**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

🙤🙧🟍🙥🙦



**ỨNG DỤNG PHÂN TÍCH DỮ LIỆU**

**CRYPTOCURRENCY HISTORICAL PRICES**

|  |  |  |
| --- | --- | --- |
| Sinh viên thực hiện: | | |
| STT | Họ tên | MSSV |
| 1 | Mã Hải Nhật | 19521953 |
| 2 | Đấu Vũ Hoàng | 19521530 |

**TP. HỒ CHÍ MINH – 5/2022**

# 1. GIỚI THIỆU

Với sự phát triển của công nghệ mới hiện nay kèm theo đó cũng làm tăng phương thức thực hiện tội phạm liên quan đến tài chính , từ đó con người đã phát triển ra hệ thống tiền ảo để lưu trữ tiền kèm với những phương thức bảo mật chặt chẽ như Bitcoin , sử dụng thuật toán SHA256 , Ethereum Ethash ,....vì có những phương thức bảo mật tuyệt vời nên tiền ảo có thể sẽ là bước nền cho sự phát triển kinh tế kỹ thuật số nên nhóm đã lựa chọn đề tài : làm ứng dụng phân tích dữ liệu về lịch sử giá tiền ảo (Cryptocurrency Historical Prices)

Thực Hiện :

* Tìm bộ dữ liệu từ Kaggle : Cryptocurrency Historical Prices Dataset
* Vẽ thiết kế hệ thống OOP bằng StarUML
* Ứng dụng có các chức năng sau :
* Sắp xếp theo yêu cầu(sắp xếp theo giá và loại coin)
* Dự đoán tương lai dựa theo vốn hóa thị trường (Market capitalization)

Kết quả:

* Hoàn thành được các chức năng sắp xếp cơ bản
* Có thể thông qua 1 số model máy học dự đoán được tương lai của loại coin

Kỳ vọng chưa hoàn thành:

* Build Application

Trong báo này, nhóm trình bày ba nội dung chính: (1) Quy trình áp dụng xây dựng hệ thống, (2) Thiết lập (3) Đánh giá

# NỘI DUNG

## Quy trình áp dụng xây dựng hệ thống

### Dataset:

Lấy từ Kaggle dataset gồm 23 tập dữ liệu .csv chứa thông tin lịch sử của 23 loại coin

Giải thích field

Có 10 field gồm : SNO(thứ tự) , Name(tên đầy đủ của loại coin) , Date(ngày ghi nhận dữ liệu) , High(đỉnh trong ngày) ,Low(đáy trong ngày) ,Open(giá mở phiên),Close(giá đóng phiên) ,Volume(lượng giao dịch) ,Marketcap(vốn hóa thị trường)

### OOP Diagram

Diagram

Description automatically generated

Hình 1 Sơ Đồ Hệ Thống

Giải Thích

* Main class : dung để chạy 2 class cũng là 2 chức năng chính của ứng dụng là sắp xếp theo yêu cầu và dự đoán
* forecast : dự đoán dựa vào Marketcap(vốn hóa thị trường) bằng các model máy học
* SortByNeed:sắp xếp theo nhu cầu thao khảo các chức năng từ các app phân tích như binance nên sẽ có các chức năng xếp theo giá(1 số người dung thích mua theo nhu cầu túi tiền của họ) , xếp theo loại coin ,show các thông số của coin trong 7 30 90 180 360 ngày vì đây không phải dự án live get dữ liệu trực tiếp nên không thể làm trong 1 ngày được
* Preprocess : class này dùng để import thư viện sử dụng trong suốt quá trình thực hiện và tiền xử lý dữ liệu sau đó sẽ chuyển cho 2 class kia

## Thiết lập

* Thư Viện (đã được đề cập ở phần phụ lục)

Giải thích 1 số thư viện

* Tensorflow.keras sử dụng để train mô hình neural network
* Threading để chạy toàn bộ hàm được gói trong 1 hàm duy nhất
* Datetime để khử field Date với định dạng bang đầu là yyyy-mm-dd H:M:S để thành các cột nhỏ hơn có thể ép về kiểu float mới có thể train mô hình máy học và loại bỏ cột name ,sno , symbol vì nó sẽ giữ định dạng string khong cho phép ép về float thì không cho phép train mô hình

## Đánh giá

### Uu điểm

Hoàn thành 1 số chức năng cơ bản về sắp xếp chọn loại

Chức năng xếp theo giá :

