Họ và tên: Trần Thái Toản

Lớp: 18IT1

MSV: 18IT042

Báo cáo tuần 1:

**Tìm hiểu 1 số ứng dụng thương mại điện tử mã nguồn mở**

Danh sách các phần mềm tham khảo:

* Shopee
* Lazada
* Sendo
* Amazon

**Shopee**

Sàn thương mại điện tử Shopee được đặt trụ tại Singapore, thuộc sở hữu của Tập đoàn SEA. Hiện tại, app bán hàng Shopee đã có hơn 40 triệu lượt tải và có mặt trên 7 quốc gia: Singapore, Thái Lan, Malaysia, Indonesia, Việt Nam, Đài Loan, Brazil.

App bán hàng Shopee có tính xã hội, cộng đồng và tương tác cao, cho phép người mua và người bán có thể trò chuyện trực tiếp với nhau. Và tính năng hashtag cho phép người dùng dễ dàng cập nhật và khám phá các xu hướng mới nhất cũng như các khuyến mãi hot

Shopee là một trong những trang ứng dụng bán hàng online, cho phép người bán đăng tải sản phẩm miễn phí và có nhiều tính năng hỗ trợ bán hàng tốt.

* Điểm nổi bật của ứng dụng Shopee:
  + Tìm kiếm dễ dàng theo tên Ngành hàng và Thương hiệu.
  + Nhận được nhiều ưu đãi độc quyền và thông báo khuyến mãi hàng ngày.
  + Mục Gợi ý hôm nay được điều chỉnh phù hợp với nhu cầu của từng người dùng.
  + Xem được các đánh giá về người bán.
  + Theo dõi xuyên suốt quá trình xử lý đơn hàng.

**Lazada**

Lazada là một sàn giao dịch thương mại điện tử lớn không chỉ tại Việt Nam mà còn cả Đông Nam Á. Ra nhập thị trường Việt Nam vào đầu năm 2012, Lazada Việt Nam là thành viên của tập đoàn thương mại điện tử đa quốc gia Lazada Group (thuộc sở hữu Tập đoàn Alibaba của tỷ phú Jack Ma). Hiện nay, App bán hàng online Lazada đã có mặt tại Indonesia, Malaysia, Philippines, Singapore, Thái Lan và Việt Nam.

* Điểm nổi bật của ứng dụng Lazada:
  + Thông báo độc quyền trên App Mobile các khuyến mãi và ưu đãi hàng ngày.
  + Gợi ý lựa chọn sản phẩm, ngành hàng, nhãn hàng cá nhân hóa cho người dùng.
  + Tính năng lọc và tìm kiếm sản phẩm mở rộng theo ngành hàng, nhãn hàng, giá, xếp hạng, màu sắc và nhiều đặc tính khác.
  + Nhận xét đánh giá và xếp hạng từ người dùng.
  + Hỏi đáp tương tác trực tiếp giữa người mua và người bán.
  + Theo dõi đơn hàng từ khi đặt hàng đến nhận hàng.

**Sendo**

Sendo.vn – Siêu chợ Sen Đỏ - Một dịch vụ chất lượng của Tập đoàn FPT. Với phương châm “Trăm người bán – Vạn người mua”, sàn giao dịch thương mại điện tử Sendo đã cung cấp đến hàng ngàn người dùng dịch vụ mua bán đảm bảo, chất lượng. Trong vòng nhiều năm trở lại đây, Sendo phát triển khá mạnh mẽ. Sendo từng đạt doanh thu 14,4% trên tổng số doanh thu của các nhóm ngành Thương mại điện tử trong năm 2014 (chỉ đứng sau Lazada với 36,1%.

