GANESH RAJ K.

848-313-8525 | ganesh_012@outlook.com | linkedin.com/in/ganeshrajk | github.com/ganeshraj-k | New Brunswick, NJ

EDUCATION

Rutgers University New Brunswick, NJ
Master's in Data Science, CGPA: 3.75 *Aug. 2022 – May 2024*

Indian Institute of Technology Indore Indore, India
Bachelor's in Computer Science and Engineering *Aug. 2015 – May 2019*

CERTIFICATIONS & SKILLS

Certifications

AWS Certified Machine Learning Specialist

AWS Certified Cloud Practitioner

Skills **Programming**: Python (3.x), R, C++ — **Databases**: tSQL, PostgreSQL, MongoDB — **Cloud Services**: AWS (SageMaker, IAM, Fargate, Glue) — **Machine Learning & Deep Learning**: XGBoost, PyTorch, TensorFlow, Keras, OpenCV — **Data Viz & Manipulation**: Tableau, Power BI, Excel (VBA), Matplotlib, Seaborn, Shiny, Scikit-learn, NumPy, Pandas — **Web & API**: Selenium, FastAPI, Flask, REST API — **LLMs/Gen AI**: Llama-Index, Langchain, VectorDB, OpenAI, Ollama, RAG, Hugging Face — **Version Control**: Git, GitHub — **Operating Systems**: Linux

EXPERIENCE

Data Scientist Feb. 2023 – Present
Rutgers UCM *New Brunswick, NJ*

- Predicted at-risk students using an ANN classification model on a dataset of over 3,100 students, including demographic, academic, and Canvas interaction data, achieving 78% accuracy.
- Provided insights to the Career Success Team via Tableau dashboards, contributing to a 21% increase in retention rates compared to the previous year.

Data Scientist June 2019 – Jan. 2022
Deloitte Consulting *Bengaluru, India*

- Achieved 86% accuracy in predicting churn, reducing churn by 4% and preserving $3 million in annual revenue through targeted offers.
- Built an XGBoost classifier on AWS SageMaker, processing 3.4 million records. Engineered features such as sentiment scores from call transcripts, and optimized the model using grid search. Applied SMOTE to balance the dataset.
- Developed an ETL pipeline with AWS Glue, extracted data from AWS RDS using MySQL, and processed data in parallel using Apache Spark.
- Implemented transaction categorization using Logistic Regression (91% accuracy), performed text processing with NLTK/spaCy, and deployed the model via Docker on AWS Fargate. AWS Comprehend analyzed transaction data.
- **Key Achievement**: Was recognized with Applause award twice for my client-centric work approach and timely deliverables.

Data Engineering Intern May 2018 – July 2018
MAQ Software *Hyderabad, India*

- Built an ETL pipeline using SSIS to consolidate over 2 million records from Azure data marts into a single data warehouse.

- Utilized SSMS for database management, executed T-SQL scripts, and ensured data quality with validation checks and stored procedures.

PROJECTS

Automate Your Job Hunt with Selenium and Gen AI | *GitHub, Medium*

- Built an Agentic RAG system using LlamaIndex and Ollama, optimizing job application queries with local cache search and fuzzy matching for faster responses. Deployed the API using FastAPI for seamless integration.
- Automated LinkedIn Easy Apply with Selenium, enabling the bot to apply to over 100 jobs in less than 2 hours by navigating listings, handling inputs, and efficiently submitting applications.

Travelers Insurance Analytics University Contest | *GitHub*

- Performed Tweedie regression using LightGBM on a zero-inflated dataset of over 29,000 records, fine-tuned parameters with grid search, and evaluated model performance with the Gini index, visualizing results using Matplotlib.
- This approach achieved a third-place finish out of 200+ teams.