Practical Training Evaluation

Presentation 2023

High Radius (Highway To Highradius)

Presented By:
Punyapu Sai Teja

highradius

CERTIFICATE
OF COMPLETION

Product and Engineering

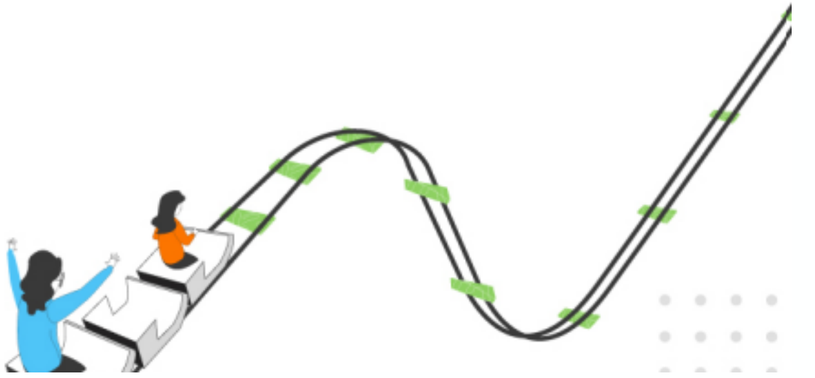
Product Essentials Program

This is to certify that **Punyapu Sai Teja** has successfully completed the Highway to HighRadius Internship Program from **31st May 2023** to **12th July 2023**, where he/she built and deployed on AI Enabled Fintech B2B Cloud Application.

During this project, he/she was involved in creating a full stack web-based product thereby developing a deep understanding of all aspects of product development such as identifying appropriate user requirements, designing a great user experience and building appropriate data models and machine learning models along with relevant UI components and backend design

Neha Srivastava

Neha Srivastava
AVP, People & Culture
HighRadius



Overview

1. Objective
2. B2B
3. Data Dictionary
4. Itinerary
5. Milestone 1
6. Milestone 2
7. Milestone 3
8. Milestone 4
9. Results
10. Conclusion

Obj...

Build a Machine Learning model to predict the order amount that customers can place in the upcoming days.

B2C World

- The B2B world operates differently from B2C or C2C.
- Businesses work with other businesses on credit.
- Accounts Receivable represents money owed by entities to the firm on the sale of products or services on credit.

Credit Check Department:

- Validates the customer and verifies the available credit limit.
- Collects purchase orders for an inventory check.
- Calculates exposure for a particular customer.
- Ensures smooth order inflow and cash inflow.

Data Dictionary for Order Amount Prediction Dataset :

Field Name	Description
CUSTOMER_ORDER_ID	Unique identifier for an order ID.
SALES_ORG	Unique identifier for the sales organization.
DISTRIBUTION_CHANNEL	Country where the shipment has been delivered.
DIVISION	Region coverage of the customer.
RELEASED_CREDIT_VALUE	Total credit value the customer possesses.
PURCHASE_ORDER_TYPE	Denotes the category of the Purchase Order.
COMPANY_CODE	Smallest organizational unit for complete accounting.
ORDER_CREATION_DATE	Date on which the order was created in the ERP.
ORDER_CREATION_TIME	Time on which the order was created in the ERP.
CREDIT_CONTROL_AREA	Organizational unit for customer credit limits.
SOLD_TO_PARTY	Unique identifier for the person/organization who placed the order.
ORDER_AMOUNT	Total sum of Purchase Prices in the Purchase Order(s).
REQUESTED_DELIVERY_DATE	Requested date of delivery by the customer.
ORDER_CURRENCY	Currency in which the order was billed and paid.
CREDIT_STATUS	Indicates the credit health of a particular customer.
CUSTOMER_NUMBER	Unique identifier for a specific customer.

Itinerary

Milestone 1
Data Sanity

- Sanitize data by handling missing/null/duplicate values

Milestone 2
Exploratory data analysis

- Visualizations will help in gaining insights and understanding the distribution, patterns, and outliers in the dataset.

Milestone 3
Feature Engineering and Selection

- Preprocess and engineer the data, ensuring it is suitable for building a robust machine learning model.

Milestone 4
ML Models and Evaluations

- Evaluate multiple ML models, compare performances, and select the best for accurate order amount prediction.

Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

Data Preprocessing...

Data Sanity using Numpy and Pandas.

Steps performed: Data cleaning, date format conversion, removing invalid records, handling special characters in order amounts, and currency conversion to USD.

Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

Exploratory Data Analysis...

