

ПОРІВНЯЛЬНИЙ АНАЛІЗ МОДЕЛЕЙ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ

Вінницький національний технічний університет

Анотація

Дана робота присвячена аналізу та порівнянню ефективності трьох провідних архітектур рекурентних нейронних мереж (RNN) глибокого навчання — LSTM, GRU та Bi-LSTM. Розглянуто застосування цих моделей для вирішення задачі прогнозування фінансових часових рядів, зокрема цін на акції. Проведено огляд архітектурних особливостей, що дозволяють цим моделям долати обмеження класичних RNN. Виконано порівняльний аналіз переваг та недоліків кожної архітектури з огляду на обчислювальну складність, швидкість навчання та точність прогнозування довгострокових залежностей.

Ключові слова: часові ряди, прогнозування, ціни акцій, глибоке навчання, LSTM, GRU, Bi-LSTM, рекурентні нейронні мережі, зникаючий градієнт.

Abstract

This work focuses on the analysis and performance comparison of three leading deep learning recurrent neural network (RNN) architectures: LSTM, GRU, and Bi-LSTM. The application of these models to the task of forecasting financial time series, specifically stock prices, is examined. A review of the architectural features that enable these models to overcome the limitations of classical RNNs is provided. A comparative analysis of the advantages and disadvantages of each architecture is conducted, considering computational complexity, training speed, and the accuracy of long-term dependency forecasting.

Keywords: time series, forecasting, stock prices, deep learning, LSTM, GRU, Bi-LSTM, recurrent neural networks, vanishing gradient.

Вступ

Точне прогнозування фінансових часових рядів залишається однією з ключових та найскладніших задач у галузі фінансової аналітики та машинного навчання. Висока волатильність, нелінійність та стохастична природа ринкових даних суттєво обмежують ефективність традиційних статистичних підходів. У цьому контексті, архітектури глибокого навчання, зокрема рекурентні нейронні мережі (RNN), здобули значну увагу завдяки їхній здатності моделювати складні часові залежності.

Однак, класичні RNN мають суттєві обмеження, такі як проблема деградації (зникаючого) градієнта, що унеможливило ефективне вивчення довгострокових патернів. Це стимулювало розробку більш просунутих архітектур, таких як LSTM, GRU та Bi-LSTM. Метою даної роботи є порівняльний огляд цих трьох ключових архітектур для визначення їх відносних переваг та недоліків у прикладному завданні прогнозування цін акцій.

Порівняльний аналіз моделей RNN

LSTM (Long Short-Term Memory). Архітектура LSTM була запропонована як фундаментальне вирішення проблеми деградації градієнта [1]. Її ключовою інновацією є впровадження комірки стану (cell state) та системи трьох вентилів (gates): вхідного (input), забування (forget) та вихідного (output). Саме цей механізм дозволяє мережі вибірково зберігати інформацію протягом тривалих періодів, ігноруючи нерелевантний "шум" та фокусуючись на значущих паттернах. Вентиль забування, зокрема, дозволяє моделі скидати застарілу інформацію, що є критичним для волатильних фінансових рядів.

Однак, така гнучка та складна архітектура має свою ціну: головним недоліком є обчислювальна складність. Велика кількість параметрів, зумовлена трьома вентилями та окремою коміркою стану, призводить до значних витрат ресурсів та тривалого часу навчання [2].

GRU (Gated Recurrent Unit). GRU є відносно новою та ефективною модифікацією LSTM, яка має на

меті спрощення архітектури при збереженні її функціональних переваг [3]. Ключові переваги цієї моделі випливають безпосередньо з її спрощення: GRU об'єднує комірку стану та прихований стан, а також замінює три вентиля LSTM на два: вентиль оновлення (update gate) та вентиль скидання (reset gate). Це призводить до суттєвого зменшення кількості параметрів. Як наслідок, GRU навчається значно швидше та вимагає менше даних для уникнення перенавчання. Дослідження демонструють, що GRU часто досягає порівнянної точності з LSTM.

