

BHAVANA JAYAPPA

bhavanasj.2000@gmail.com | +1 404 449 6618 | [LinkedIn](#)

Summary:

Data Scientist with 2 years of experience applying machine learning, statistical modeling, and data visualization to solve business problems. Skilled in Python, SQL, and end-to-end ML pipelines with proven ability to translate complex analytical insights into actionable strategies. Experienced in collaborating with technical and non-technical teams to deliver scalable, data-driven solutions that enhance decision-making and organizational efficiency.

Education:

Georgia State University, Atlanta

Aug 2024 – May 2026

Master of Science, Data Science and Analytics

Dr. Ambedkar Institute of Technology, India

Aug 2018 – July 2022

Bachelor of Technology, Electronics & Communication Engineering

Experience:

Georgia State University – Graduate Assistant (Data Scientist)

Aug 2024 – Present

- Extracted, cleaned, and analyzed 15,000+ student records using Python (Pandas, NumPy) and SQL, improving accuracy of program evaluation reports.
- Built 5+ interactive dashboards in Tableau to track engagement and retention trends, enabling administrators to make data-driven decisions faster.
- Applied classification and clustering models to uncover behavioral patterns, helping administrators identify key retention drivers.
- Collaborated with faculty and data teams to communicate statistical findings in clear, actionable formats.

Tata Consultancy Services Ltd - Data Scientist

Oct 2022 - July 2024

- Designed and optimized ML-driven analytics pipelines using Python, SQL, and AWS, improving data processing efficiency and enabling 10% faster system performance.
- Built and monitored KPI dashboards in Power BI, applying Agile development practices to deliver business-ready insights for 200+ stakeholders across sales and finance.
- Engineered MySQL database schema and ETL pipelines with 100+ entities, improving query performance and scalability across operational systems.
- Collaborated with software engineers and data teams to document model lifecycle, establish model monitoring protocols, and ensure compliance with governance standards.

Projects:

Credit Risk Modeling for Loan Approval | Python, Scikit-learn

- Modeled loan default risk on 500K applications using Random Forest, XGBoost, and logistic regression, improving predictive accuracy by 15% and reducing false positives by 20%.
- Implemented end-to-end ML pipeline from data preprocessing to model evaluation, following CRISP-DM and MLOps principles for version control and governance tracking.
- Delivered interpretable model insights for risk assessment teams, reducing false-positive loan rejections by 20%.

Rail Track Defect Classification | Python, TensorFlow, CNN

- Trained a CNN on over 120k+ rail-track images using TensorFlow and Keras, achieving 95% defect detection accuracy and enabling near real-time inference (~2s/image).
- Documented feature extraction and validation workflow under Agile sprint cycles, contributing to iterative model improvements and scalable computer vision deployment.

Skills:

Programming & Frameworks: Python, C++, PySpark, TensorFlow, Keras, Scikit-learn, PyTorch

Databases & Cloud: SQL, MySQL, PostgreSQL, MongoDB, AWS, GCP

Machine Learning & AI: Deep Learning, NLP, LLMs, RAG, Generative AI, Prompt Engineering, Feature Engineering, Time Series Forecasting, Model Validation, Computational Statistics, Graph-Based ML, MLOps

Data Engineering: ETL Pipelines, Spark, Kafka, Hadoop, Data Governance, Model Governance

Visualization & BI: Tableau, Power BI, Excel (Power Query, KPI Dashboards), Matplotlib, Seaborn

Tools & Practices: Git/GitHub, Jupyter, CRISP-DM, Agile/Scrum, CI/CD, REST APIs, Jira

Soft Skills: Communication, Collaboration, Problem Solving, Analytical Thinking