

Đại học quốc gia Thành phố Hồ Chí Minh

Trường Đại học Công nghệ thông tin

Khoa công nghệ phần mềm

Dự đoán giá vàng

Giảng viên hướng dẫn

Nguyễn Công Hoan

Sinh viên thực hiện

Phạm Hùng Vỹ - 15521037

Trần Phú Vinh - 15521020

Đào Đức Huy - 15520295

Phan Hữu Chí - 15520065s

13/06/2019

Version: 1.0

Lời cảm ơn

Nhóm chúng em hoàn thành được tốt đồ án môn học này, không thể không nói đến công lao của thầy Nguyễn Công Hoan. Chúng em xin chân thành cảm ơn thầy đã tận tâm hướng dẫn chúng em. Bên cạnh đó, nhóm cũng xin gửi lời cảm ơn chân thành đến các anh chị khóa trên, các bạn trong và ngoài lớp đã sẵn lòng chia sẻ tài liệu cũng như kinh nghiệm từng trải của bản thân để nhóm chúng em học tập và tránh mắc những sai lầm, tiết kiệm được thời gian trong quá trình thực hiện đồ án.

Tuy nhiên, do kiến thức và khả năng của chúng em còn nhiều hạn chế, do đó không tránh khỏi những thiếu sót, yếu kém. Chúng em rất mong nhận được những ý kiến đóng góp quý báu của thầy cô và các bạn học cùng lớp để đồ án được hoàn thiện hơn và rút ra được kinh nghiệm. Sau cùng, chúng em xin kính chúc quý thầy cô ở Khoa Công nghệ Phần mềm, đặc biệt là thầy Nguyễn Công Hoan thật dồi dào sức khỏe để tiếp tục thực hiện sứ mệnh cao đẹp của mình là truyền đạt kiến thức cho thế hệ mai sau. Nhóm xin chân thành cảm ơn!

Mục lục

Tổng quan

1.1 Giới thiệu

Dự đoán giá cổ phiếu là việc mà các nhà kinh tế thường xuyên bị hỏi. Tuy nhiên việc dự đoán chính xác là một việc khó có thể làm vì có rất nhiều yếu tố ảnh hưởng. Như cổ phiếu của một công ty chịu rất nhiều ảnh hưởng từ hoạt động công ty, chính phủ, quyết định của các nhà đầu tư. Một số cá nhân, tổ chức có ảnh hưởng cực kì lớn tới giá cổ phiếu. Với sự phát triển của internet thì thông tin được lan truyền rất nhanh, thông tin dường như gây ảnh hưởng đến giá cổ phiếu ngay lập tức

Đối với vàng thì thực sự không có một cá nhân tổ chức nào có thể nắm được quyền lực cụ thể. Có rất nhiều nhân tố, một dòng twitter của tổng thống Mỹ Donald Trump không thể nào có những ảnh hưởng đáng kể tới giá vàng như việc tuyên bố cấm cửa Huawei của ông.

Thời trước đã có nhiều phương pháp để dự đoán giá vàng bằng phương pháp xây dựng mô hình thuần thống kê như

- Autoregressive (AR)
- Moving Average (MA).
- Autoregressive Moving Average (ARMA)
- Autoregressive Integrated Moving Average (ARIMA)
- Seasonal Autoregressive Integrated Moving Average (SARIMA)

Những phương pháp này yêu cầu người dùng cần có rất nhiều kiến thức để tinh chỉnh tham số, đưa ra các dự đoán

Ngày nay, với sự phát triển của Machine Learning, có nhiều thuật toán mới tiến bộ, không yêu cầu người dùng phải có kiến thức rộng lớn như cây quyết định, học sâu.

Trong cây quyết định có thuật toán nổi tiếng như XGBoost, LGB. Những thuật toán yêu cầu người dùng có khả năng đưa ra nhiều feature nhất có thể để có thể đưa ra dự đoán chính xác

Học sâu có các mạng dòng họ Recurrent Neural Networks thường được sử dụng vì khả năng ghi nhớ thông tin trước đó. Phổ biến nhất là LSTM, GRU. Những thuật toán này không cần người dùng đưa ra những feature nhưng những mạng này thường yêu cầu dữ liệu rất nhiều và sức mạnh tính toán rất lớn để đưa ra dự đoán chính xác.

Cần đây, Facebook có đưa ra thuật toán Prophet. Sử dụng các tham số mặc định, Prophet đưa ra các dự đoán chính xác đánh ngạc nhiên mà không tốn nhiều tài nguyên để tính toán.

1.2 Đặt bài toán

Bài toán đặt ra cho biết thông tin giá vàng trên sàn chứng khoán hiện tại (GLD ETF). Làm sao để dự đoán được giá vàng đóng phiên của 90 ngày tiếp theo Dự đoán được giá vàng sẽ giúp cho nhà đầu tư đưa ra được quyết định mua vào bán ra ngắn hạn, tạo điều kiện để tích lũy số tiền. Còn việc dự đoán xa hơn 90 ngày thì rất khó có thể chính xác và ít có ý nghĩa. Như lời khuyên từ Jack Welch cựu CEO cựu General Electric, bạn nên để tất cả các trứng vào một giỏ và trông chừng nó. Hiếm có ai mà mua vàng để sinh lời mà lại không theo dõi giá cả hằng ngày.

1.3 Xây dựng mô hình

1.3.1 Dữ liệu

Được lấy từ package python `fix_yahoo_finance`. Gốc này lấy thông tin từ trang Yahoo Finance và chuyển thành Panda Dataframe. Bảng này gồm 7 cột:

- Date: ngày
- Open: giá lúc mở phiên
- Close: giá lúc đóng phiên
- High: giá cao nhất trong ngày
- Low: giá thấp nhất trong ngày
- Adj close: giá đóng phiên
- Volume: số lượng bán ra