



# **XÂY DỰNG MÔ HÌNH HỒI QUY: ĐÁNH GIÁ KHẢ NĂNG BÁN ĐƯỢC MẶT HÀNG THỜI TRANG THỂ THAO.**

**Thành viên: Phạm Minh Tuấn\*, Nguyễn Trần  
Lâm, Nguyễn Năng Anhh**

# Mục tiêu

Trong kinh doanh thời trang **thể thao**, việc đánh giá khả năng bán hàng là quan trọng. Mô hình hồi quy giúp **dự đoán** trên bộ dữ liệu '**Adidas and Nike product**'

Dự án tập trung vào việc phát triển một **mô hình hồi quy** để đánh giá và dự đoán **khả năng bán được hàng** của sản phẩm. Bằng cách sử dụng dữ liệu chi tiết về sản phẩm, mục tiêu là xây dựng một công cụ linh hoạt giúp doanh nghiệp hiểu rõ hơn về xu hướng tiêu thụ và tối ưu hóa chiến lược kinh doanh.



# TIỀN XỬ LÝ DỮ LIỆU

- Kiểm Tra Phân Phối Dữ Liệu
- Chuẩn Hóa Dữ Liệu
- Giữ Lại Các Đặc Trưng Quan Trọng
- Kiểm Tra và Loại Bỏ Ngoại Lệ



## TRÍCH SUẤT ĐẶC TRƯNG

- **'description'**: mô tả để đánh giá độ chi tiết hoặc phong cách mô tả của sản phẩm.
- **'images'**: Sử dụng kỹ thuật xử lý ảnh để trích xuất thông tin như màu sắc chủ đạo, kiểu dáng.
- **'price'** các khoảng giá hoặc tính toán giá trung bình cho từng thương hiệu. Giúp nhận biết mức độ cao hoặc thấp cấp độ giá.
- **'rating'**: Tính xác suất đánh giá tích cực hoặc tạo ra đặc trưng liên quan đến tỉ lệ đánh giá tích cực so với tổng số đánh giá.

Tiếp tục mở rộng thêm các đặc trưng như **'color'** và **'brand'**

# Mô hình hồi quy

Lựa chọn mô hình:

- Linear Regression
- Decision Trees
- Random Forests
- Neural Network
- Support Vector Machines (SVM)
- ...

# ĐÁNH GIÁ MÔ HÌNH: HIỂU RÕ VÀ ĐO LƯỜNG HIỆU SUẤT

Đo Lường Hiệu Suất:

- MSE, R-Square, MAE
- Confusion matrix
- ROC Curve và AUC
- Cross- Validation



## Hướng Phát Triển

- Tham chiếu đến quyết định kinh doanh sử dụng kết quả từ mô hình để hỗ trợ quyết định tài chính, marketing, hoặc chiến lược sản phẩm.
- Phát hiện xu hướng của các sản phẩm đã thành công để hiểu rõ xu hướng thời trang không chỉ bao gồm việc theo dõi các mốt mới mẻ mà còn bao gồm việc tìm hiểu sở thích, phong cách sống của đối tượng mục tiêu



# Hướng Phát Triển

Danh sách những sản phẩm thêm vào		Thông số đánh giá
Bảng tùy chỉnh các trường dữ liệu đưa vào	<b>NÚT BẮT ĐẦU ĐÁNH GIÁ</b>	
Bảng thông báo lỗi		Danh sách những sản phẩm tiềm năng

Thiết kế cơ bản khi triển khai sản phẩm





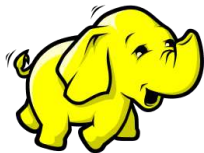
# CÔNG NGHỆ



NumPy



scikit  
*learn*



pandas

Mojo



seaborn



TensorFlow

**Thanks!**

# DỰ ĐOÁN XU HƯỚNG TIÊU THỤ NHÓM SẢN PHẨM TẠI CỬA HÀNG THEO KHU VỰC

Đội: FinTea

Thành viên: *Phạm Minh Tuấn\*, Nguyễn Trần Lâm, Nguyễn Năng Anh*



# NỘI DUNG

1. Lý do chọn đề tài
2. Phương hướng triển khai

# VẤN ĐỀ ĐẶT RA

Việc biết trước xu hướng tiêu thụ sản phẩm của thị trường giúp được gì cho doanh nghiệp?

