

ЛАБОРАТОРНА РОБОТА №3. ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

Теоретична довідка

Набори даних часових рядів, як-от ціни на фондовому ринку, метеорологічні дані або дані про продажі та ін. використовуються для прогнозування майбутніх значень шляхом вивчення історичних моделей і тенденцій.

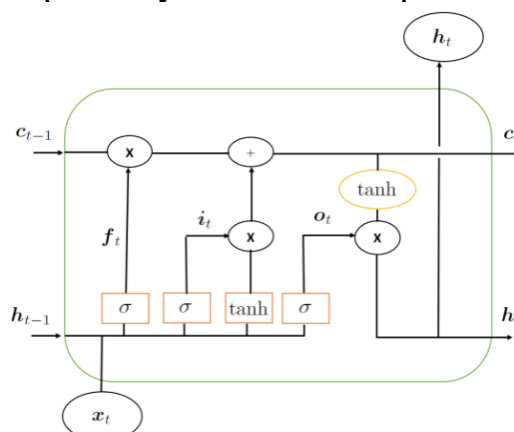
Можна передбачити часові ряди за допомогою різноманітних моделей, починаючи від простих статистичних методів, таких як ковзаючі середні та експоненціальне згладжування, до більш просунутих методів машинного навчання, таких як ARIMA, LSTM. Для точного прогнозування ці моделі враховують сезонність, тенденції та інші характерні закономірності в даних часових рядів.

Серед інших програм передбачення часових рядів використовується у фінансових прогнозах, прогнозуванні попиту. У цій лабораторній роботі буде досліджено та реалізовано прогнозування часових рядів за допомогою LSTM.

Як зазначалося вище, прогнозування часових рядів зазвичай виконується за допомогою мереж LSTM (довгокороткочасної пам'яті), які є потужними варіантами рекурентних нейронних мереж (RNN).

Опис архітектури LSTM

Модуль довготривалої короткострокової пам'яті (LSTM) був розроблений для подолання слабких місць рекурентної нейронної мережі (RNN), таких як проблема зникаючих градієнтів. RNN складається з ряду повторюваних модулів нейронної мережі, де кожен модуль складається зі структури, яка включає єдиний \tanh -шар. \tanh використовується для стиснення значень між -1 і 1 . Ми зберігаємо вхідні дані в прихованому стані h_t , який також містить дані з попереднього прихованого стану h_{t-1} . LSTM складається з такої ж серії повторюваних модулів, що й RNN, за винятком того, що структура модуля LSTM є складнішою. Кожен модуль містить чотири шари, а не єдиний шар, як для RNN, і це допомагає вирішити проблему зникаючого градієнта RNN.



Три шлюзи керують даними, які будуть додані або видалені до стану комірки. Мережа LSTM обчислює відображення від вхідного вектора $x = (x_1, \dots, x_t)$ до вихідного вектора $y = (y_1, \dots, y_t)$, де спочатку вхідний вектор активації шлюзу і значення кандидатів стану комірки пам'яті C_t обчислюються за допомогою такого набору рівнянь:

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i),$$

$$C_t = \tanh(W_c * [h_{t-1}, x_t] + b_c).$$

де σ – логістична сигмоїдна функція, W – відповідні ваги, b – відповідні зміщення в цих рівняннях. $[h_{t-1} + x_t]$ стосується конкатенації попереднього прихованого стану та векторів введення. C_t — це вектор нових значень-кандидатів, створених кожним шаром \tanh , які потрібно додати до стану. Потім визначається вектор активації пропуску f_t :

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f),$$

Шлюз забування значень слугує для видалення даних із стану комірки, а також допомагає скинути комірки пам'яті. Оцінені значення i_t , C_t і f_t потім використовуються для визначення нового стану комірки пам'яті C_t .

$$C_t = f_t * C_{t-1} + i_t * C_t,$$

Стан як конвеєр, який проходить по всьому ланцюжку. Потім вектор стану комірки C_t використовується для обчислення вектора активації вихідного затвора o_t , який потім можна використовувати для визначення вихідного сигналу LSTM.

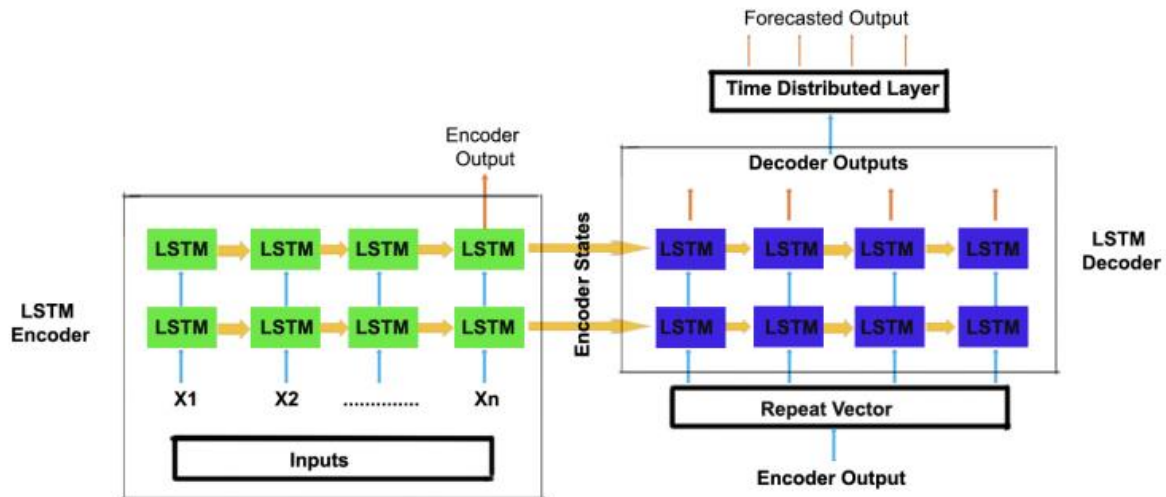
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o),$$

