GANESH RAJ K

|  |  |
| --- | --- |
| Email: [ganesh\_012@outlook.com](mailto:ganesh_012@outlook.com) | LinkedIn: [linkedin.com/in/ganeshrajk/](https://www.linkedin.com/in/ganeshrajk/) |
| Phone: +1-848-313-8525 | GitHub: [github.com/ganeshraj-k](https://github.com/ganeshraj-k) |

**E****DUCATION**

* ***Rutgers University***, Master’s in Data Science (GPA – 3.79) **Aug 2022 - May 2024**
* ***Indian Institute of Technology Indore***, Bachelor of Technology in Computer Science and Engineering **June 2015- May 2019**

**SKILLS AND CERTIFICATIONS**

|  |  |  |
| --- | --- | --- |
| ***Programming:*** Python, R, C++ | ***Database:*** Postgres, SQL, MongoDB | ***Data Visualization:*** Power BI, Tableau |
| ***Libraries:*** scikit-learn, pandas, NumPy, seaborn, matplotlib, OpenCV | ***Frameworks*:** PyTorch, TensorFlow, Flask | ***Tool***: AWS, Azure |

* AWS Machine learning Specialist Certification
* AWS Cloud Practitioner Certification

**PROFESSIONAL EXPERIENCE**

***Deloitte*, *Business Analyst* June 2019 – January 2022, Bangalore**

* Employed data mining techniques and logistic regression to accurately predict customer churn in the banking sector.
* Leveraged Tableau for creating comprehensive data visualizations and dashboards, uncovering insights and patterns pertinent to the client’s business objectives.
* Presented data-driven recommendations that enabled the client to target their strategies effectively, resulting in a 33% reduction in churn the following quarter.
* Engineered an automated boat detection system leveraging Mask R-CNN for segmentation and Optical Flow in OpenCV for image mask change detection.
* Enhanced image quality using advanced dehazing, utilized GeoTIFF for accurate geo-location, and deployed YOLO for swift, precise object detection. Automated the workflow using AWS Lambda, SageMaker, and CloudWatch, reducing manual workload by 35 hours weekly.
* Categorized customers of a restaurant chain with over 350 locations using k-means clustering to assess safety perception and risk profiles.
* In parallel, developed Tableau dashboards to analyze demographic trends, revealing insights that were crucial in shaping more targeted marketing strategies presented to the clients.
* Implemented an R Shiny-based employee search tool, applying R for data processing and NLP to extract and tag skills from resumes. Created R Shiny visualizations to display employee data trends and skill distributions.
* Managed a real-time Excel database for over 700 employees, ensuring up-to-date tracking of their status.
* Led a machine learning project to enhance medical record management for a provider with 2+ million records. Established an ETL pipeline via AWS Glue for efficient data handling and leveraged AWS Comprehend for NER, boosting record query speed and case segmentation accuracy.

***MAQ Software, Database Management Intern* May 2018 – July 2018, Hyderabad**

* Implemented an ETL pipeline in SQL Server Management Studio and SSIS to clean, structure, and consolidate three Data Marts into one. Also developed triggers and stored procedures to identify and resolve errors and inconsistencies in tables during the transfer process.

***UCM Rutgers, Unit Computing Specialist* February 2023 – present, New Brunswick**

* Developed Python scripts for Excel data cleaning, reducing manual work by 30+ hours quarterly.
* Created Tableau dashboards for budget analysis and IT ticket trends, identifying common issues to support decision-making.

**PROJECTS**

***Chatbot model with a personality:***

* Crafted a seq2seq AI model to mimic Chandler Bing’s dialogue style from “Friends,” using an 8,700-dialogue dataset and implemented a 2-layer LSTM with a dropout layer, which achieved a BLEU score of 0. 63.
* Preprocessed the data using Python and created Tableau dashboards to analyze the show’s data, such as average season ratings, episode rankings, character interactions, and dialogue frequency, which informed further model development and data understanding.

***2024 Travelers Insurance Analytics University Contest:***

* Performed Tweedie regression on a zero-inflated dataset with over 29,000 records, fine-tuning parameters via grid search and assessing model efficacy with the Gini index.
* Executed comprehensive EDA and data visualizations to discern distributions of individual attributes, which informed model selection. This methodical approach garnered us third place among 200+ teams and recognition on campus for precise insurance claim predictions. precise insurance claim predictions.

***Bloomberg Global University Trading Challenge:***

* Secured third place in the Bloomberg Trading Challenge, generating a profit of over $55,000 from a $500,000 capital across 12 positions and executing 200+ trades in two months using the Bloomberg Terminal.
* Leveraged the terminal’s news search and sentiment analysis to inform trading strategies and monitor stock metrics.

***Twitter Search****:*

* Created a web application for querying a dataset of ~120,000 tweets from 13,000 users, utilizing a combination of Postgres (relational) and MongoDB (non-relational), with a local cache of 200 trending tweets.
* Applied NLP techniques for efficient search, including synonym search and Levenshtein distance, and managed API requests and the web app using Flask.

***Ear image recognition***

* Implemented image recognition on ear images using a CNN model with SIFT features for biometric authentication. Captured and preprocessed high-resolution images of students, achieving an accuracy of 91% and a precision of 73%.

**VOLUNTEERING**

* **Club Head, AVANA**: Led AVANA, the campus social welfare club at IIT Indore, spearheading initiatives such as cleanliness drives, weekend teaching workshops for underprivileged children, and blood donation drives.
* **Volunteer Team Lead, Breathe India:** During the COVID-19 second wave, I facilitated access to medical resources by daily updating our website with information on available beds, nebulizers, and medicines, and promptly connected individuals to the necessary healthcare services.