Nếu một doanh nghiệp trả lời được câu hỏi trên và xác định được trước xu hướng tiêu thụ tăng hoặc giảm sẽ giúp cho doanh nghiệp **có chiến lược kinh doanh phù hợp trong tương lai**. Điều này đồng nghĩa với việc giúp doanh nghiệp **luôn bước nhanh hơn so với thị trường**.

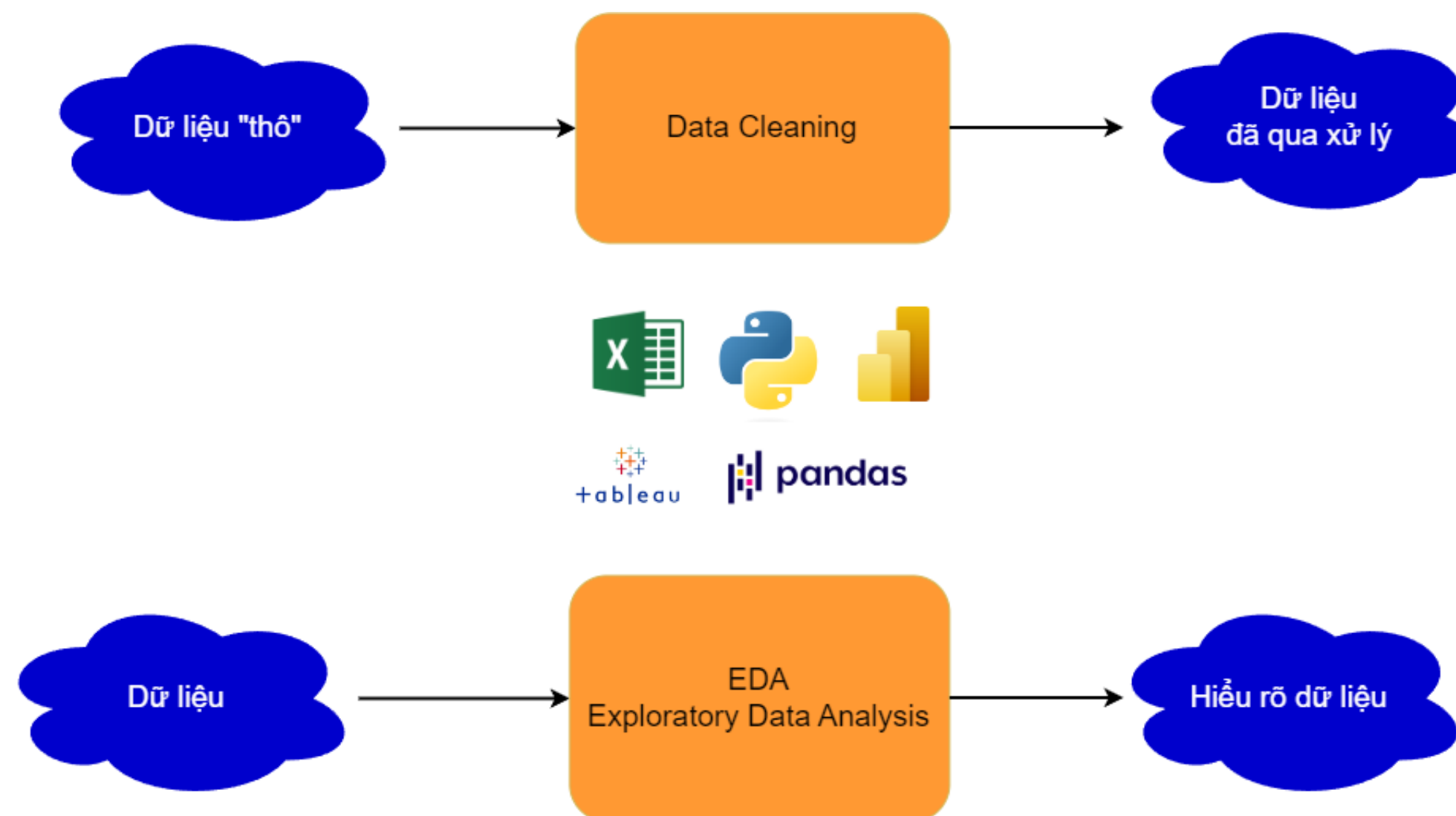
Theo một bài nghiên cứu của McKinsey có tên “**Prediction: The future of customer experience**”. Cho thấy, *các doanh nghiệp tại Mỹ sử dụng mô hình dự đoán có thể tăng cường hiệu suất tài chính lên đến 20% trong năm 2021 chủ yếu thông qua việc tối ưu hóa chiến lược kinh doanh.*

