

ĐẠI HỌC BÁCH KHOA HÀ NỘI
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



BÁO CÁO TUẦN 6

Môn học: Project II

*Chủ đề: Xây dựng mô hình dự đoán giá cổ phiếu bằng mạng
Transformer*

Giáo viên hướng dẫn:

Đỗ Tuấn Anh

Sinh viên thực hiện:

Lê Minh Triết

Mã số sinh viên:

20220045

Hà Nội - 2025

MỤC LỤC

Contents

MỤC LỤC.....	3
1. Công việc đã hoàn thành trong tuần	4
1.1. Thay đổi Multi-head Attention bằng Local Attention, Causal Attention và Restricted Causal Attention	4
1.2. Triển khai các attention mới trên Kaggle	4
2. Các vấn đề gặp phải	4
2. Dự kiến các công việc tuần tới	4

1. Công việc đã hoàn thành trong tuần

1.1. Thay đổi Multi-head Attention bằng Local Attention, Causal Attention và Restricted Causal Attention

Từ báo cáo tuần trước, một trong những vấn đề của mô hình Transformer mà em đã xác định là cơ chế Multi-head Attention có vẻ là không phù hợp. Lý do em đưa ra là do giá cổ phiếu có tính tương quan ngắn hạn nên việc tập trung vào toàn bộ chuỗi đầu vào là không hoàn toàn hợp lý. Chính vì vậy, tuần này em thử nghiệm 3 attention khác. Đó là:

Local Attention: Tập trung vào n ngày xung quanh giá trị đầu vào. Nó sẽ có một tham số window xác định cửa sổ nhìn thấy của Transformer. Ví dụ, khi model xét tới ngày thứ i , nó sẽ chỉ chú ý tới các ngày trong khoảng i -window tới i +window.

Causal Attention: Tập trung vào các ngày trước ngày hiện tại. Cả Multi-head Attention và Local Attention đều có thể nhìn vào các giá trị các ngày trong tương lai. Với dữ liệu thời gian điều này có vẻ không ổn. Causal Attention giúp Transformer sẽ không học gì từ các ngày trong tương lai.

Restricted Causal Attention: Đây là sự kết hợp của Local Attention và Causal Attention, nó sẽ giúp mô hình chỉ tập trung vào n ngày trước ngày hiện tại.

Việc triển khai các mô hình Transformer chỉ đơn giản là thay bộ Attention ở Encoder Layer trong kiến trúc, chi tiết em đã để ở

<https://github.com/trietp1253201581/StockPrediction/blob/main/model/transformer.py>

1.2. Triển khai các attention mới trên Kaggle

Em đã triển khai 3 loại transformer trên Kaggle, tuy nhiên gặp phải một số vấn đề em sẽ đề cập tới trong phần 2. Việc triển khai có thể xem ở version 4 của Notebook

<https://www.kaggle.com/code/trietp1253201581/stock-prediction-test-model>

2. Các vấn đề gặp phải

Trong quá trình triển khai 3 mô hình Transformer mới, em gặp phải một vấn đề là **loss trên tập validation và test biến động mạnh qua mỗi lần huấn luyện khác nhau**. Mặc dù em đã cố gắng đặt seed cho quá trình random nhưng sự biến động mạnh vẫn xảy ra (với val-loss, có ngày em huấn luyện đạt 0.06, có lần huấn luyện chỉ đạt 0.20).

Thực chất vấn đề này đã xảy ra từ triển khai mô hình Transformer tuần trước, tuy nhiên sau khi huấn luyện lại 1 lần em đạt được loss tốt và không để ý nhiều. Nhưng tuần này vấn đề tỏ ra nghiêm trọng hơn.

Vì vậy, tuần tới em dự định sẽ tập trung giải quyết vấn đề này.

2. Dự kiến các công việc tuần tới

- Giải quyết tính ổn định của các mô hình Transformer.
- Tối ưu tham số cho các mô hình với các Attention mới.