AMDARI®

Data Science Project

Jewelry Price Optimization with ML

Data Science Project

Specialization: Demand Forecasting

Business Focus: jewelry industry

Tool: Python, MLFlow & Git

AMDARI®

Gain hands-on experience in building machine learning models to optimize pricing strategies.
Learn to use MLflow for model tracking and deployment in real-world scenarios.
Develop skills in data analysis, feature engineering, and understanding customer pricing behavior.
Understand the integration of pricing models into business workflows for actionable insights.

Learning Skill

- Data preprocessing and feature engineering.
- Building regression models and demand forecasting.
- Experiment tracking and deployment with MLflow.
- Collaborative coding and project management with Git.
- Interpreting and communicating model insights for business strategy refinement.

AMDARI®

Business Introduction

Company: Gemineye
Industry: luxury goods and jewelry industry

The focal company is a luxury jewelry retailer known for its craftsmanship, quality, and innovation. With a global presence, the brand aims to cater to a diverse clientele. The company has consistently leveraged technology to enhance customer experiences, offering online customization and seamless e-commerce options. However, its pricing strategies currently rely on manual adjustments, leading to inefficiencies in capturing optimal revenue.

Achievements:

- Leading in bespoke jewelry sales.
- Recognition for excellent customer service.

Unique Aspects:

- Seasonal pricing variations.
- Customization options influencing price variability.

AMDARI®

Business Problem

The company faces challenges in setting optimal prices for its jewelry products. Key issues include:

- Overpricing, which risks losing price-sensitive customers.
- Underpricing, which diminishes profit margins.
- Difficulty in dynamically adjusting prices based on market trends, customer preferences, and competitive actions.

Obstacles:

1. Inconsistent pricing strategies across regions and product lines.
2. Lack of a data-driven approach to predict demand elasticity.
3. High competition necessitating quick responses to market trends.

AMDARI®

Rationale for cohort analysis

Why Jewelry Price Optimization with ML?

1. **Maximized Revenue**: Data-driven pricing ensures optimal revenue generation by balancing volume and margins.
2. **Competitive Edge**: Dynamic pricing enables quick adaptation to market changes.
3. **Improved Customer Retention**: Personalized pricing strategies cater to different customer segments.
4. **Efficient Decision-Making**: Automating pricing decisions reduces reliance on manual intervention.
5. **Actionable Insights**: An ML-driven approach provides deeper insights into customer behavior and demand patterns.

AMDARI®

Why Jewelry Price Optimization with ML?

Rational for the Project:

01

Maximized Revenue:
Data-driven pricing ensures optimal revenue generation by balancing volume and margins.

02

Competitive Edge:
Dynamic pricing enables quick adaptation to market changes.

03

Improved Customer Retention:
Personalized pricing strategies cater to different customer segments.

04

Efficient Decision-Making:
Automating pricing decisions reduces reliance on manual intervention.

05

Actionable Insights:
An ML-driven approach provides deeper insights into customer behavior and demand patterns.

AMDARI®

The features contained in the dataset are:
- Order datetime: The time at which the order was placed
- Order ID: Identifiers for the different orders placed
- Purchased product ID: Identifiers for the different product ordered for
- Quantity of SKU in the order: Quantity of jewelry pieces ordered for
- Category ID: Identifier for the jewelry category
- Category alias: Name of jewelry category e.g. earring
- Brand ID: Identifier for jeweler brand
- Price in USD: Jewelry price in US Dollars
- User ID: Identifier for user/customer
- Product gender (for male/female) (Target gender for jewelry piece)
- Main Color: Overall color of jewelry piece
- Main metal: Main metal used for mounting
- Main gem: Main gem mounted on jewelry piece

AMDARI ®

Project Enhancement

To enhance the project deliverables, you are required to implement a Version Control system using git and set up experiment Tracking and Monitoring system using ml-flow

MLFlow: utilized by Data Scientist & Machine Learning Engineers to Track experiments, hyperparameter tuning, and model iterations. It also provides Logging which helps us in comparing results for reproducibility

Git : you are required to set a version control system using git that will enable you track version changes and code history from your work on github

Tech Stack

1. Programming language – Python
2. Git & Github

- Numpy: For performing mathematical operations over data
- Pandas: For Data Analysis and Manipulation
- Matplotlib.pyplot: For Data Visualization
- Seaborn: For Data Visualization
- Scikit-learn: For Machine Learning
- Mlflow: for experiment tracking and monitoring

AMDARI®

Data Science Project Scope

Ingest Data
Ingest the data and perform data cleaning

Exploratory Data Analysis
the data is thoroughly explored and analysed to gain insights and understand its characteristics. This involves statistical analysis, data visualization, and other exploratory techniques to identify patterns, correlations, anomalies, and potential issues in the data.

Feature Engineering
Feature engineering is the process of creating or selecting relevant features (input variables) from the available data that will be used for modeling. It may involve feature extraction, transformation, scaling, or the creation of new features to improve the predictive power of the models.

Model Development
Various modeling techniques are applied to the prepared data in this stage to build predictive or descriptive models. This can include machine learning algorithms, statistical models, or other analytical approaches. The models are trained, validated, and fine-tuned to optimize their performance. And Mlflow is utilized here to log and track model performance and artefacts

Model Evaluation and Selection
Once the models are developed, they are evaluated using appropriate evaluation metrics and validation techniques. This involves assessing their performance, generalizability, and robustness. The best-performing model(s) are selected for further deployment or refinement.

Reporting and Strategy Recommendations
Generate a good report and develop strategies and recommendations.