# ROAD MAP



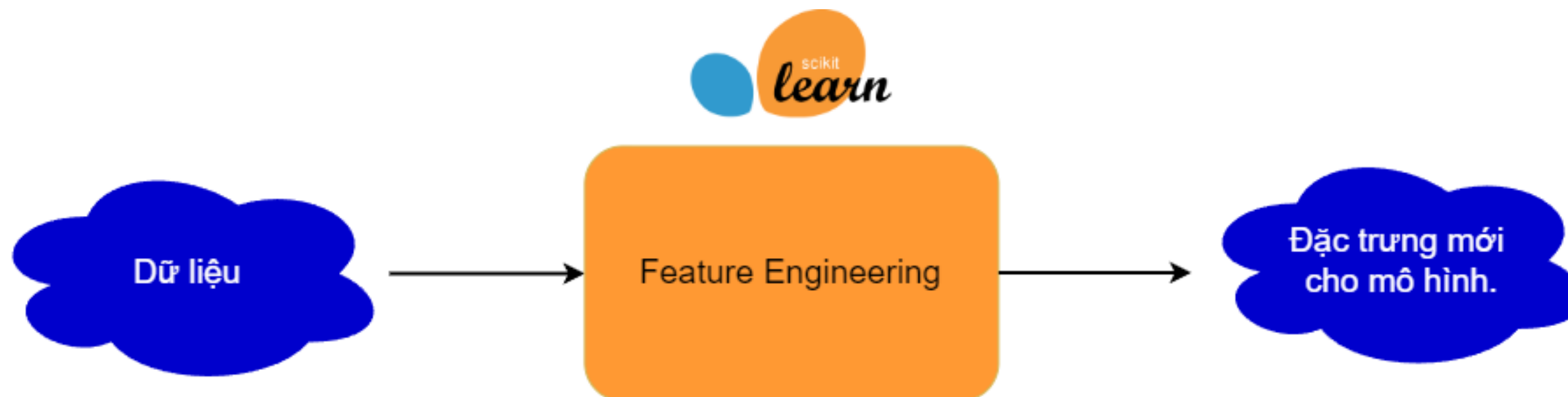
# DATA CLEANING & EDA

Trong ý tưởng xây dựng mô hình này, chúng tôi sử dụng bộ dữ liệu số 2 “Sales and inventory Data of Vietnam Retails”