Попри це, спрощення має й теоретичні обмеження: у випадках із надзвичайно довгими послідовностями або коли потрібен дуже точний контроль над потоком інформації, дещо простіша архітектура GRU може теоретично поступатися LSTM.

Bi-LSTM (Bidirectional LSTM). Двонаправлений LSTM (Bi-LSTM) є розширенням, яке дозволяє мережі обробляти дані у двох напрямках одночасно: вперед (з минулого в майбутнє) та назад (з майбутнього в минуле). Принципова перевага такого підходу полягає в тому, що стандартний LSTM враховує лише минулий контекст для прогнозування. Bi-LSTM, завдяки другому, зворотному шару, отримує повний контекст послідовності для кожної часової точки. Численні дослідження підтверджують, що це часто дозволяє моделі виявляти патерни, пропущені однонаправленою моделлю, що призводить до найвищої точності прогнозування [4].

Разом з тим, ця перевага у точності створює два суттєвих недоліки. По-перше, обчислювальна складність є найбільшою, оскільки кількість параметрів фактично подвоюється. По-друге, що є більш суттєвим обмеженням для прогнозування цін акцій, для обробки поточної точки модель вимагає доступу до майбутніх точок (інформація зі зворотного шару). Це робить її ідеальною для ретроспективного аналізу історичних даних, але її застосування для прогнозування в реальному часі є ускладненим [5].

Висновки

Проведений аналіз демонструє, що не існує єдиної "найкращої" архітектури; вибір залежить від конкретних вимог задачі: доступних обчислювальних ресурсів, обсягу даних та мети моделювання.

Модель GRU є оптимальною відправною точкою, пропонуючи найкращий баланс між точністю та обчислювальною ефективністю. Вона особливо корисна в умовах обмежених ресурсів або при необхідності швидкої ітерації та тестування гіпотез.

На противагу їй, LSTM залишається надійним та потужним вибором, коли обчислювальні ресурси не є критичним обмеженням, а дані містять дуже складні та тривалі залежності, що вимагають більш тонкого налаштування потоків інформації.

У свою чергу, коли пріоритетом є не швидкість навчання, а виключно максимальна точність, Bi-LSTM показує найвищий потенціал завдяки повному контекстному аналізу. Проте, як зазначалося, його використання виправдане переважно для пакетної обробки або ретроспективного аналізу історичних даних, де доступний повний часовий ряд.

Таким чином, вибір оптимальної моделі є компромісом між потребами у точності, наявними ресурсами та специфікою часового ряду, що аналізується.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms [Електронний ресурс] – Режим доступу до ресурсу: <https://www.mdpi.com/2673-2688/2/4/30>
2. Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market [Електронний ресурс] – Режим доступу до ресурсу: <https://www.mdpi.com/2674-1032/3/4/29>
3. A Novel Multivariate Bi-LSTM model for Short-Term Equity Price Forecasting [Електронний ресурс] – Режим доступу до ресурсу: <https://arxiv.org/html/2409.14693v1>
4. Unveiling the Precision of Deep Learning Models for Stock Price Prediction: A Comparative Analysis of Bi-LSTM, LSTM, and GRU [Електронний ресурс] – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/10366646>
5. Enhancing Stock Price Forecasting Accuracy Using LSTM and Bi-LSTM Models [Електронний ресурс] – Режим доступу до ресурсу: https://www.itm-conferences.org/articles/itmconf/pdf/2025/01/itmconf_dai2024_04008.pdf

Морозов Дмитро Павлович – студент групи 2KN24-м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: 01-24-175.stud@vntu.edu.ua

Паночішин Юрій Миколайович – кандидат технічних наук, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця, e-mail: y.panochyshyn@vntu.edu.ua

Morozov Dmytro P. – student of group 2KN24-m, Faculty of Intellectual Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: 01-24-175.stud@vntu.edu.ua

Panochyshyn Yuriy M. – Candidate of Technical Sciences, Associate Professor of the Department of Computer Science, Vinnytsia National Technical University, Vinnytsia, e-mail: y.panochyshyn@vntu.edu.ua