ ІНДИКАТОРІВ У PYTHON

¹Вінницький національний технічний університет;

Проведено аналіз методів прогнозування фінансових часових рядів, розглянуто моделі з використанням інтелектуальних технологій. Запропоновано власну модель для прогнозування часових рядів з використанням масивів індикаторів у Python. Прогнозування часових рядів є застосування моделі для прогнозування майбутніх значень на основі раніше спостережених значень. І хоча регресійний аналіз часто використовується для перевірки теорії про те, що поточне значення одного чи кількох незалежних часових рядів впливає на поточне значення іншого часового ряду, цей аналіз часових рядів не називається «аналізом часових рядів», а фокусується на порівнянні значень одного часового ряду в різні моменти часу. В якості тестових даних було взято курс валютної пари EUR/USD за 03.08.19 - 03.03.20. Після огляду існуючих методів прогнозування часових рядів за допомогою Python було обрано модель LSTM. Проаналізувавши отримані результати, отриманий прогноз сигналів можна використовувати як систему підтримки рішень. Точність прогнозування наперед досягнута за рахунок використання LSTM моделі, нормуванню вхідного датасету, використанням допоміжних технічних індикаторів та підбором параметрів для моделі прогнозування. Після огляду існуючих методів прогнозування часових рядів за допомогою Python було обрано модель LSTM. Використовуючи LSTM, моделі прогнозування часових рядів можуть передбачати майбутні значення на основі попередніх послідовних даних. Це забезпечує більшу точність для прогнозистів попиту, що призводить до кращого прийняття рішень для бізнесу. Проблему з обчисленнями вирішує cuDNN від NVIDIA. З досвіду, cuDNN може запровадити в 10+ разів більшу швидкість, ніж стандартне використання безпосередньо CUDA. Виходячи з того, що передбачені значення ціни Close не є досить точними, було спроектовано сигнали для рекомендацій покупки/продажу позицій. Даний сигнал являє собою значення від 0 до 1, де 0 є купівля, а 1 – продаж.

Ключові слова: прогнозування, фінансовий часовий ряд, інтелектуальні технології, моделі прогнозування, системи прийняття рішень.

Вступ

Часовий ряд – це серія точок даних, проіндексованих у хронологічному порядку. У більшості випадків часовий ряд – це послідовність, взята в рівновіддалені моменти часу, які слідують одна за одною. Таким чином, це послідовність даних у дискретному часі [1]. *Мета роботи* полягає у підвищенні ефективності прийняття рішень на фінансових часових рядах за рахунок інтеграції інтелектуальних технологій Python та застосування масивів технічних індикаторів. Питання розробки нових підходів задля вирішення задачі прийняття рішень на фінансових часових рядах є актуальними на сьогодні, особливо за використанням інтелектуальних технологій.

Постановка задачі

Прогнозування є одним із найскладніших, але водночас найактуальніших і популярних завдань аналізу даних. Складність процесу прогнозування пов'язана з необхідністю аналізу та оцінки великих масивів даних, ускладненням методів, появою нових підходів до процесів прогнозування, застосування нових технологій тощо [2]. Тому сучасний стан розвитку методів прогнозування тісно пов'язаний з розвитком інформаційних технологій. Так звані прогнозні інформаційні системи, які відображають цей зв'язок у рамках економетрики, фінансової математики та статистики, проявляються в широкому діапазоні наукових застосувань, а також у таких сферах, як виробництво, фінансове планування в економіці та торгівлі. Сьогодні вони є частиною процесу управління

складними системами та системами прийняття управлінських рішень і використовуються аналітиками, серед іншого, для оцінки ризиків фінансових інвестицій.

Основними компонентами системи прогнозування часових рядів є база даних з постійно поповнюваними ретроспективними даними, комплексна модель прогнозування та методи оцінки його якості, які згруповані відповідно до постановки завдання прогнозування.

Функціонування такої системи здійснюється в режимі діалогу з особою, яка приймає рішення.

Вхідні дані, представлені у вигляді дата-сетів, в які входять також масиви показників індикаторів. Крім того відомі моменти за певний проміжок часу, в які було б бажано прийняти те чи інше рішення. Необхідно створити підхід, базуючись на методах прогнозування часових рядів і за використанням бібліотек Python, та навчити нейронну мережу прогнозувати майбутні значення на основі даних в реальному часі: цінних вимірів та набору значень індикаторів.