# FEATURE ENGINEERING

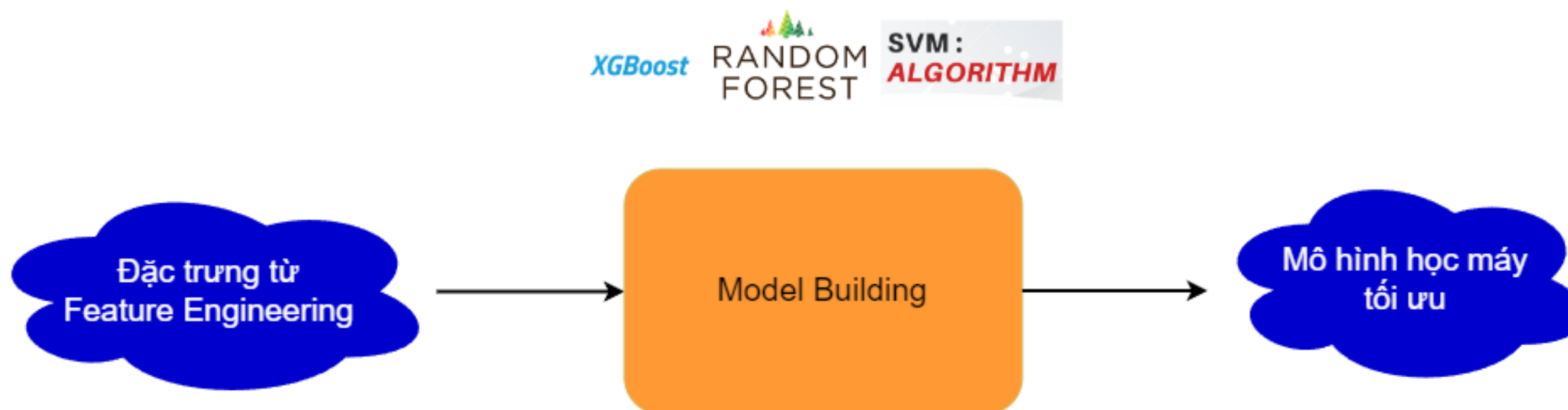
Tạo và chọn đặc trưng mới từ dữ liệu.





