

PRUDHVIRAJU SRIVATSAVAYA

E-mail: prudevds@gmail.com

Mobile: +91 8985632426

Professional qualified Data Scientist/Data Analyst with overall 6+ years of experience in Analytics including Machine Learning, Data Mining, Statistical Analysis and ETL Testing

PROFESSIONAL SUMMARY

- ✓ A dynamic professional with 2+ years of experience in Machine Learning and 4 years of experience in ETL Testing
- ✓ Experienced at creating data Regression models, using predictive data modelling, and analysing data mining algorithms to deliver insights and implement action-oriented solutions to complex business problems.
- ✓ Involved in all phases of data science such as data extraction, data cleaning, statistical modelling, data visualization, model building, evaluation, testing, deploying and monitoring
- ✓ Experience with machine learning algorithms Linear Regression, Logistic Regression, SVM, Naive Bayes, KNN, Ensemble algorithms, Boosting algorithms, Dimensionality Reduction Algorithms Recommendations and Clustering algorithms
- ✓ Experience in NLP algorithms
- ✓ Good understanding on deep learning algorithms
- ✓ Good experience in performing Descriptive analysis and Exploratory Data Analysis (EDA)
- ✓ Rich knowledge in Statistics and Probability
- ✓ Good understanding on Hadoop framework
- ✓ Hands on experience in ETL Testing and Salesforce Testing
- ✓ Hands on experience with Oracle, SQL Server, Hive, HBase and Cassandra
- ✓ Experience in preparing the test plan and test estimates
- ✓ Experience on executing the projects in Agile Methodology
- ✓ Proficiency in writing complex SQL queries to validate business logics
- ✓ Strong understanding of database design
- ✓ Having good experience on SDLC, STLC and Defect life cycle

AREAS OF EXPERTISE

- ✓ **Operating Systems:** Windows Family, UNIX
- ✓ **Programming Languages:** Core Java, Python, R basics
- ✓ **Databases:** Oracle, SQL Server
- ✓ **Big Data:** HDFS, Hive, HBASE, Kafka, Flume
- ✓ **Reporting Tools:** Tableau
- ✓ **Bug Tracking Tools:** QC 10.0, Bugzilla, TFS
- ✓ **Query Languages:** SQL, HiveQL, CQL
- ✓ **ETL Tools:** Informatica
- ✓ **Other Tools:** Sql Developer, UltraEdit, SSH Client, Notebook, Spyder, WinScp, Eclipse

POC's EXECUTED

- ✓ Amazon Fine Food Reviews
- ✓ Personalized cancer diagnosis
- ✓ Taxi demand prediction in New York city
- ✓ Microsoft malware detection
- ✓ Netflix movie recommendation system
- ✓ Stack overflow tag predictor
- ✓ Quora question payer similarity problem
- ✓ Self driving car

PROJECTS EXECUTED

Project: 1

Title: Yanmar Analysis

Description: Project aims to predict the engine failures before docking. The process includes understanding the data and find out what are all caused parameters for the damage and send an acknowledge to the customer before damage so that they will do require maintains.

Roles and Responsibilities:

- ✓ Work with the stakeholders to identify opportunities for leveraging data to drive business solutions
- ✓ Involved in finding the right metric to be used to measure the performance of the model
- ✓ Analysing the data to identify the patterns and trends
- ✓ Processing, cleaning and verifying the integrity of data used for analysis
- ✓ Performed exploratory data analysis using seaborn, matplotlib and bokeh
- ✓ Involved in model training, evaluation and testing
- ✓ Involved in deploying the model in pre-production and production, monitoring and retraining the model
- ✓ Creating a summarized document with the finding using visualizations and other means

Project: 2

Title: Management Awareness System

Description: The Management Awareness System (MAS) is the process of systematic analysis of a Detroit Police Department member's performance patterns to alert supervisors of the possible need for intervention or special recognition. This System is a thorough, reliable and extensible platform for the acquisition, analysis, reporting and monitoring of early warning behavioural characteristics of members of Detroit Police Department.

Roles and Responsibilities:

- ✓ Work with the stakeholders to identify opportunities for leveraging data to drive business solutions
- ✓ Involved in finding the right metric to be used to measure the performance of the model
- ✓ Analysing the data to identify the patterns and trends
- ✓ Processing, cleaning and verifying the integrity of data used for analysis
- ✓ Performed exploratory data analysis using seaborn, matplotlib and bokeh
- ✓ Involved in model training, evaluation and testing
- ✓ Involved in deploying the model in pre-production and production, monitoring and retraining the model
- ✓ Creating a summarized document with the finding using visualizations and other means

Project: 3

Title: PANGEA (Platform for Analytics and Next Generation Engineering Applications)

Description: Pangea offers a workbench approach by enabling analytics stakeholders to collaborate, build and operationalize analytics solutions for business problems. The business problem is introduced into the system as a Service request (SR) and it follows a systematic workflow to ingest data, build, deploy and monitor the analytics models that generate desired insights for solving the given business problem.

Roles and Responsibilities:

- ✓ Involved in designing the architecture of the project
- ✓ Explaining the functionality of the Algorithms to developer
- ✓ Involved in building Augustus scoring engine
- ✓ Building use cases in Pangea and Python to compare the performance of the scoring engine
- ✓ Involved in defining the metrics and tuning parameters to be used for each algorithm
- ✓ Involved defining the functions to be used in the Pre-processing and transformations
- ✓ Involved in setting up the Hadoop environments and deploying the application

Project: 4

Title: Franklin – Underwriting (Commercial Insurance)

Description: Data Management in the insurance domain of one of the largest Insurance Providers in US with the DATA WAREHOUSE testing. The Data is transformed in the Data Warehouse to support business analysis and reporting. Data warehousing is not just the data but also the architecture and the tools to collect, query, analyse and present information. There is a repository of data where data is coming from different sources like input files from BETA applications/operational data store tables. After transforming the data it will be loaded in a BDW brokerage data warehouse which is a historical data store.

Roles and Responsibilities:

- ✓ Reviewing and understanding the requirements and Regular audience on all the client calls during requirements gathering
- ✓ Reviewing the technical design and functional design documents
- ✓ Creating test scenarios and test cases
- ✓ Creating requirement traceability matrix against requirements, test scenarios and test cases
- ✓ Running and monitoring the ETL jobs to load data from source to target
- ✓ Creating complex SQL queries to validate business logics
- ✓ Executing the SQL queries to compare the data of Target systems against source systems
- ✓ Executing the test cases, reporting and tracking the defects till closure using defect tracking tools (ALM, JIRA)
- ✓ Preparing daily, weekly and months status reporting document

EMPLOYMENT NARRATION

- ✓ Since Jun 4th, 2012 to Sep 19th, 2014 worked with LIVA IT Services Pvt. Ltd, Hyderabad as Software Test Engineer.
- ✓ Since Sep 25th, 2014 to Mar 17th, 2016 worked with Capgemini India Hyderabad as Associate Consultant.
- ✓ Since March 21st, 2016 to Aug 12th, 2016 worked with Infosys Limited as Test Analyst.
- ✓ From Oct 25th, 2016 associated with HCL Technologies Limited as Test Analyst.

CERTIFICATIONS

- ✓ Completed ISTQB Foundation Level Certification
- ✓ Completed SQL Foundation Level Certification

ACADEMIC

- ✓ B.Tech in ECE from Pondicherry university -6.86 (CGPA)
- ✓ Intermediate from STPP GOVT JUNIOR COLLEGE, Yanam -79.60%
- ✓ S.S.C from Regency public School, Yanam -70.33%

I hereby confirm that the information given above is true to the best of my knowledge.