

Radhika Ganesh

Irving, Texas | github.com/radhikaganesh29 | radhikaganesh292024@gmail.com | +1 (930) 333-4998 | [LinkedIn](#) | [Portfolio](#)

EDUCATION

Indiana University Bloomington, United States

August 2023 – May 2025

Master of Science in Computer Science

CGPA: 3.87/4

Courses: Applied Machine Learning, Advanced Database Concepts, Data Mining, Elements of AI, Information Visualization

Sardar Patel College of Engineering, Mumbai, India

August 2017 – May 2021

Bachelor of Technology in Electrical Engineering

CGPA: 9.46/10

Ranked 3rd in Electrical Department based on overall CGPA after 8 semesters of undergraduate studies

Relevant Courses: Image Processing, Artificial Intelligence, Numerical Techniques and Programming (MATLAB)

TECHNICAL SKILLS

- **Data Analysis Tools:** SAS, AWS QuickSight, MATLAB, Power BI (DAX, Power Query/M), Tableau, Google Analytics
- **Big Data & Distributed Computing:** Apache Hadoop, Apache Hive, Apache Spark, Snowflake
- **Statistical Analysis:** Hypothesis Testing, Regression Analysis, Time Series Analysis, Bayesian Statistics, t-test, A/B Testing
- **Programming Languages:** Python, C/C++, SQL, R
- **Databases:** SQL Server, MySQL, PostgreSQL, NoSQL (MongoDB)
- **ML libraries:** Keras, Matplotlib, NLTK, NumPy, Pandas, Scikit, Seaborn, TensorFlow, Scipy, OpenCV
- **Development Tools & Platforms:** Git, GitHub Copilot, Docker, DBT, Bash, Postman, Jira, Confluence, Agile Methodology
- **Courses and Certifications:** [Deep Learning Specialization](#), [Python](#), [Machine Learning](#), [Data Science](#)

PROFESSIONAL EXPERIENCE

Data Analyst | Project 990, Bloomington, Indiana, United States

February 2025 – Present

- Engineered ETL pipelines using Alteryx, Python, MySQL, and Snowflake, orchestrated with DBT, to process 100,000+ IRS Form 990 tax filings and Business Master File records for downstream analysis.
- Leveraged Hive, Hadoop, and Apache Spark to query and store distributed datasets, reducing processing latency by 35% and ETL execution time from 6 hours to under 2 hours.
- Extracted and classified 1,000+ mission statements into 10+ philanthropic categories using RoBERTa, improving data accuracy.
- Built 5+ interactive dashboards in Tableau and AWS QuickSight, delivering insights on grant distribution, impact, and trends.

Software Engineering Intern - Data Solutions | SafeSpace, Bloomington, Indiana, United States

June 2024 – December 2024

- Analyzed user interactions using SQL and Python scripts, identifying trends driving a 10% expected retention increase.
- Integrated Google Analytics with event tracking, improving journey visibility and reducing 3 key drop-offs by 18% post-redesign.
- Devised 7+ Django API endpoints to store structured/unstructured data (images, videos, audio) for sentiment and behavioral analysis.
- Created 5+ dashboards in Power BI, Tableau, and Excel to visualize journal and media trends; collaborated on A/B testing to refine recommendations and boost personalization.

Data Analyst - Credit Risk | Axis Bank, Mumbai, India

August 2021 – June 2023

- Leveraged Python, SAS Enterprise Guide, SQL, and AWS Redshift to extract key risk metrics across 5 loan categories, creating visualizations with SAS Viya, Tableau, Power BI, and QuickSight for enhanced decision-making.
- Automated risk dashboard generation using AWS S3 for storage, AWS Glue for ETL, and AWS QuickSight for visualization, saving 48 hours of manual labor monthly and improving operational KPIs.
- Employed decision trees for multivariate analysis to support data-driven Auto Loan policy decisions, leading to a 15% drop in defaults and a 10% rise in approvals.
- Formulated and executed an application score cutoff strategy for Two-Wheeler Loans, using Glue for ETL, Redshift for data processing, and QuickSight for visualizing results, cutting the default rate by 30%.

PROJECTS

Home Credit Default Risk | Python, PyTorch Lightning

- Achieved 91.94% accuracy in loan default prediction using logistic regression; boosted performance with ensemble models (72.78% validation accuracy, 74.93% AUC) and optimized MLPs in PyTorch Lightning via tuning.

Retail Analytics Dashboard | Snowflake, Tableau

- Implemented a Snowflake-based data pipeline and Tableau dashboards to analyze global retail data, uncovering \$60K+ in product losses and customer churn risks through RFM segmentation and profitability analysis.

Electricity Load Forecasting | Apache Spark (PySpark + MLlib)

- Constructed a time-series forecasting pipeline using PySpark and MLlib on German electricity load data, achieving a 28% RMSE reduction with Gradient Boosted Trees by engineering lag features to capture daily and weekly consumption patterns.

Occupation Reporting Tool | Microsoft Power BI

- Crafted a Power BI dashboard with 10+ visuals across 3+ pages to analyze 1,000+ occupations using O*NET and BLS data, covering job families, education, salaries, and employment trends across all U.S. states and 380+ metro areas.