# MODEL BUILDING

Xây dựng và lựa chọn mô hình phù hợp

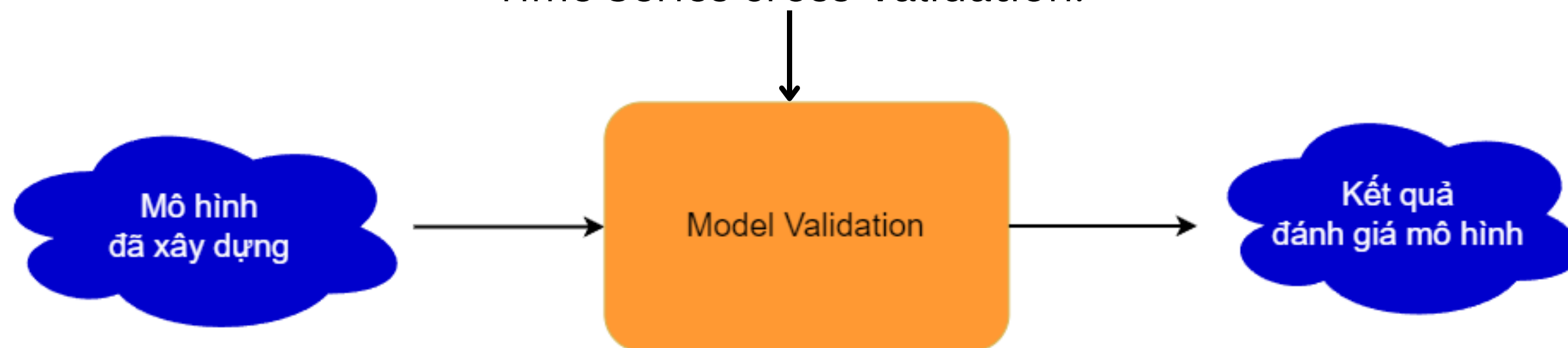


# MODEL VALIDATION

Xây dựng và lựa chọn mô hình phù hợp

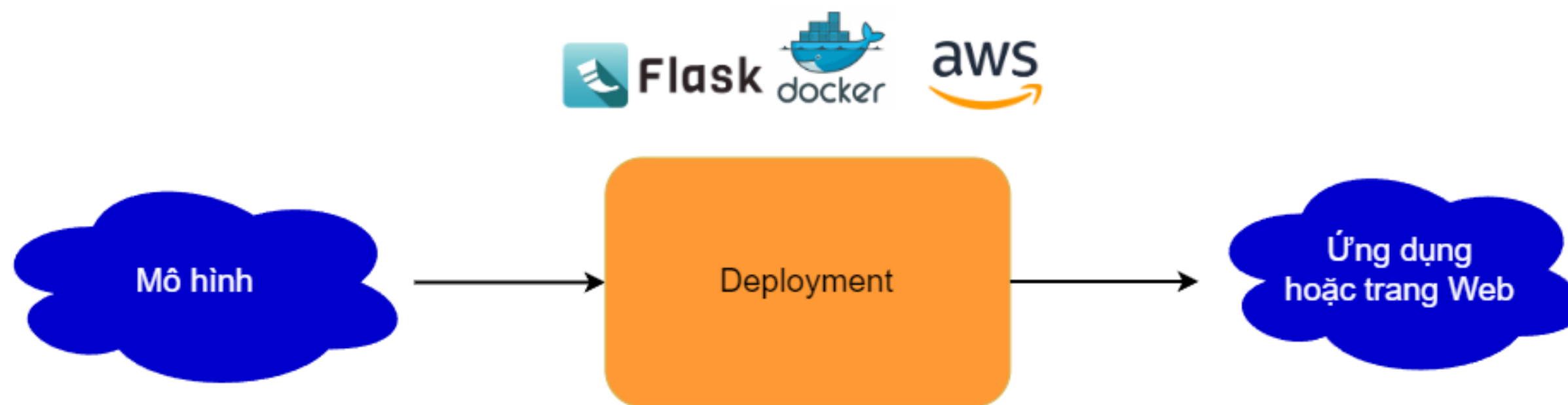
Cross-validation trong Scikit-learn

Time Series cross-validation.



