

# FPT Notebook

Code được thực hiện trên Google Colab, Anh/Chị có thể sẽ cần chỉnh sửa một chút ở phần “Libraries and Data” để chạy code trên máy tính khác.

Hoặc Anh/Chị có thể truy cập link: và bấm vào nút “Open in Colab”

Thời gian chạy code khoảng 3 phút.

## Notebook có 5 phần chính:

1/ **Libraries and Data:** cài đặt, import libraries cần thiết và dữ liệu.

2/ **Data Preprocessing:** preprocess data để phù hợp với mô hình.

(sử dụng giá và volume theo ngày, thay vì theo phút như dữ liệu gốc, kiểm tra missing values, etc).

3/ **Modelling:** xây dựng mô hình. Có 3 mô hình: Tree-based models (Decision Tree và Gradient Boost) và Deep learning (GRU).

**Ý tưởng đối với Tree-based models:** sử dụng dữ liệu của 5 ngày gần nhất và volume của 1 ngày gần nhất để dự báo giá ngày hôm sau.

Kết quả	Training set shape	Test set shape	MSE	2 important features
Tree Regression	(300, 6)	(194, 1)	1.34	Giá của 2 ngày gần nhất.
Gradient Boosting	(300, 6)	(194, 1)	0.75	Giá của 2 ngày gần nhất.

**Ý tưởng đối với Deep Learning:** dùng dữ liệu của 3 ngày gần nhất để dự báo giá ngày hôm sau.

Kết quả	Training set shape	Test set shape	MSE
GRU	(296, 3, 1)	(195, 3, 1)	44.2

**Những điều cần thử nghiệm thêm:**

- Dùng MSE là chưa đủ để đánh giá độ chính xác. Giả sử giá hiện tại là 100, nếu mô hình dự báo giá ngày mai là 110 hoặc 90 thì MSE đều là 100 nhưng nó không giúp được chúng ta quyết định mua hay bán. Nên kết hợp thêm với một hình classification.
- Số lần giá cổ phiếu FPT tăng và giảm trong dữ liệu là 253 và 246 lần tương ứng với tỷ lệ tăng giảm hằng ngày là 50%. Mô hình classification kết hợp phải có độ chính xác lớn hơn 50% (nếu không việc mua bán cũng giống việc tung đồng xu).
- Thử nghiệm Reinforcement Learning (code bị lỗi ở mục Draft).
- Tuning models