

## SKILLS

**Languages:** SQL | Python | MATLAB | HTML | CSS | Javascript | PL/SQL |

**Tools:** Azure Devops | Databricks | Power BI | AWS | Apache Spark | Kafka | Snowflake | Tableau | Oracle | Informatica

**Technical Skills:** Machine Learning | GenAI | Database Management & Administration | EDA | Data Visualization | Data Analytics | ETL

## EDUCATION

Master Of Science in Data Science, **University at Buffalo – Buffalo, NY**

**Aug 2024- Dec 2025(Expected graduation)**

Courses - Numerical Math, Probability Theory, Programming Database Fundamentals, Statistical Learning and Data Mining

Bachelor of Engineering in Biotechnology, **New Horizon College of Engineering – Bangalore, IN**

**3.62 GPA | 2015-2019**

Courses – Engineering Mathematics, C programming, Biostatistics, Bioinformatics

## EXPERIENCE

**Tata Consultancy Services, Data Engineer**

**Jul 2022 – Aug 2024**

- Developed stored procedures in SQL with minimal runtime to avoid timeouts in webpages. Modeled tables, views and indexes in SQL according to business requirements.
- Created and modified complex Views in Oracle for the Sales Compensation and Financial Controls reports using Datawarehouse Tool Wherescape Red (Client ETL tool) and deployed using Azure Devops.
- Leveraged advanced SQL techniques (CTEs, Joins) to extract data and improve efficiency. Designed mappings between source and target database (SQL) and implemented ETL job applications using Informatica. Created Workflow mapping, database connections using Informatica Designer.
- Worked on troubleshooting slow performing views and tuned them reducing their runtime significantly by 70%.
- Collaborated with cross functional teams for optimization strategies in Agile environment.

**Tata Consultancy Services, Database Administrator**

**Dec 2019 – Jul 2022**

- Demonstrated expertise in managing and optimizing large databases, performance tuning of TSQL stored procedures significantly reducing query response time by 10% and a daily workload reduction by 4 hours.
- Migrated 300+ 2012 SQL server databases to 2019 SQL Version and setup of always on environment disaster recovery management and configuration.
- Automated plans and jobs like index rebuild, update statistics, backup/restore for smooth functioning of databases. Configured and maintained Log shipping of Production databases for on time disaster recovery.
- Reviewed execution plans of slow running queries and implemented indexes to optimize query efficiency by 26%.
- Troubleshot failing jobs and provided permanent fix in production and reduced escalation by 75%.
- Performed data manipulation and reduced the number of incidents to 20%.

## PROJECTS

**Loan approval Prediction(Dec 2024-Dec 2024)**

- Developed and evaluated a loan approval prediction system using machine learning models including Logistic Regression, RidgeClassifier, Random Forest, and XGBClassifier achieving an F1 score of 0.8350 for XGBClassifier.
- Designed and deployed a containerized API with FastAPI and Docker, hosted on Railway for real-time loan prediction and integrated a Streamlit application for user interaction.
- Implemented end-to-end ML lifecycle tracking with MLFlow and DagsHub including hyperparameter tuning, feature selection and PCA for dimensionality reduction.

**Transforming Health Care: A Machine Learning Approach for Heart Attack Prevention(Sep 2024-Dec 2024)**

- Developed a predictive model using Random Forest and other machine learning algorithms (Logistic Regression, Decision Trees, XGBoost) on a large dataset (400,000+ rows) to identify key health and demographic risk factors for heart attacks, achieving 94.84% accuracy and a 42.42% F1-score.
- Performed data preprocessing and EDA, including handling missing values, balancing class distribution (SMOTE/ROSE), feature scaling, and visualizing key insights namely correlations, age and smoking status impact. Key predictors included Had Angina, Age Category, and BMI.
- Fine tuned the models through cross-validation and hyperparameter optimization while addressing class imbalance and overfitting, delivering actionable insights for heart attack prevention strategies.

**Predicting the reasons of Attrition of the employees of a company(Dec 2023-Jan 2024)**

- Created two machine learning models (Logistic Regression and Random Forest) to predict the cause for attrition of employees in a company. Attained 89% accuracy in Logistic Regression Model.

**Creating and training models to predict the successful recovery of a rocket's part(Oct 2023-Nov 2023)**

- Developed four Machine Learning Models (Logistic Regression, Support Vector Machine, Decision Tree Classifier, and K Nearest Neighbors) for new SpaceY company who wants to bid against Tesla's SpaceX company to predict successful recovery of a rocket's part as SpaceX is successful in it.
- Collected data from public SpaceX API and SpaceX Wikipedia page and created models. All models produced similar results with accuracy rate of about 83.33%.

## CERTIFICATIONS & ACHIEVEMENTS

- **IBM Data Science Professional Certification:** 6 months coursework on learnt Machine learning techniques, Data visualization and Data analysis using Python.
- **Leadership:** Mentored 3 team members and performed code reviews weekly.
- **AWS Cloud Practitioner:** Cloud Concepts completed on LinkedIn Learning.
- **Agile Development Practices:** Agile practices completed on LinkedIn Learning.
- Team Member of the University at Buffalo Data Analytics Club : AI exploration with practices in Dataset analysis.

## AWARDS

- Best performance grade 'A' for 4 consecutive years & Best Team Award in 2023