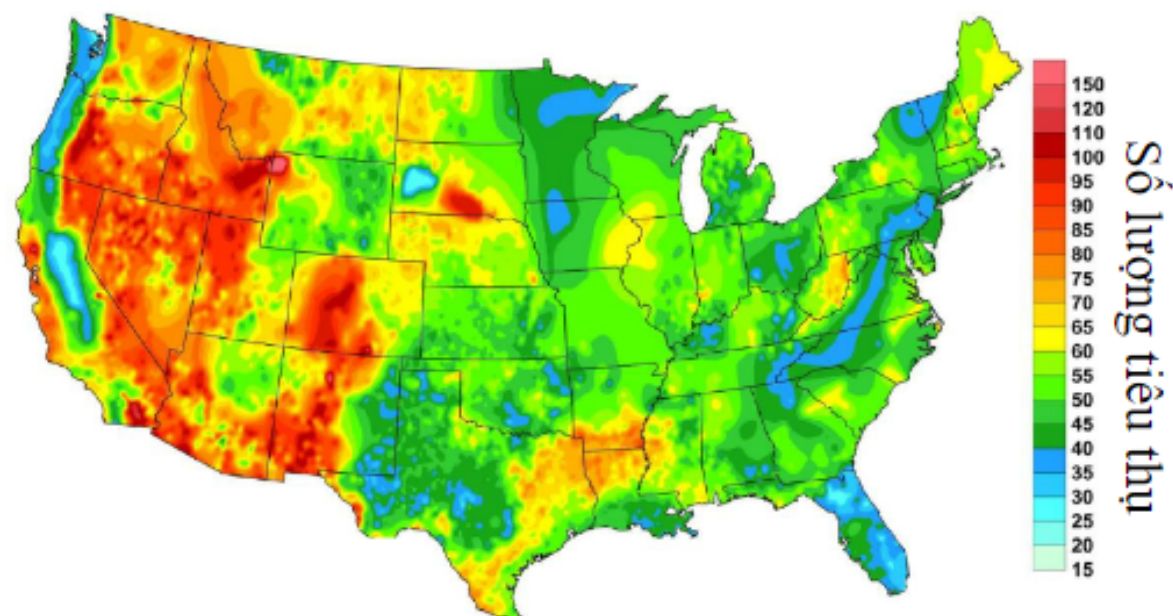
# DEPLOYMENT

Triển khai mô hình vào sản phẩm thực tế

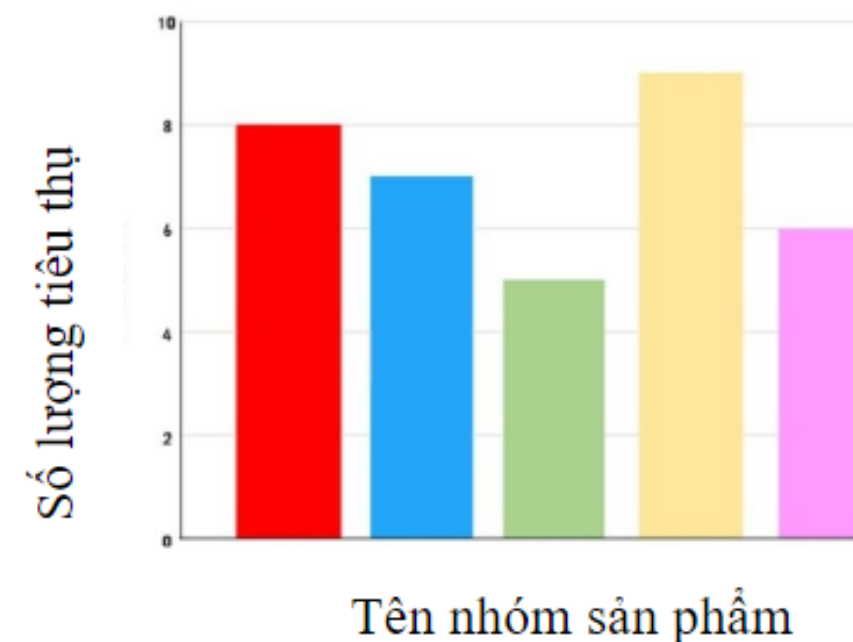


# MÔ TẢ OUTPUT DEMO

Trang dự đoán xu hướng tiêu thụ nhóm sản phẩm tại cửa hàng theo khu vực



*Số lượng tiêu thụ theo khu vực*



*Số lượng tiêu thụ theo nhóm sản phẩm*

Khoảng thời gian

Tháng

Năm

Tải liên file dữ liệu

**THANKS FOR  
WATCHING**



*Ý tưởng: Nhận diện sản phẩm  
tiềm năng trong lúc khách hàng  
lựa chọn bằng Camera quan sát*

Thành viên: Phạm Minh Tuấn\*, Nguyễn Trần Lâm,  
Nguyễn Năng Anh

**VIETNAM  
DATATHON  
2023**



# Ý tưởng giúp cải thiện vấn đề gì?

- Vấn đề xuất phát từ các số liệu sản phẩm bán hàng chỉ phản ánh các **sản phẩm khách hàng đã thanh toán**
  - => Chưa phản ánh được các **sản phẩm khách hàng đã quan tâm** trong quá trình chọn lựa
- Ý tưởng tận dụng dataset 5: Shelf behaviors CCTV data
  - Cụ thể, Hành động loại 4: Kiểm tra sản phẩm khi giữ trên tay



# Ý tưởng giúp cải thiện vấn đề gì?

- **Phân tích hành động** kiểm tra sản phẩm của khách hàng
  - => Đây là một **sản phẩm** khách hàng quan tâm trong lúc lựa chọn
    - => Một **sản phẩm** tiềm năng cần được **nhận diện** để cửa hàng tiện lợi cung cấp đúng nhu cầu sản phẩm cho khách hàng
      - => Ngoài việc chỉ xét những sản phẩm đã thanh toán, nay còn **bổ sung thêm những sản phẩm** khách hàng quan tâm trong lúc lựa chọn
        - => Cải thiện vấn đề **“Đáp ứng đúng sản phẩm khách hàng cần”** trong cửa hàng tiện lợi

