Project Proposal Team 85

VIETNAM DATATHON 2023

Table of contents

01

Project overview

Context

Brands are concerned about performance and profit potential to maximize returns.

Challenge

Determining effective sales strategy to optimize revenue.

Solution

Predicting product sales to help brands gain insights into product potential revenue and develop impactful sales strategies.

Overview

Dataset

Sales and Inventory Data of Vietnamese Retailers

Problem

Predict the quantity of products that can be sold

Solution

Use linear regression to predict the quantity of products that can be sold based on corresponding attributes, especially the selling price and time.

Question: Why predict the quantity of products sold instead of other factors like price or revenue forecasting?

Answer:

- Predicting the sales quantities of products is the initial step towards adjusting prices and optimizing revenue.
- Furthermore, it also helps enhance inventory management for better control and reduced stock imbalances.

02

Building MVP

Development Phases

1. Requirement

Specify the functional and non-functional requirements

2. Get Data

Acquire and preprocess data

3. Train Model

Develop a regression model to predict the product sales

4. Test Data

Evaluate the model's performance

5. Deploy

Create API endpoints and front-end application

Phase 1: Requirements Specification

Functional

User inputs:

- Specific product.
- Price setting.
- Sales time or period.

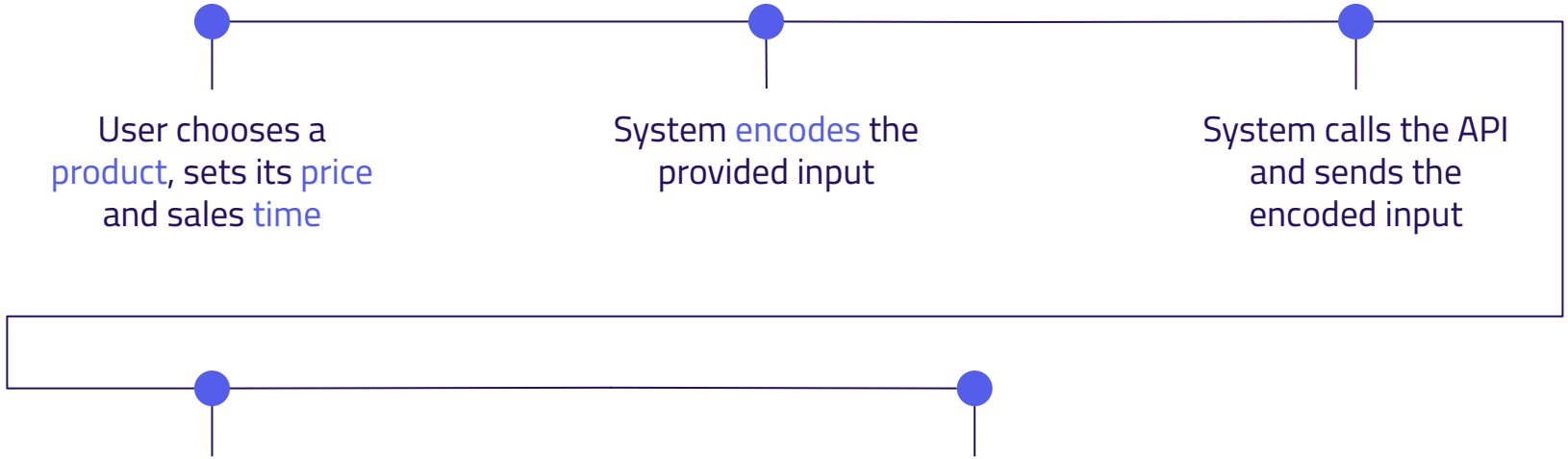
System Output:

- Predicted sold quantity of the product based on the provided inputs.

Non-functional

Accuracy: assessed using the Mean Absolute Percentage Error (MAPE) metric. Initially targeting a 10% error rate, gradually enhance to achieve 5% and 1%.

Process demo



Phase 2: Get Data

Data Acquisition

Sales and Inventory Data of Vietnamese Retailers

Data Cleaning

Handle missing values, outliers, and inconsistencies.

Dimensionality Reduction

Identify and retain the most relevant features from the product characteristics.

Normalization

Normalize numerical features (like price) and encode categorical variables (like product categories) into numerical representations.

Phase 3: Train Linear Regression Model

Input features:

Product Information	Product ID and the details of product information such as category, color and size.
Time	The period during which the product is expected to be sold.
Price	The price of the product during the given time period.

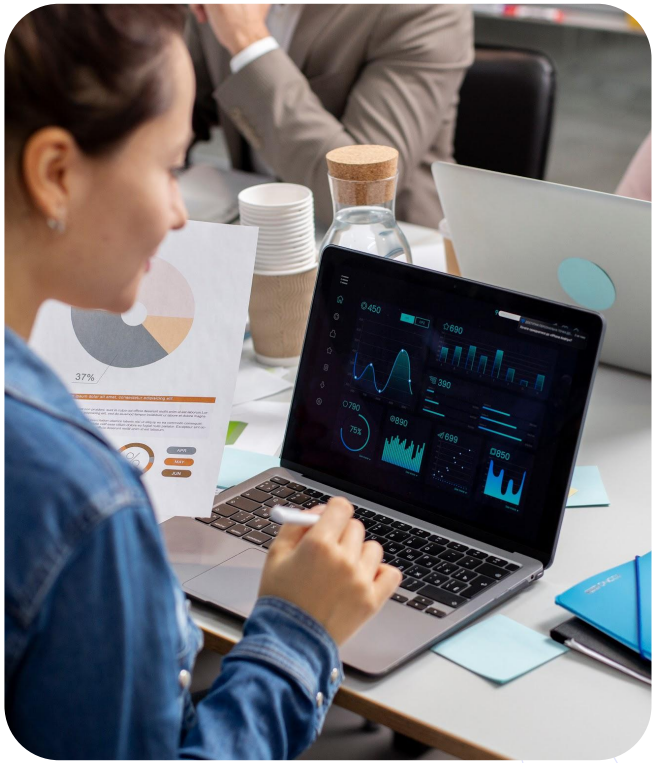
Target variable:

Quantity sold	The number of units of the product sold during the given time period at a particular price.

Phase 4: Test Data

Evaluate the performance of the model using test data to ensure the model to generalize well

Metric: Mean Absolute Percentage Error (MAPE)



Phase 5: Deploy

API

Build API endpoints to process input (price, time, product details) and provide predicted sales quantities.

Front-end app

Develop a user interface to gather the required information from users and display predictions received from the API.

03

Development Potential

Development Potentials

Core Feature Development

To swiftly develop our MVP and evaluate its functionality, we're focusing solely on these input features:

- product information,
- sales price
- time

Feature Expansion

As the model evolves, additional features can be incorporated to enhance its predictive capabilities, such as:

- store location
- distribution channel
- launch and sales season

Transition to Autoregressive

We transit the linear regression model to an autoregressive model to leverage historical data, allowing for real-time predictions, ensuring more accurate and responsive predictions in dynamic business environments.

04 Applications

Applications

- Inventory management
- Price optimization
- Sales forecasting
- Understand customer behavior and demand
- Optimize production based on predicted sales

Thanks!

Do you have any questions?
nmdat21@clc.fitus.edu.vn