

Порівняння LSTM та GRU рекурентних нейронних мереж для прогнозування цін на акції

Порівняння ефективності та швидкості роботи алгоритмів глибокого навчання в прогнозуванні цін фондового ринку. Для дослідження було запропоновано рекурентні нейронні мережі - LSTM та GRU.

Фондовий ринок є важливою сферою економіки, а його дослідження – це одна з найскладніших проблем останніх років. Коливання цін на акції має нелінійний характер, оскільки включає багато факторів – фізичні фактори проти раціональної та ірраціональної поведінки тощо. Поспання всіх аспектів робить фондовий ринок нестабільним та важким для точного прогнозу. Мета прогнозування ціни на фондовому ринку – це збільшити прибуток, мінімізуючи ризики.

Застосування традиційних методів, таких як фундаментальний та технічний аналіз, може не забезпечити надійність прогнозу, здебільшого використовують регресійний аналіз. З розвитком штучного інтелекту, все частіше для прогнозу застосовують методи глибокого навчання та нейронні мережі.

Глибине навчання — це набір алгоритмів для моделювання високих рівнів абстракцій в дані, використовуючи глибокий граф із деякими шарами обробки, які побудовано з кількох лінійних або нелінійних перетворень.

Методи глибокого навчання LSTM та GRU є популярним рішенням в області прогнозування цін на ринку акцій. Саме тому ці мережі були взяті для порівняння їх ефективності.

Особливість штучних нейронних мереж у прогнозуванні часових рядів полягає в їх здатності до нелінійного моделювання, без будь-яких відомих вхідних даних про статистичний розподіл даного часового ряду. Модель штучної нейронної мережі формується адаптивно.

Рекурентні нейронні мережі (РНМ) — це клас штучних нейронних мереж, де зв'язки між вузлами утворюють спрямований граф уздовж часової послідовності. Це дозволяє йому демонструвати часову динамічну поведінку.

РНМ застосовується до послідовних даних та можуть бути використані для побудови моделей регресії. Є дві основні частини, вхідні дані та прихований стан, які збирають результати з попереднього вузла. Ці дві частини поєднуються, і результати вводяться в наступний вузол. РНМ мають так звану проблему довгострокової залежності, тобто виникають труднощі з ідентифікацією та вивченням інформації минулого часу в межах даних, що знаходяться на довгих відстанях часових кроків.

Довга короткочасна пам'ять (LSTM) - це особливий різновид архітектури рекурентних нейронних мереж, що вирішує проблему РНМ з даними часових рядів з довгою послідовністю. Адже, довга короткочасна пам'ять має стан

комірки, яка утримує стан кожного вузла, щоб він міг змінити значення, отримані з попереднього стану, і має блок керування потоком.

Керуючий рекурентний блок (GRU) має ту ж загальну функціональність, що і алгоритм LSTM. Кожен вузол має процес, поєднуючи ворота забуття та входні ворота в шлюз оновлення. Продуктивність еквівалентна алгоритму LSTM. Однак GRU має простішу архітектуру та менше параметрів для генерації послідовних даних.

LSTM та GRU мають внутрішні механізми (ворота), за допомогою яких можна регулювати дані входних та вихідних блоків пам'яті. Реалізація цих воріт була виконана за допомогою логістичної функції для обрахунку даних в діапазоні [0;1].

Часовий ряд — це ряд точок даних, проіндексованих в хронологічному порядку. Зазвичай сукупність таких даних представлена лінійною діаграмою. Дані можуть бути річні, кварталні або щомісячні. Для отримання об'єктивного прогнозу цін на акції дані повинні відповідати таким вимогам: однаковий крок за часом для досліджуваного періоду цін; для кожного часового відрізка значення зберігаються у той самий момент; виключення виникнення пропуску даних.

Модельовання машинного навчання - це процес побудови та оцінки моделей прогнозування. Для створення моделі, яка реалізує нейронну мережу, що використовується для прогнозування часових рядів фінансових ринків, була застосована послідовна модель.