Результати дослідження

Для вирішення проблеми запропоновано даний алгоритм дій:

1) інтеграція з торговим терміналом MT5 для отримання масиву даних: пакет MetaTrader використовується для зручного та швидкого отримання біржової інформації в Python. Отримані таким чином дані можна використовувати для статистичних обчислень і машинного навчання.

2) прогін даних, технічних індикаторів (запропоновані MT5 або створені за допомогою Python), формування дата сету для моделі та нейронної мережі[3]: дані потрібно привести до вигляду, який “розуміє” нейронна мережа, виділити основні дані масиву, доповнити масив технічними індикаторами (пошук входів за стратегією Елдера [4]).

3) навчання рекурентної нейронної мережі та її тестування, прогнозування за допомогою моделі LSTM: за допомогою цих інструментів будуємо графік, оснований на попередніх даних, таким чином визначаємо майбутню ціну валютної пари.

4) аналіз результатів та тестування моделі з різними наборами даних.

В якості тестових даних було взято курс валютної пари EUR/USD за 03.08.19 - 03.03.20 (Рис. 3). Використовується один ряд – Close. Для формування масивів індикаторів слід використовувати функції Pandas TA, що дозволяють видобути такі індикатори, як ковзке середнє. У даному випадку масив індикаторів утвориться на основі пари середніх з використанням оцінки різних часових вимірів (2MA systems). Суть методу, що створить індикатори для датасету, полягає у співставленні двох ковзких середніх на різних часових вимірах (такі, як H4, M15). В залежності від цього співставлення можна визначати сигнали відкриття/закриття позицій. Для більш точних прогнозів можливе застосування RSI та MFI індикатори. Для збору індикаторів, їх обчислення запропонована бібліотека ta-lib, що містить більше 150 індикаторів. Після огляду існуючих методів прогнозування часових рядів за допомогою Python було обрано модель LSTM [5].



Рис. 1. Набір даних з використанням технічних індикаторів

Використовуючи LSTM, моделі прогнозування часових рядів можуть передбачати майбутні значення на основі попередніх послідовних даних. Це забезпечує більшу точність для прогнозистів попиту, що призводить до кращого прийняття рішень для бізнесу.

За допомогою використання бібліотек Pandas та Pandas TA стає можливим не тільки створити DataFrame, а й наповнити його технічними індикаторами. Таким чином, не потрібно підраховувати та вивантажувати індикатори, використовуючи функціонал MT5, достатньо викликати відповідні методи з бібліотеки Pandas TA (Рис. 1):

При тренуванні спостерігалось зменшенні втрат впродовж кожної епохи – такий результат досягнуто в першу чергу великою вибіркою даних та підбором параметрів моделі.

Передбачення складає собою ціну Close на 12 проміжків уперед. На рисунку 2 показано детальне порівняння прогнозованих даних (синій колір) з реальними даними (червоний).



Рис. 2. Порівняння прогнозованих даних із реальними

Слід зауважити що розбіжність результатів сягає у середньому 150 пункта, для даного часового проміжку досить багато. При цьому можна спостерігати збереження тренду, що є позитивним результатом.

Виходячи з того, що передбачені значення ціни Close не є досить точними, було спроектовано сигнали для рекомендацій покупки/продажу позицій. Даний сигнал являє собою значення від 0 до 1, де 0 є купівля, а 1 – продаж.

Змінивши вихідну змінну моделі прогнозування та додавши даних сигнал у набір даних було отримано графік, зображений на рисунку 3. Сигнал купівлі/продажу реалізований таким чином, що дає значення у проміжку від 0 до 1, де значення, наближені до нуля – є сигналом купівлі, а наближені значення до одиниці – сигнал продажу.



Рис. 3. Спроектовані сигнали купівлі/продажу

Проаналізувавши отримані результати, отриманий прогноз сигналів можна використовувати як систему підтримки рішень. Точність прогнозування насамперед досягнута за рахунок використання LSTM моделі, нормуванню вхідного датасету, використанням допоміжних технічних індикаторів та підбором параметрів для моделі прогнозування.