$$h_t = o_t * \tanh(C_t).$$

Накопичуваний seq2seq автокодувальник послідовності LSTM

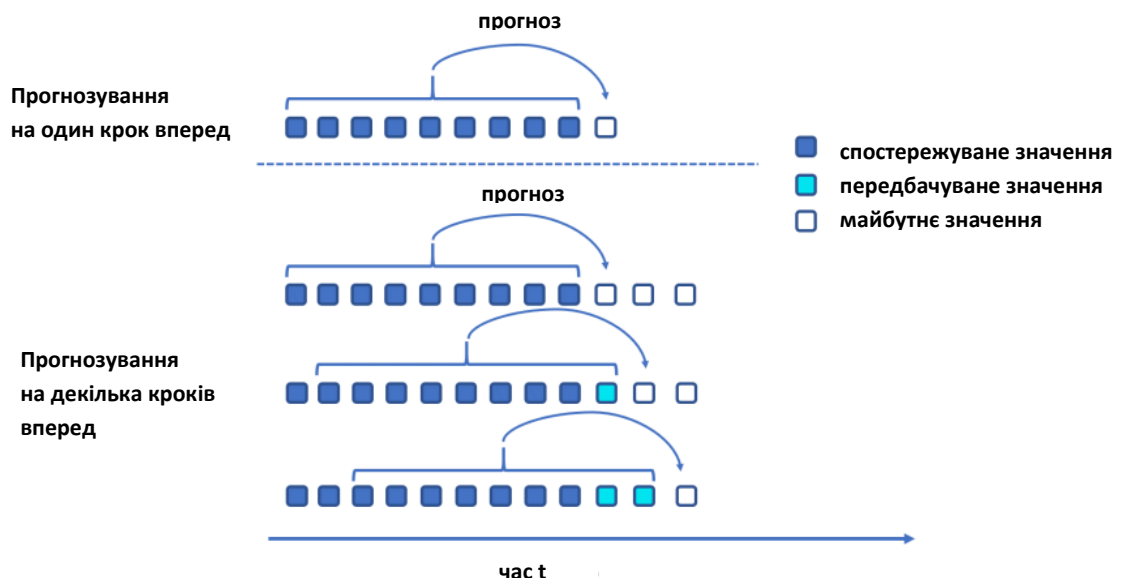
Модель LSTM відображає вхідну послідовність у вихідну послідовність. Модель seq2seq містить два LSTM, які є кодувальником (encoder) і декодувальником (decoder). Кодувальник змінює задану вхідну послідовність на контекстний вектор, який підсумовує вхідну послідовність. Контекстний вектор подається як вхідні дані для декодера, а кінцевий стан кодувальника є вихідним станом декодувальника для прогнозування вихідної послідовності. Ми можемо використовувати послідовність для послідовного навчання для багатоетапного прогнозування часових рядів, додавши два шари: повторюваний векторний шар і розподілений у часі щільний шар. Рівень повторюваного вектора використовується для повторення вектора контексту, який ми отримуємо від кодувальника, щоб передати його як вхідні дані в декодувальник. Ми повторимо це для n -кроків (n – це кількість майбутніх кроків, які ми хочемо спрогнозувати). Вихід, отриманий від декодера для кожного кроку часу включено. Щільно

локалізований час застосує повністю пов'язаний щільний шар на кожному часовому кроці та розділить вихід для кожного часового кроку. Щільно локалізований час є обгорткою, яка дозволяє використовувати шар для кожної часової частини вводу. Накопичування додаткових шарів на частині кодувальника та частині декодувальника моделі послідовності може покращити здатність нашої моделі вивчати складніші картини наших даних часового ряду в прихованих шарах, беручи дані на різних рівнях (див. рисунок нижче).



Завдання для самостійної роботи

1. Для набору даних згідно варіанту виконати попередню обробку та підготовку даних, зокрема розділити набір даних на 3 частини (для навчання 80%, для тестування 10% і для перевірка 10%), та при потребі нормалізувати та згладити.
2. Виконати прогнозування часових рядів із застосуванням нейронної мережі на один крок вперед.
3. Виконати прогнозування часових рядів із застосуванням нейронної мережі на декілька кроків вперед.



4. Визначити точність прогнозування із використанням метрик MAPE, RMSE, AME.

Варіанти завдань

	Набір даних
1	https://raw.githubusercontent.com/jbrownlee/Datasets/master/daily-min-temperatures.csv
2	https://raw.githubusercontent.com/jbrownlee/Datasets/master/daily-max-temperatures.csv
3	https://raw.githubusercontent.com/jbrownlee/Datasets/master/daily-total-female-births.csv
4	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-airline-passengers.csv
5	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-robberies.csv
6	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-sunspots.csv
7	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly_champagne_sales.csv
8	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-car-sales.csv
9	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-mean-temp.csv
10	https://raw.githubusercontent.com/jbrownlee/Datasets/master/monthly-writing-paper-sales.csv
11	https://raw.githubusercontent.com/jbrownlee/Datasets/master/yearly-water-usage.csv