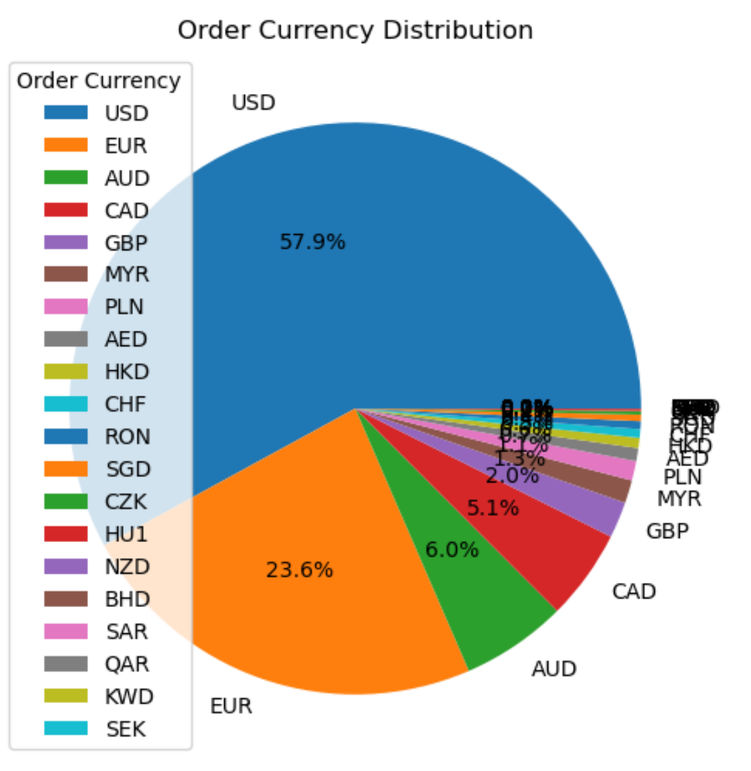
EDA to gain insights from the data.

Visualizations: Histogram on DISTRIBUTION_CHANNEL, Pie Chart on ORDER_CURRENCY, Line charts on PURCHASE_ORDER_TYPE and DISTRIBUTION_CHANNEL, Line plot on ORDER_CREATION_DATE and amount_in_usd, Boxplot on ORDER_AMOUNT, and Barchart on COMPANY_CODE and ORDER_AMOUNT.

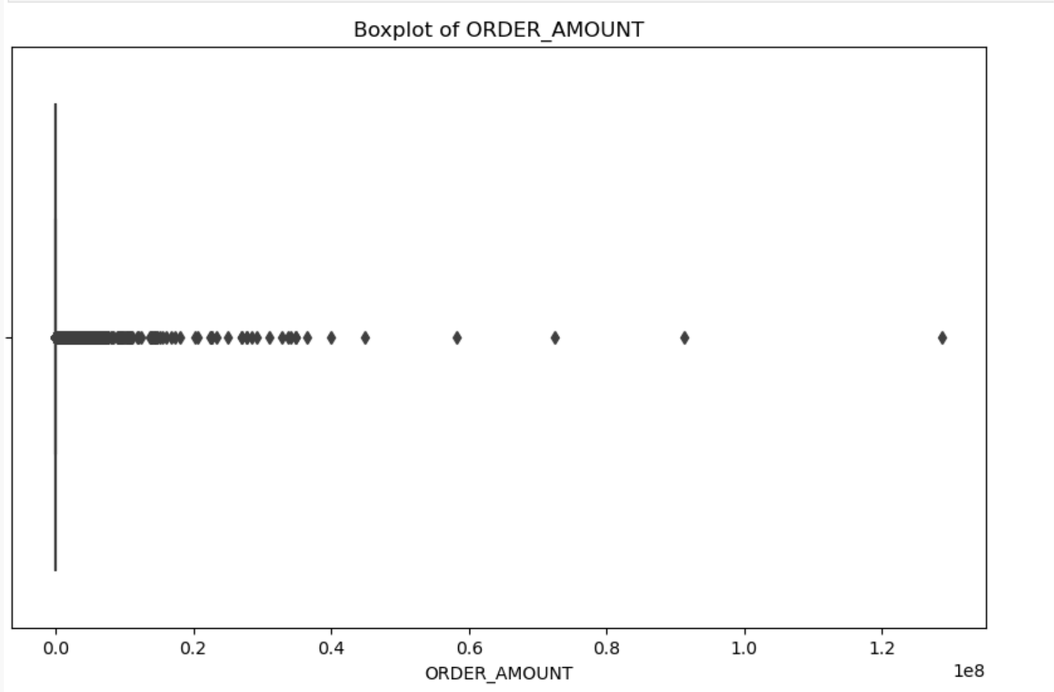
Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

Exploratory Data Analysis...