* Ưu điểm của ứng dụng bán hàng online Sendo
  + Trực thuộc bởi tập đoàn FPT vậy nên Sendo đủ tự tin là một sàn thương mại điện tử uy tín và đáng tin cậy
  + Sendo cho phép khách hàng tạo gian hàng hoàn toàn miễn phí
  + Phát triển nhất đang là ngành hàng về thời trang, công nghệ
  + Độ bảo mật thông tin khách hàng tốt

**Tìm hiểu về hệ thống dự đoán giá hoặc doanh thu và tìm kiếm bộ dữ liệu phù hợp**

**link tham khảo:**

[**https://github.com/mgangrade7/Grocery-Sales-Forecasting/blob/master/Portfolio/Portfolio\_Grocery%20Sales%20Forecasting.ipynb**](https://github.com/mgangrade7/Grocery-Sales-Forecasting/blob/master/Portfolio/Portfolio_Grocery%20Sales%20Forecasting.ipynb)

Dự báo doanh thu là một khía cạnh chính của quản lý mua hàng.Dự báo rất quan trọng trong việc xác định mức dự trữ tồn kho và ước tính chính xác nhu cầu hàng hóa trong tương lai là một thách thức không ngừng. Nếu hàng hóa sẵn có nhiều hơn nhu cầu thì lợi nhuận tổng thể có thể bị ảnh hưởng. Do đó dự báo doanh thu hàng hóa rất quan trọng có thể có ý nghĩa để giảm thiểu tổn thất.

Trong bài này, người ta để cố gắng dự báo doanh số sản phẩm dựa trên các mặt hàng và một số biến phụ thuộc khác như ngày lễ.

Người ta sử dụng rất nhiều mô hình hồi quy để dự báo doanh số bán hàng khác nhau như hồi quy tuyến tính, cây quyết định,... Hơn nửa để tối ưu kết quả còn sử dụng multilayer perception và LightGBM.

Dữ liệu có hình dạng của nhiều tập tin. Đầu tiên, dữ liệu đào tạo (Train.csv) về cơ bản chứa doanh số theo ngày, cửa hàng và mặt hàng. Dữ liệu thử nghiệm (Test.csv) chứa các tính năng tương tự mà không có thông tin bán hàng mà chúng tôi được giao nhiệm vụ dự đoán. Đào tạo và thử nghiệm dựa theo ngày.

**Các thư viện được sử dụng:**

* **Numpy**: Numpy là một thư viện cho ngôn ngữ lập trình Python, thêm hỗ trợ cho các mảng và ma trận lớn, đa chiều, cùng với một bộ sưu tập lớn các chức năng toán học cấp cao để hoạt động trên các mảng này.
* **Pandas**: Pandas là một thư viện được viết cho ngôn ngữ lập trình Python để thao tác và phân tích dữ liệu. Cụ thể, nó cung cấp các cấu trúc dữ liệu và hoạt động để thao tác các bảng số và chuỗi thời gian.
* **Matplotlib**: Nó là một thư viện vẽ đồ thị rất mạnh mẽ hữu ích cho những người làm việc với Python và NumPy.
* **Scikit-learn:** Scikit-learn (Sklearn) là thư viện mạnh mẽ nhất dành cho các thuật toán học máy được viết trên ngôn ngữ Python.
* **XgBoost:** XGBoost là một thuật toán rất mạnh mẽ, tối ưu hóa về tốc độ và hiệu năng cho việc xây dựng các mô hình dự đoán.
* **Keras:** Keras được coi là một thư viện ‘high-level’ với phần ‘low-level’ (còn được gọi là backend) có thể là TensorFlow, CNTK, hoặc Theano.
* **LightGBM:** LightGBM là một khung tăng cường độ dốc sử dụng các thuật toán học tập dựa trên cây

**Bộ dữ liệu:**

**Trong bài này bộ dữ liệu chính bao gồm dữ liệu bán theo ngày, store và item information, cho dù mặt hàng đó đang được quảng bá, cũng như doanh số đơn vị. Dữ liệu kéo dài 5 năm từ 2013 đến 2017.**

**Bộ dữ liệu:**

[**https://www.kaggle.com/c/favorita-grocery-sales-forecasting/data**](https://www.kaggle.com/c/favorita-grocery-sales-forecasting/data)

**Bộ dữ liệu có 8 bảng khác nhau:** train.csv, test.csv, sample\_submission.csv, stores.csv, items.csv, transactions.csv, oil.csv, holidays\_events.csv.