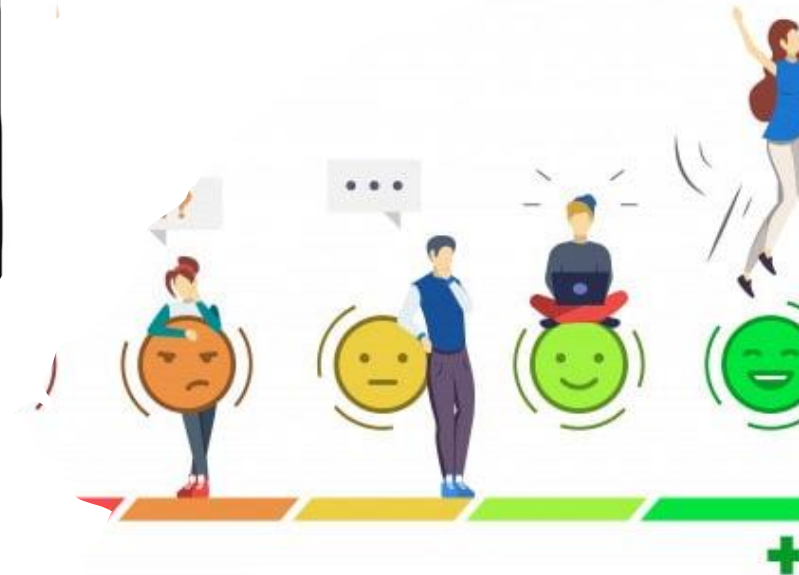
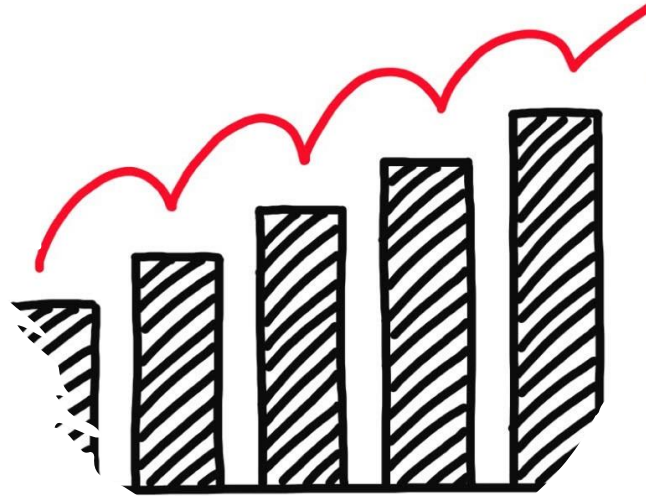




# “Đáp ứng đúng sản phẩm khách hàng cần” có thực sự quan trọng?

- Điều này sẽ giúp:
  - **Tăng tính hài lòng** của khách hàng khi mua ở cửa hàng tiện lợi
  - **Tối ưu** hóa doanh số sản phẩm bán được
  - **Giảm thiểu** Hụt và Chết sản phẩm

SALES





## Cách nào để cải thiện “Đáp ứng đúng sản phẩm khách hàng cần”?

- Ứng dụng trí tuệ nhân tạo để giải quyết bài toán **“Nhận diện sản phẩm tiềm năng trong lúc khách hàng lựa chọn bằng Camera quan sát”** => Sản phẩm tiềm năng => Xác định tập sản phẩm khách hàng cần => cải thiện **“Đáp ứng đúng sản phẩm khách hàng cần”**
- **Mô tả đầu vào:**
  - Các khung hình được gán nhãn hành động loại 4 trong bộ dữ liệu Shelf behaviors CCTV data
  - Các khung hình này còn sẽ được gán bounding box bao quanh sản phẩm trong lúc khách hàng đang giữ trên tay

Cách nào để cải thiện  
“Đáp ứng đúng sản  
phẩm khách hàng cần”?