* List các coin có thể mua
* Trực quan hóa biểu đồ các loại coin đó theo nhãn mang tính tương quan

Graphical user interface

Description automatically generated

Hình 2 Trực quan hóa dữ liệu bằng các nhãn có tính tương quan

chọn loại coin và trực quan hóa theo thời gian móc từ ngày cuối cùng thu thập của bộ dữ liệu

ở đây là thông số của đồng Aave

Graphical user interface, chart, line chart

Description automatically generated

Hình 3 thông số toàn thời gian

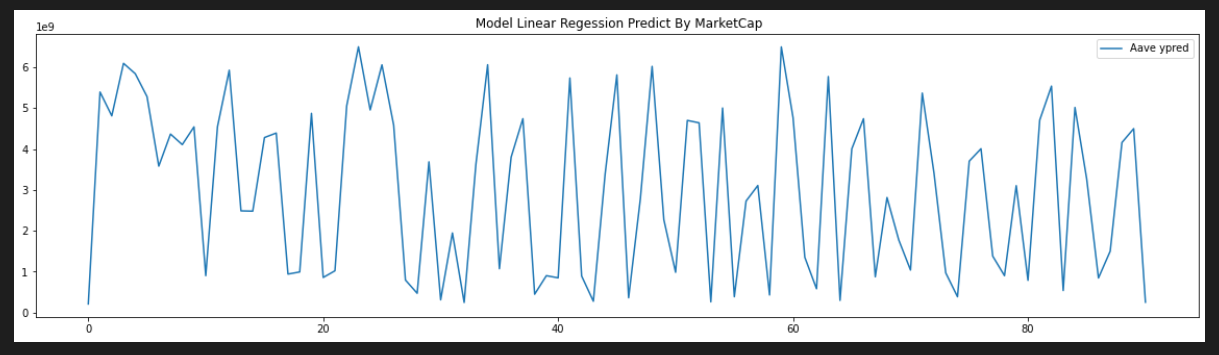
Chart

Description automatically generated

Hình 4 thông số trong 7 ngày

Dự đoán tương lại , dự đoán dựa trên Marketcap(ở đây chọn coin Aave)

