

Raj pravin Rajendran

Data Scientist & Machine Learning Engineer | Deep Learning | Natural Language processing | Data Visualisation | Python | SQL | Business Intelligence | Data Analyst

6 years of industry experience, including 2 years as a Data Scientist at **AMAZON** and **LIBERTY GLOBAL**, where I used ML algorithms and predictive analysis. In addition, I have substantial expertise and understanding in Python with libraries like as **PANDAS, NUMPY, SKLEARN, TENSORFLOW, and SEARBORN**. I worked with tools such as **Pycharm and Jupyter Notebook**. Stakeholder recognised and appreciated the efficient solutions provided. Obtaining the **STAR** performance award for completing a crucial project Successfully

✉ rajpravin12893@gmail.com

☎ 07747265013

📍 London, United Kingdom

🔗 rajpravin1208.github.io

WORK EXPERIENCE

Data Scientist/ Business Intelligence Engineer

Liberty Global/ Virgin Media

10/2022 - Present

London, United Kingdom

Achievements/Tasks

- Extracted Data from **2+ sources** and converted them into one combined source for reporting purposes using **ETL** techniques and built Business intelligence dashboard in **looker**, saving **2 FTE** of manual reporting work.
- Conducted in-depth analysis and filtration of extensive datasets comprising **100,000** rows, submitted in multiple batches by stakeholders. Leveraged data-driven insights to drive a remarkable **7%** increase in sales for Virgin Media customers..
- Successfully orchestrated the migration of **10 terabytes** of critical data from Oracle Fusion to BigQuery on the **Google Cloud Platform** (GCP).Ensured seamless data transfer, minimizing downtime, and optimizing data accessibility for real-time analytics.
- Successfully connected Looker to BigQuery, enabling the creation of interactive dashboards , ultimately reducing the workload of **2 FTEs**
- Utilized **VERTEX AI** for model deployment and implemented a Python script for seamless model retrieval and prediction, resulting in cost savings, improved efficiency, faster decision-making, and enhanced customer experiences.
- **Optimized Reporting Queries** : Conducted analysis of existing pipelines, applying advanced query optimization techniques and load balance analysis. Reduced daily reporting query execution time by 45 minutes, resulting in annual savings of **240 hours** and improved cluster performance.

Senior Catalog Analyst

Amazon

09/2016 - 09/2021

Chennai, India

EXPERTISE

- Performed efficient feature engineering on a dataset with **148 columns** using Python libraries, including NumPy, Pandas, and Seaborn. This optimization reduced model training time from **1 hour to 35 minutes**.
- Build and trained the model using **Random Forest Classifier, XGBoosting** and **Bagging** classification. Segmented data by creating clusters using K-means and **Hierarchical** algorithm.
- Analysed model prediction accuracy using **Classification report, Confusion matrix, AUC curve**.
- Applied "**Market Basket Analysis/Association rule mining**" on a large customer data to discover the patterns for bundle packing which helped in increasing the sales to **17%**.
- Analysed **30000+** responses to a consumer feedback to evaluate brand perception and given feed back to the vendor to improve the quality of the product which decreased the negative comments from **60% to 23%**.
- Training new joiners and **mentoring** them about the basic workflow and process.
- On Boarded **Vendor calling initiative** , which improved the vendor TAT by 45% resulting in immediate resolution for existing issues on Catalog.

EDUCATION

MSc. Data Science

University of Westminster - London, 7.9 GPA

09/2021 - 09/2022

London, UK

SKILL SET

PROGRAMMING LANGUAGES - Python, SQL ,R

DATA VISUALIZATION - Tableau, Seaborn, Matplotlib, PowerBI and Looker

TECHNIQUES - Machine Learning algorithms, Natural Language Processing, Clustering, Prediction, ETL - Extract, Transform, Load, Time Series Analysis, Neural Network

LIBRARIES/TOOLS - Numpy, Scipy, Pandas, Scikit-learn, Keras, Tensorflow

AWS - S3 , redshift

GCP - Big Query,Vertex AI, DOC AI, Google cloud functions, Looker, Oracle - Fusion Analytics, ERP systems

Version Control: Git (GitHub) - Experience with GitHub for repository management, branching, merging, and pull requests.

GLOBAL PROJECTS

Invoice Matching

Liberty Global

- Utilized **GCP tools (Doc AI, BigQuery, looker), Python, and SQL** to automate invoice-PO matching, which help in reducing 10 positions in a team.

Vendor Compliance - Dispute Reduction:

Amazon

- Channelizing and categorizing vendor errors during the time of shipment. This effort lead to **\$60K** savings annually across all market regions.

Recommended System for Books Department

Amazon

- Implemented a content-based recommendation system on the application which recommends optimal options to users based on the history data using Pandas library.
- Applied **Content-based recommendation system** algorithm for the books category which increased in the sale of **8%**.

Title Guidelines

Amazon

- Have fixed the title style gate issues for few product lines, which reduced the overall work inflow by **13%** of the title update issues across US market place

Removing the invalid data from the product description on Amazon catalog page using python:

Amazon

- Finding and replacing the invalid data from the detail page via Python and this project helped in headcount savings of **3.17 FTE**

Sentiment analysis of Mobile Phones brands:

Amazon

- Using the **NLTK** library and the **Natural Language Processing (NLP)** approach, I reduced negative sentiments toward a mobile phone brand from **60% to 23%** .