

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



BÁO CÁO TUẦN 8

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. Tăng tính ổn định của các mô hình Transformer	4
1.2. Tối ưu tham số và triển khai Local Transformer và Restricted Causal Transformer	4
2. Dự kiến các công việc tuần tới	4

1. Công việc đã hoàn thành trong tuần

1.1. Tăng tính ổn định của các mô hình Transformer

Qua việc tìm hiểu các phương pháp giúp tăng tính ổn định của Transformer, em đã lựa chọn các phương pháp sau:

1. Lưu giữ trạng thái tốt nhất (có val loss nhỏ nhất) trong quá trình huấn luyện vào file checkpoint, sau đó khi huấn luyện xong có thể load lại trạng thái này. Đây là cách làm chủ yếu em lựa chọn, và khi triển khai nó thì 1 mô hình Transformer (với các siêu tham số đầu vào, tham số huấn luyện giống nhau) qua các lần huấn luyện khác nhau sẽ cho kết quả không sai khác quá nhiều.
2. Thực hiện dừng sớm (early stopping) nếu không cải thiện.
3. Cấu hình Scheduler chính xác hơn trong quá trình huấn luyện.

Nhờ các phương pháp trên, độ ổn định của các mô hình Transformer đã được tăng lên đáng kể. Nếu lúc trước với cùng tham số nhưng với các lần huấn luyện khác nhau, loss trên tập test có thể nhảy từ 5700 lên 13000, thì giờ luôn dao động quanh ngưỡng 5700 (sai khác không quá 1000).

Ngoài ra, để tránh overfitting, em cũng đề xuất và sẽ thử nghiệm trong tuần sau phương pháp chọn lựa 2-3 checkpoint tốt nhất từ quá trình huấn luyện và chọn lựa mô hình cuối cùng từ đó (sẽ ưu tiên các mô hình có dropout cao hơn để tránh overfitting).

1.2. Tối ưu tham số và triển khai Local Transformer và Restricted Causal Transformer

Em đã tối ưu và triển khai Local Transformer (trong version 6+8+11) và Restricted Causal Transformer (trong version 9+10+11) của notebook.

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

Kết quả thu được có thể tóm tắt là: Local Transformer cho kết quả khá tốt trên train set và valid set, nhưng trên test set thì Restricted Causal Transformer tỏ ra tốt hơn.

Điều này xảy ra có thể là do Local Transformer có thể nhìn về tương lai, trong khi Restricted Causal Transformer thì không.

Do vấn đề giới hạn thời gian huấn luyện GPU miễn phí trên Kaggle, em mới chỉ thử tối ưu hai mô hình trên với khoảng 17-20 lần thử tham số mỗi mô hình. Vì vậy, trong tuần tới nếu có thể sẽ tối ưu thử mô hình còn lại và huấn luyện lại theo chiến lược ở phần 1.1

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

2. Dự kiến các công việc tuần tới

- Thử nghiệm Positional Encoding không tuần hoàn.
- Tiếp tục ổn định các mô hình Transformer.