

Vipul Shrivastava

vipulmks89@gmail.com

+91 84316 88715

Bengaluru, IN

[LinkedIn](#)

Experienced data science professional with excellent analytical and interpersonal skills. Seeking for the position of Data Scientist / ML Engineer, with 2+ years of hands-on experience utilizing statistical modelling and machine learning techniques. Proficient in performing exploratory data analysis, data cleaning, building and evaluating models to implement analytical solution.

PROFILE SUMMARY

- Ability to work with large data sets and present findings / insights to key stakeholders
- Good understanding in applied statistics skills, such as distributions, inferential statistical, hypothesis testing, regression, predictive analytics etc.
- Expertise in analysing complex, high-volume, high-dimensionality data from varying data sources
- Extensive knowledge of working with SQL databases like MySQL and Oracle.

TECHNICAL SKILLS

Programming Languages: Python, SQL

Packages: NumPy, Pandas, Seaborn, Matplotlib, SciPy, Scikit-Learn

Statistics/ML: Linear/Logistic Regression, Decision Tree, Random Forests, k-means Clustering, Naive Bayes
Gradient Boosting, Time Series Forecasting (ARIMA, SARIMA), Recommendation System

Analytics Toolset: Jupyter Notebook, Altair Knowledge Studio, Altair Monarch

Visualization Tools: Tableau, Altair Panopticon

Databases: Oracle, MySQL

PROFESSIONAL EXPERIENCE

Altair Engineering | Associate Data Scientist

Aug '20 – Mar'21

Altair Engineering | Data Science Trainee

Aug '19 – Jul '20

Project: - Machine Predictive Maintenance

Roles & Responsibilities

- Performed data pre-processing techniques such as data cleaning, data wrangling, data preparation and data quality issues handling to make data useful for creating machine learning models
- Identified and observed any existing pattern between independent variables using exploratory data analysis
- Created multiple classification model (Logistic Regression, Decision Tree, Random Forest) and identified the probability of machine failure
- Tested & evaluated predictive models and monitored outcomes as per the requirement

Proof of Concept (PoC): - RFM (Recency, Frequency, Monetary) Analysis

With increase in number of warranty claims, a proof of concept was conducted on warranty claim data to cluster it into low, medium, or high priority based on RFM analysis.

Proof of Concept (PoC): - Recommendation System

Prepared and demonstrated the use of recommendation system that bank can rely upon to cross-sell or upsell their banking products. The PoC involved identifying which products the existing customers will likely purchase next based on their past behaviour and that of similar customers.

- Interpreted and analysed data to draw conclusions for managerial action and strategy
- Presented data and findings to team in order to improve strategies and operations
- Developed PL/SQL stored procedure, functions and style sheets to reduce data retrieval time
- Analysed existing SQL queries for performance improvements as per the business requirement

- Created and validated SQL script for database tables, views, and stored procedures
- Developed SQL queries using joins, sub-queries, and views to perform different database operations
- Provided development support for existing systems, troubleshoot data and system issues
- Performed testing and debugged existing SQL and PL/SQL Scripts for modification

FREELANCING EXPERIENCE

- Worked directly with client stakeholders to develop technical solutions for business cases
- Developed responsive website from Photoshop Document (PSD) file. And used CSS/Bootstrap to design multiple responsive web applications
- Performed testing to identify bugs and technical issues before and after deploying
- Monitored website performance and rectified front-end-related technical issues

CAPSTONE PROJECT

Domain: BFSI | *Tech Stack:* Python, Jupyter Notebook

Objective: - To determine the factors affecting credit risk and to build scorecard to acquire the right customers using past demographic & credit bureau data

- Worked independently and as part of team to achieve proposed objective
- Collect, clean, transform and validate data as a part of data pre-processing
- Presents data in the form of charts, plots, and tables for immediate reference
- Built **Logistic Regression** and **Random Forest models** to predict customer default and build application scorecard to identify good and bad customers.

EDUCATION

- **Post Graduate Diploma in Data Science | IIIT Bangalore & upGrad | Sep '18 – Sep '19 with 3.86/4 CGPA**
- **M. Tech. – Computer Science | Manipal Institute of Technology | Manipal, IN | Aug '11 – Aug '13 with 8.83 GPA**
- **B. E. – Computer Science | Mahakal Institute of Technology | Ujjain, IN | Jul '06 – Jun '10 with 73.78%**