Висновки

В роботі проведено аналіз існуючих методів прогнозування фінансових цінових рядів, розглянуто можливості застосування інтелектуальних технологій для прогнозування. Запропоновано власний алгоритм із застосування бібліотек обробки даних до прогнозування, інтеграції з торговим

терміналом та використанням технічних індикаторів, що показують задовільні результати прогнозування на невеликих проміжках часу.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] Сохацька, О. М. Фундаментальний та технічний аналіз цін товарних та фінансових ринків: навч. посіб. / О. М. Сохацька, І. В. Роговська-Ішук, С. І. Вінницький. – К. : Кондор, 2012. – 305 с.
- [2] Берзлев О. Ю. *Адаптивні комбіновані моделі прогнозування біржових показників* / О. Ю. Берзлев, М. М. Маляр, В. В. Ніколенко // Вісник Черкаського держ. технолог. ун-ту. Серія: технічні науки. – 2011. – № 1. – С. 50-54.
- [3] Кветний, Р. Н., Кабачій, В. В., Чумаченко, О. О. // Імовірнісні нейронні мережі в задачах ідентифікації часових рядів - 2010. - 2-3 с. / [Електронний ресурс] - Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/4609>.
- [4] С.І. Бакай, В.В. Кабачій // *Водель прийняття рішень на основі пари середніх з використанням оцінки різних часових вимірів* [Електронний ресурс] – Режим доступу: <http://ir.lib.vntu.edu.ua/bitstream/handle/123456789/13106/Buckeye%20%D0%9A%D0%B0bachiye.pdf?sequence=1&isAllowed=y>
- [5] Oraz R. LSTM Model. Medium [Електронний ресурс] – Режим доступу: <https://towardsdatascience.com/lstm-for-time-series-prediction-de8aeb26f2ca>
- [6] LSTM – PyTorch 2.4 documentation. *PyTorch*. [Електронний ресурс] – Режим доступу: <https://pytorch.org/docs/stable/generated/torch.nn.LSTM.html>

Кабачій Владислав Володимирович — к.т.н, доцент кафедри автоматизації та інформаційно-вимірювальної техніки, e-mail: kabachij.v.v@vntu.edu.ua

Маслій Роман Васильович — к.т.н, доцент, доцент кафедри АІТ, факультет комп'ютерних систем та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: romas4580@gmail.com

Кулик Ярослав Анатолійович – к.т.н., доцент кафедри Автоматизації та інтелектуальних інформаційних технологій, Факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: kulyk.y.a@vntu.edu.ua

Вінницький національний технічний університет;

В.В. Кабачій¹

Р.В. Маслій¹

Я.А. Кулик¹

Time series forecasting using indicator arrays in Python

¹Vinnitsia National Technical University;

An analysis of methods of forecasting financial time series was carried out, models using intelligent technologies were considered. A proprietary model for forecasting time series using arrays of indicators in Python is proposed. Time series forecasting is the application of a model to predict future values based on previously observed values. And although regression analysis is often used to test the theory that the current value of one or more independent time series affects the current value of another time series, this time series analysis is not called "time series analysis" but focuses on comparing the values of one time series in different moments of time. The rate of the EUR/USD currency pair for 08/03/19 - 03/03/20 was taken as test data. After reviewing existing methods for time series forecasting using Python, the LSTM model was chosen. After analyzing the obtained results, the received signal forecast can be used as a decision support system. The accuracy of forecasting is primarily achieved due to the use of the LSTM model, normalization of the input dataset, the use of auxiliary technical indicators and the selection of parameters for the forecasting model. After reviewing existing methods for time series forecasting using Python, the LSTM model was chosen. Using LSTM, time series forecasting models can predict future values based on past sequential data. This provides greater accuracy for demand forecasters, leading to better business decisions. NVIDIA's cuDNN solves the computational problem. From experience, cuDNN can deliver 10+ times faster than standard use of CUDA directly. Based on the fact that the predicted Close price values are not accurate enough, signals were designed to recommend buy/sell positions. This signal represents a value between 0 and 1, where 0 is a buy and 1 is a sell.

Keywords: forecasting, financial time series, intelligent technologies, forecasting models, decision-making systems.

Kabachiy Vladyslav V. — Candidate of Technical Sciences, Assistant Professor of the Department of Automation and Information-Measuring Equipment, e-mail: kabachij.v.v@vntu.edu.ua

Maslii Roman V. – Candidate of Technical Sciences, Associate Professor, Department of Computer Systems and Automation, Vinnitsia National Technical University. Vinnitsa, e-mail: romas4580@gmail.com.

Kulyk Yaroslav Anatoliyovych – Candidate of Technical Sciences, Associate Professor of Automation and Intelligent Information Technologies Department, Faculty of Intelligent Information Technology and Automation, Vinnitsia National Technical University, Vinnitsia, e-mail: kulyk.y.a@vntu.edu.ua ;