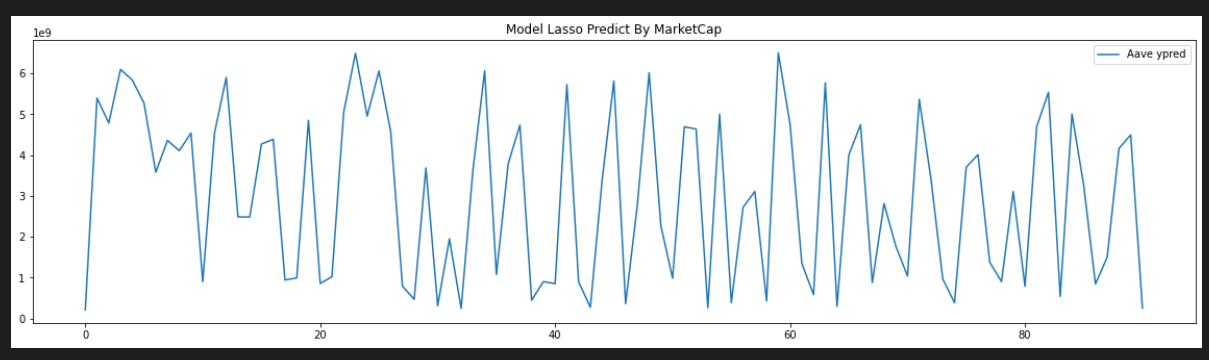
* Linear Regession



Hình 5 Linear Regession Predict By Marketcap

Độ chính xác : 0.9992740314803963

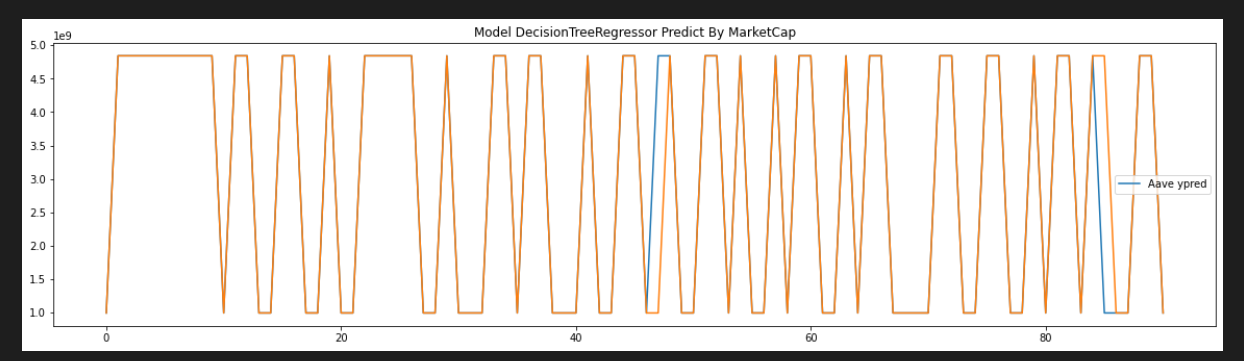
* Lasso



Hình 6 Lasso Predict By Marketcap

Độ chính xác : 0.8179056112619283

* Decision Tree Regressior



Hình 7 Decision Tree Regressior Predict B y Marketcap

Độ chính xác : 0.8286966513908162

### Khuyết điểm

Sắp xếp

* Do chưa thể build application nên chưa có giao diện chỉnh chu khó nhìn

Forecast

* tổ chức dữ liệu chưa tốt phải call từng tên biến

# KẾT LUẬN

## Uu điểm

có thể sử dụng được các chức năng cơ bản

* giống các ứng dụng phân tích kinh tế ngoài thị trường như sắp xếp mã coin theo giá , loại coin , thể hiện biểu đồ trực quan hóa dữ liệu trong 7 30 90 180 360 ngày
* cỏ thể dự đoán tương lai bằng cách train mô hình máy học test dựa trên Marketcap field

## Khuyết điểm

Chưa thể build thành 1 app hoàn chỉnh

Sử dụng dữ liệu chưa hợp lý (đặt từng biến cho từng dataframe gây dài dòng)

**(Không tính số trang tài liệu tham khảo)**

**TÀI LIỆU THAM KHẢO**

[Cryptocurrency Historical Prices | Kaggle](https://www.kaggle.com/datasets/sudalairajkumar/cryptocurrencypricehistory/code?resource=download) – tác giả : sudalairajkumar – ngày tham khảo đầu tiên : 24/04/2022 – mục đích : thu thập bộ dữ liệu

[Stack Overflow - Where Developers Learn, Share, & Build Careers](https://stackoverflow.com/) – mục đích : giải quyết bug

[scikit-learn: machine learning in Python — scikit-learn 1.0.2 documentation](https://scikit-learn.org/stable/) – mục đích : tìm hiểu mô hình máy học

[all\_currency\_predict\_in\_one | Kaggle](https://www.kaggle.com/code/aklimarimi/all-currency-predict-in-one) – tác giả: AKLIMA AKTER RIMI -ngày tham khảo: 05/05/2022 :-mục đích : tham khảo cách sử dụng model neural\_network

[Điểm danh top 5 phần mềm phân tích chứng khoán (cafef.vn)](https://cafef.vn/diem-danh-top-5-phan-mem-phan-tich-chung-khoan-20210521171807633.chn) – ngày tham khảo :28/01/2022 -mục đích : tham khảo app phân tích kinh tế

**PHỤ LỤC**

Contents

[1. GIỚI THIỆU 0](#_Toc103209052)

[2. NỘI DUNG 0](#_Toc103209053)

[2.1. Quy trình áp dụng xây dựng hệ thống 0](#_Toc103209054)

[2.1.1. Dataset: 0](#_Toc103209055)

[2.1.2. OOP Diagram 1](#_Toc103209056)

[2.2. Thiết lập 2](#_Toc103209057)

[3. KẾT LUẬN 5](#_Toc103209058)

[3.1. Uu điểm 5](#_Toc103209059)

[3.2. Khuyết điểm 5](#_Toc103209060)

[PHỤ LỤC PHÂN CÔNG NHIỆM VỤ 2](#_Toc103209061)

Text

Description automatically generatedText

Description automatically generated

Hình 8 Các thư viện được sử dụng trong quá trình phát triển

Hình 9 Khử Field



Hình 10 Tổ chức dữ liệu chưa hiệu quả

**PHỤ LỤC PHÂN CÔNG NHIỆM VỤ**

|  |  |  |
| --- | --- | --- |
| **STT** | **Thành viên** | **Nhiệm vụ** |
| 1 | Mã Hải Nhật | Code , làm sơ đồ OOP |
| 2 | Đấu Vũ Hoàng | Làm báo cáo, slide |