Data Scientist

Presently working as Data Scientist in banking domain (Data Analytics) department at NTT data services.

• Total 6+years' experience in Analytics field and 3 years in Machine Learning/Data Science with large data sets of structured and unstructured data, data cleansing, Exploratory data Analysis to extract the hidden information and convey the same to business stake holder using predictive modelling and data visualization.

• Adept in statistical programming languages like Python.

• Proficient in managing entire data science project life cycle and actively involved in all the phases including data acquisition, data cleaning, features scaling/extraction, statistical modelling (decision trees, classification techniques, regression models, neural networks, SVM, clustering), dimensionality reduction using Principal Component Analysis, testing and validation.

•Strong Knowledge of Pandas, Numpy, Sklearn model.

•Adept and deep understanding of statistical modelling, univariate/multivariate analysis, model building, problem analysis, model Calibration and Deployment.

•Have good hands-on knowledge in the area of deep learning and strong understanding of Neural Network optimizations.

• Have good knowledge with TensorFlow, Keras.

• Extensive experience in SQL and applying datawarehouse concepts.

• Good knowledge of data visualizations using statistical programming (Python).

• Good industry knowledge, analytical &problem-solving skills and ability to work well with in a team as well as an individual.

Contact

Skills

Statistical Analysis

Data mining

Machine Learning

Neural networks

Deep Learning (Text/Image)

Dimensionality Reduction

Prediction Analytics

Cluster Analysis

Work History

2019-05 - Current **Data Scientist**
NTT data services

2014-08 - 2019-05 **Associate**
Cognizant Technology Solutions, Bangalore, Karnataka

Projects

Data Science projects:

Providing the interest rate and bank loans:

Determining the interest rate and bank loans for merchant using the machine learning and NN techniques. The goal is to provide the bank loans with suitable interest rate to merchant by determining its performance .

.

Classification techniques

Tree-based Methods

Ensemble Methods

Support Vector Machine

Python

SQL

UNIX Scripting

Datawarehous e

ETL Tool

Roles & responsibility:
1. Data cleansing, Data munging using SQL, Python
2.Performed the exploratory analysis using the Pandas and Numpy packages.
3. ANN is used to determine whether the merchant performance and using the concept of linear regression to determine the interest rate.
4.Saved the model using Pickle and deployed using Azure DevOps.

Tools Used: - Python, SQL,Unix

Algorithm used: - Keras(sigmoid function), OLS,SVC,XGboost,Keras.

Exception report analysis (US bank):

To do exploratory analysis on the settled transaction by merchant. The exception is raised when the transaction is settled in 7 days. To know the percentage of it we have performed the exploratory analysis

Roles & responsibility:
1. Data cleansing, Data munging using SQL, Python
2.Performed the exploratory analysis using the Pandas and Numpy packages.
3.Built the graphs on the basis of exceptions raised for a week.
4.Saved the model using Pickle and deployed using Azure DevOps.

Tools Used: - Python, SQL, Excel
Package used: - seaborn, matplotlib, Numpy, Pandas.

Customer Segmentation (Advance Auto parts)

This Project was to analyse behaviour of customer from real time. Data includes real data and demographic data of users. It includes the segmentation of customer group and classify it demographically

Roles & responsibility:
1.Desigined a model that segment the customer on the basis of the demographics.
2. Deep cleaning of data with help of domain experts and other feedback from clients using Regex.
3. Regular meeting with Domain Experts for building formation and algorithm.
4. Built Multiclass classification.
5. Validation is done on generalized daily data collected.

Tools Used: - Python, Excel, SQL, Files
Algorithm used: - Kmean.

AAP Forecasting

AAP Subscription Activity is a project on Data Analytics to capture the data for different states of subscription of customers so that it can be used for inferential analysis and to build multivariate forecasting model to understand Seasonality and Trends in data.

ROLES AND RESPONSIBILITIES:

1.Data cleansing, Data munging using Python.
2. Univariate and Multivariate analysis on Data using Matplotlib, seaborn packages.
3. Machine learning to infer meaningful insight from data.
4.App store subscription multivariate forecasting using SARIMAX algorithm.
5.OLS regression to identify driving factor for growth in Subscriptions.
6.Extracting, analysing, specifying and validating business process requirements.

Tools Used: - Python, SQL
Algorithm used: - Sarimax, OLS regression, Time series analysis

Datawarehouse Project

MGM Resorts International has been working to expand their ability to use data and analytics to ensure
better decisions while increasing revenues across their business lines. The data warehouse, called the Unified Business View (UBV) and Data Marts will be a central repository for data from Hotel, Human Resources Expenses, Casino, Retail, Food & Beverage, and Entertainment business lines and will enable enterprise-wide reporting and analytics

Roles & responsibility:

1. Requirement Gathering, Analysis and Development part for establishing an Enterprise Data Warehouse from the core to the Mart level.
2. UBV (Data Warehouse) implementation helped to integrate data from various source systems into a single data mart which enables the business to get proper non-redundant reporting from a single point.
3. Involvement in Unit testing for the code Developed and coordinating with Quality Assurance team to get the code refineries checked at the very microscopic level.
4. Experience in working in Production Support environment with complete system interaction and end to end view on the process followed.
5. Involvement in Requests and Incidents creation/maintenance, Raising/closure of change request and change orders.

6. Experience in handling and documenting essentialities like Approach document, Unit Test cases, Implementation Plan, roll out Plan required for code movement from lower environment to Production.
7.Developed innovations, Automations, Enhancements with high value in terms of appreciation and revenue