Prudhviraju Srivatsavaya
Ph: +91 8985632426

**PRUDHVIRAJU SRIVATSAVAYA**

**E-mail:** prudivds@gmail.com | **Mobile:** +91 8985632426

Professionally qualified Applied Machine Learning Engineer 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
- ✓ Design and develop highly scalable, machine learning applications, processing massive amounts of data
- ✓ Involved in all phases of machine learning project life cycle such as data extraction, data cleaning, transform, feature development, statistical analysis, build solutions and deploy machine learning models
- ✓ Experience with machine learning algorithms Linear Regression, Logistic Regression, SVM, Naive Bayes, Ensemble, XGBoosting, PCA, TSNE, Recommender systems and Clustering algorithms
- ✓ Experience and in-depth understanding in NLP algorithms
- ✓ Strong understanding on deep learning algorithms (ANN, CNN and RNN) and Keras
- ✓ Good experience in performing descriptive analysis to describe or summarize data in ways that are meaningful and useful
- ✓ Extensive experience in performing Exploratory Data Analysis (EDA) to summarize their main characteristics, often with visual methods
- ✓ Work with massive data sets (structured and unstructured) and design, develop and implement solutions.
- ✓ Collaborate with other engineers, data scientists, product owners and other stakeholders in designing and developing solutions
- ✓ Leverage industry knowledge and stay close to technology developments in the machine learning communities
- ✓ Proficiency in writing complex SQL queries to validate complex business logics
- ✓ Experience in ETL Methodology for Supporting data extraction, and testing in BI applications
- ✓ Hands on experience with Oracle, SQL Server, Hadoop, Sqoop, Hive, HBase and Cassandra
- ✓ Hands on experience on executing the projects in Agile and Scrum development Methodologies

## AREAS OF EXPERTISE

- ✓ **Machine Learning:** Supervised Learning, Unsupervised Learning, NLP and Deep Learning
- ✓ **Big Data Technologies:** Hadoop Eco Systems (Hive, Sqoop, HBase, Flume), Horton Works
- ✓ **Programming Languages:** Core Java, Python, R(basics)
- ✓ **Python Libraries:** Sklearn, Matplotlib, Seaborn, Bokeh, NumPy, SciPy, Pandas, NLTK, Keras, TensorFlow
- ✓ **Databases:** Oracle, SQL Server, MySQL
- ✓ **No SQL Databases:** HBase
- ✓ **Query Languages:** SQL, HiveQL
- ✓ **ETL Tools:** Informatica
- ✓ **Reporting Tools:** Tableau, Power BI
- ✓ **Bug Tracking Tools:** QC 10.0, Bugzilla, TFS
- ✓ **Other Tools:** Sql Developer, UltraEdit, SSH Client, Jupyter Notebook, Spyder, WinScp, Eclipse

## POC's EXECUTED

- ✓ Personalized cancer diagnosis
- ✓ Amazon fine food reviews
- ✓ Taxi demand prediction in Ney York city
- ✓ Microsoft malware detection
- ✓ Netflix movie recommendation system
- ✓ Stack overflow tag predictor
- ✓ Quora question payer similarity problem

## PROJECTS EXECUTED

**Project: 1**
**Title:** Predictive Maintenance for Yanmar Marine Engine
**Client:** Yanmar India Pvt Ltd
**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.
**Responsibilities undertaken:**
✓ Work with the stakeholders in understanding the business problem and leveraging data to drive business solutions
✓ Involved in finding the right algorithm and metric to be used to measure the performance of the model
✓ Involved in identifying the train, cross validation and test data sets using time based splitting
✓ Performed data cleaning activities such as fill in missing values, smoothing noisy data, identify, treat or remove outliers, removing quasi constant columns and resolve inconsistencies
✓ Involved in identifying trend and distribution change of the features which vary with time using exploratory data analysis
✓ Identify the intervals of trend shift and calculate RML (remaining maintenance life) for those intervals and use this data to train the model
✓ Research, experiment and develop Machine Learning models for regression problems
✓ Involved in model training, evaluation and testing in different environments
✓ Coordinating with the UI team to perform end-to-end testing with respective API's
✓ Involved in deploying the model in production, monitoring and retraining the model
✓ Creating a summarized document with the finding using visualizations and other means. Sharing the insights with the respective stakeholders.

**Project: 2**
**Title:** Management Awareness System
**Client:** City of Detroit, MI, USA
**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.
**Responsibilities undertaken:**
✓ Work with the stakeholders to identify opportunities for leveraging data to drive business solutions
✓ Involved in finding the right algorithm and metric to be used to measure the performance of the model
✓ Removing symbols, numbers using regular expressions and stop words that carry less important meaning than keywords in Peer Description column
✓ Applied Lemmatization on Peer Description to remove inflectional endings and to return the base or dictionary form of a word
✓ Applied tokenization on the Peer Description and removed certain parts of speech using the Named Entity Recognition
✓ Converting peer description into vectors using Ngrams in TFIDF and BOW
✓ Research, experiment and develop Machine Learning models for classification problems
✓ Involved in model training, evaluation and testing in different environments
✓ Coordinating with the UI team to perform end-to-end testing with respective API's
✓ Involved in deploying the model in production, monitoring and retraining the model
✓ Creating a summarized document with the findings using visualizations and other means. Sharing the insights with the respective stakeholders.

**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.
**Responsibilities:**
✓ Involved in designing the architecture of the project
✓ Explaining the functionality of the Algorithms to developers
✓ 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)
**Client:** Zurich North America
**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.