Ganesh Raj K

◆ ganesh_012@outlook.com ◆ 848 313 8525 ◆ linkedin.com/in/ganeshrajk ◆ github.com/ganeshraj-k

EDUCATION:

- **MS in Data Science**, Rutgers University | CGPA – 3.75 May 2024
- **BTech in Computer Science and Engineering**, Indian Institute of Technology Indore May 2019

EXPERIENCE:

Rutgers UCM, *Data Analyst* **Feb 2023 – Present, New Brunswick**

- Harnessed learning data from canvas Api to predict new student course outcomes/CGPA grade using a classification model which scored a 78 percent success rate in classifying new students.
- The dataset consisted of 12 diverse features including demographic academic, behavioral, parent participation for the predictive modeling.
- Extracted data using canvas Api, preprocessed using python, and used ANN for the classification model.

Deloitte Consulting, Data Analyst **June 2019 – Jan 2022, Bengaluru**

Medical Data NER:

- Enhanced query speed for a medical record database with over 2 million records.
- Used Amazon Comprehend with Python to perform Named Entity Recognition on the DynamoDB dataset.
- Added recognized entities as tags using AWS Glue for the ETL process.

Restaurant Chain:

- Categorized restaurant patrons based on their dining preferences from survey data using K-Means clustering.
- Generated detailed Tableau visualizations to correlate the cluster results with their risk and safety behaviors.
- This analysis helped strategize marketing along with promotion of safety protocols. which led to an 74% increase in take away orders the next quarter.

Banking:

- Mitigated the lockdown-induced customer churn by constructing a multivariate logistic regression model to Identified churn-prone customers and key contributing factors.
- Performed EDA using matplotlib and communicated results to stakeholders. helped them target marketing accordingly resulting in a 30% churn reduction in the next quarter.

Geospatial Intelligence:

- Addressed the challenge of manually identifying docked vessels by developing an object detection system using Mask R-CNN and OpenCV for change detection in Python.
- Accessed high-definition GIS satellite imagery from the Sentinel API in Python and dehazed the images for better results.
- Extracted the geolocation data of the detected objects using QGIS.
- Automated the entire process using AWS Lambda and CloudWatch, saving over $100k in labor costs.

Employee Search:

- Crafted an employee search tool with an R Shiny interface, utilizing R for data handling and text extraction from resumes.
- Applied text search and NLP to match and tag employees with specific skills

***Key Achievement**:* Was recognized with Applause award twice for my client centric work approach and timely deliverables.

MAQ Software, Data Engineer **May 2018 – July 2018, Hyderabad**

- Established an ETL pipeline using SQL Server Management Studio and SSIS, consolidating three large Azure data marts with over 2 million records into one. Developed triggers and stored procedures in place to identify inconsistencies during the transfer and maintain data integrity.

PROJECTS:

Chatbot with a personality: [github]

- Built a generative AI (Gen AI) model chatbot to replicate Chandler Bing's dialogue style from "Friends," utilizing an extensive dataset of 8,700 dialogues.
- The model, featuring a seq2seq with 2-layer LSTM network with a dropout layer, achieved a BLEU score of 0.63.

Twitter Search: [github]

- Designed a web application with a local cache of 200 trending tweets, leveraging a combination of Postgres

(relational) and MongoDB (non-relational) to query a dataset of about 120,000 tweets from 13,000 users.
- Applied NLP techniques for efficient search, including synonym search and Levenshtein distance, and managed API requests with Flask

Traveler's insurance modeling contest: [github]
- Conducted Tweedie regression using LightGBM on a zero-inflated dataset of over 29,000 records, fine-tuned parameters using grid search, and assessed model efficacy with the Gini index. This systematic approach secured a third-place finish among 200+ teams.

Gen AI Linkedin Application bot: [github]
- Automated applying to jobs on LinkedIn easy apply using selenium in python.
- For the question and answering, built a RAG API using LLamaindex for querying with GPT turbo as baseline model and used FastAPI to make and deploy the API Client.

SKILLS

- **Programming & Scripting**: Python, R
- **Machine Learning & Data Science:** PyTorch, scikit-learn, pandas, numpy
- **NLP & Generative AI:** LangChain, LlamaIndex, RAG, vectorDB, FastAPI
- **Cloud & DevOps:** AWS (S3, EC2, SageMaker, DynamoDB, Glue, IAM)
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Plotly

CERTIFICATIONS

- AWS Machine Learning Specialist
- AWS Cloud Practitioner