Pie Chart on ORDER_CURRENCY



Boxplot on ORDER_AMOUNT



Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

Feature Engineering and Selection...

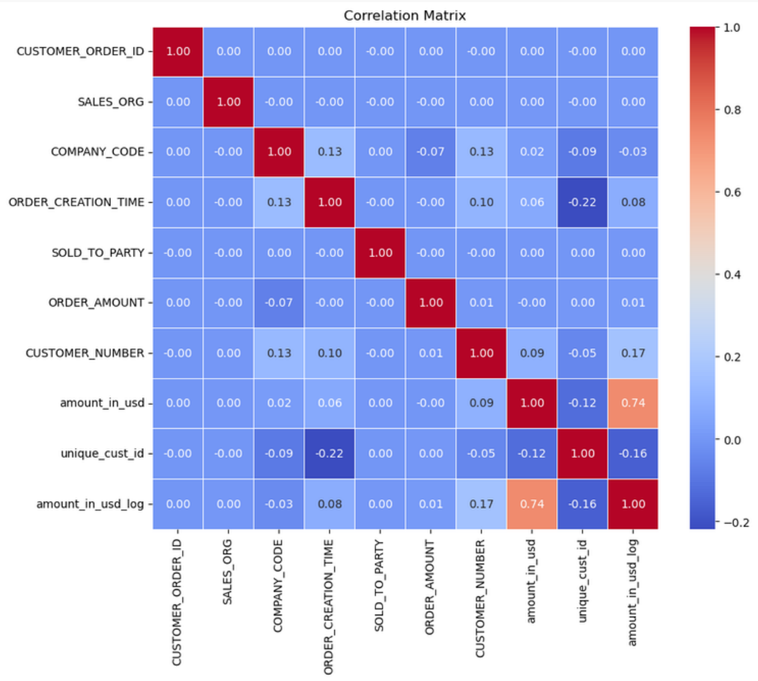
Preprocess and engineer the data, ensuring it is suitable for building a robust machine learning model.

- **Outlier detection and replacement in the "amount_in_usd" column.**
- **Label encoding or One-hot Encoding on categorical columns, Log Transformations on continuous columns and creating new features by grouping existing columns.**
- **Heatmap to find correlation between columns.**

Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

Feature Engineering and Selection...

Correlation Matrix



One-hot encoding

CREDIT_CONTROL_AREA

	NR01	NR02	NR03	NR04	SR01	SR02	SR03	SR04
0	0	0	0	0	0	1	0	0
1	0	0	1	0	0	0	0	0
2	1	0	0	0	0	0	0	0
3	0	0	0	0	0	1	0	0
4	0	0	0	0	1	0	0	0
...	...	...	...	...	...	...	...	...
1101920	0	1	0	0	0	0	0	0
1101921	0	0	0	1	0	0	0	0
1101922	1	0	0	0	0	0	0	0
1101923	0	0	0	0	0	0	0	1
1101924	0	1	0	0	0	0	0	0

Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

ML Models and Evaluations...

ML Models and Evaluations.

- **Models used: Linear Regression, Random Forest, and AdaBoost**
- **Evaluation metrics: MSE, RMSE, and R-Square to compare model accuracies.**

Milestone 1 | Milestone 2 | Milestone 3 | Milestone 4

ML Models and Evaluations...

LINEAR REGRESSION

```
Mean Squared Error (MSE): 0.8069919785088896
Root Mean Squared Error (RMSE): 0.8983273225884257
Mean Absolute Error (MAE): 0.664170097738177
R-squared (R^2): 0.8460657921989152
```

Random Forest

```
Mean Squared Error (MSE): 134007.80500227807
Root Mean Squared Error (RMSE): 366.07076501993174
Mean Absolute Error (MAE): 256.4054669760091
R-squared (R^2): 0.2095625017987025
```

Ada Boost

```
Mean Squared Error (MSE): 0.13765957627791936
Root Mean Squared Error (RMSE): 0.3710250345703365
Mean Absolute Error (MAE): 0.24683036892333282
R-squared (R^2): 0.9737413525971732
```

Results

ADA BOOST

```
print(best_model)
```

```
AdaBoostRegressor(learning_rate=0.1, loss='square', n_estimators=100)
```

CONCLUSION

In conclusion, our project aimed to develop a Machine Learning model to predict the order amount that customers might place in the upcoming days, with a focus on B2B operations. I successfully accomplished this goal through a series of well-defined milestones, each contributing to the overall success of the project.