---

- **Mô tả đầu ra sản phẩm demo tối thiểu:**

### Trang nhận diện sản phẩm tiềm năng

Nút thêm  
video

Video Camera quan  
sát quay lại quá trình  
khách hàng đang chọn  
lựa sản phẩm

Nơi hiển thị video đã nhận diện  
được sản phẩm bằng bounding  
box bao quanh

Liệt kê các sản phẩm trong quá  
trình khách hàng lựa chọn. Hiển  
thị các thông tin sản phẩm (*Hình  
ảnh sản phẩm - Phút bắt đầu giữ  
sản phẩm - Phút kết thúc giữ sản  
phẩm - Tổng thời gian giữ sản  
phẩm*)

# Cách nào để cải thiện “Đáp ứng đúng sản phẩm khách hàng cần”?

- Các bước thực hiện từ đầu vào đến đầu ra:
  - 1. Preprocessing:
    - Các vấn đề cần xử lý: resizing, normalizing, augmenting videos (tập trung các hành động type 3, 4 trong bộ dữ liệu)
    - Công cụ/ tài nguyên có sẵn: OpenCV, Pillow
  - 2. Annotation:
    - Các vấn đề cần xử lý: sẽ cần đánh dấu thêm sản phẩm khách hàng lấy là sản phẩm nào ? Vì dữ liệu chỉ tập trung đánh dấu các hành động cho khách hàng (không bao gồm sản phẩm)
    - Công cụ/ tài nguyên có sẵn: Labelbox, Labelimg, VGG Image Annotator
  - 3. Feature Extraction:
    - Các vấn đề cần xử lý: cần trích xuất đặc trưng từ các frame trong video, có thể cần lấy thông tin chuyển động dựa theo sử dụng 'optical flow'
    - Công cụ/ tài nguyên có sẵn: TensorFlow, Pytorch, Pre-trained models (ResNet, Inception, 3D CNNs model zoos)

# Cách nào để cải thiện “Đáp ứng đúng sản phẩm khách hàng cần”?

- Các bước thực hiện từ đầu vào đến đầu ra:
  - 4. Model Selection:
    - Các vấn đề cần xử lý: Chọn model phù hợp để nhận diện hành động định mua sản phẩm/ có hành vi quan tâm đến 1 sản phẩm cụ thể
    - Công cụ/ tài nguyên có sẵn: một số cấu trúc model có thể phù hợp cho video (RNNs, LSTMs, 3D CNNs, Attention)
  - 5. Training & Validation :
    - Các vấn đề cần xử lý: so sánh hành động dự đoán với `Type` hành động, tối ưu hàm loss cho tập train + validation
    - Công cụ/ tài nguyên có sẵn: TensorFlow, Pytorch, GPU Kaggle
  - 6. Fine-Tuning:
    - Các vấn đề cần xử lý: Đánh giá kết quả và thay đổi tham số điều chỉnh để cải thiện kết quả
    - Công cụ/ tài nguyên có sẵn: Scikit-learn, GridSearchCV, RandomizedSearchCV, Wandb
  - 7. Testing:
    - Các vấn đề cần xử lý: Đánh giá kết quả dự đoán trên tập test bao gồm: thông số đánh giá (accuracy, precision, recall, and F1), quan sát kết quả thật khi đã dự đoán.
    - Công cụ/ tài nguyên có sẵn: Scikit-learn, Wandb
  - 8. Deployment:
    - Các vấn đề cần xử lý: Đưa model lên server để dự đoán real time video (có thể hiển thị kết quả bằng demo)
    - Công cụ/ tài nguyên có sẵn: Docker, FastAPI, Streamlit

THANK FOR  
WATCHING