Першою і однією з найважливіших стадій глибинного навчання являється збір даних, в процесі якого має бути сформована навчальна вибірка. Набори даних для навчання нейронної мережі повинні бути достатньо повними, мати великий обсяг і потрібну інформацію для подальшого знаходження мережею особливостей та встановлення зв'язку.

У процесі підготовки даних часових рядів до їх подачі у процес прогнозування, була використана техніка розсувного вікна, яка бере вихідні дані і перетворює індекс часу в атрибут у форматі перерізу.

Для навчальної та тестової вибірки було використано дані компанії Google з офіційного сайту фінансового ринку NASDAQ за 5 років, з 1 квітня 2015 року по 1 квітня 2020 року. Обраний набір даних це інформація на початку та в кінці кожного дня про ціну на акції, максимальну і мінімальну вартість акції, та кількість акцій, які були продані за добу. Створення часового ряду була застосована функція генерування часових рядів. Після цього ми запускаємо функцію для обчислення конкретних значень та генерування навчальних зразків для нашої нейронної мережі, яку потрібно навчити. Потім розподіл даних на навчальний набір та набір тестів. Ціна закриття акцій використовувалася як цільова для прогнозування на наступний день. В якості вихідних даних використовувались приблизні ціни.

Параметри методів глибокого навчання підібрані наступним чином: вхідний вузол = '1215', вихідний вузол = '1', прихований вузол = '700', епоха = '150', функція активації = 'ReLU', методом оптимізації - ''.

Алгоритм оптимізації Adam - це розширення стохастичного градієнтного спуску, яке нещодавно стало широко застосовуватись для програм глибокого навчання. Його використовують замість класичної стохастичної процедури

градієнтного спуску для оновлення вагових коефіцієнтів мережі ітеративно на основі навчальних даних.

При аналізі результатів роботи тестові дані використовувались для оцінки узагальнення моделей прогнозу.

Ефективність прогнозування вимірювали за показниками MAE та MSE для аналізу відповідності історичних даних короткостроковому прогнозуванню цін на акції даних часових рядів. За допомогою MAE (Mean Absolute Error) можна дізнатись середню абсолютну величину помилок у наборі прогнозів для безперервних змінних. MSE (Mean Squared Error) вимірює середнє квадратичне відхилення похибки прогнозу.

Час обробки LSTM становив 13,028 сек. Час обробки GRU - 10,980 сек.

Прогнозування на основі тестового набору є лінійним графіком, як показано на рис.1.

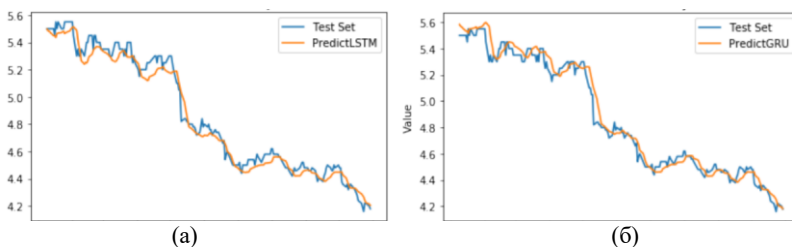


Рис. 1. Моделі прогнозу (а) LSTM та (б) GRU.

Враховуючи напрямок руху тренду як тестових, так і прогнозованих даних, виявляється, що передбачувані значення поєднуються з реальними значеннями тесту. У порівнянні ефективності прогнозування між LSTM та GRU, модель LSTM може передбачати з меншою похибкою, ніж модель GRU. Однак обробка LSTM займає більше часу.

Список літератури

1. J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015
2. Бідюк П.І. Аналіз часових рядів: навчальний посібник. К: Політехніка, 2010. 317 с.
3. Кухарев С. О., Олексіенко Г. О. Моделі прогнозування часових рядів на прикладі вартості акцій // Міжнародний науковий журнал "Інтернаука". — 2019. — №9.