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, Excel, MS Visio |
| ***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*, *Data Scientist* June 2019 – January 2022, Bangalore**

* Developed a logistic regression model to predict customer churn in banking.
* Evaluated model performance and feature relevance using k-fold cross-validation and conducted in-depth EDA to understand feature distributions.
* Insights derived from the model informed client strategies, contributing to a 33% reduction in churn.
* 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.
* Led the efforts to categorize customers of a restaurant chain with over 350 locations by safety perceptions and risk profiles using K-means clustering and RFM analysis.
* Integrated NLP for sentiment analysis of feedback on safety during lockdowns. Utilized Tableau for visual analytics, aligning sentiment with demographics to devise precise marketing strategies, resulting in a better targeted marketing strategy.
* 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.
* Rolled out as an internal project for an organization with over 700 staff, enhancing the efficiency of talent identification.
* 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**

* Wrote python scripts for Excel data cleaning, cutting down 30+ hours of manual work each quarter and regularly analyzed office hardware specs via NLP for timely upgrades.
* Administered Active Directory, maintaining security and permissions, while managing software updates for the department through Kace, ensuring optimal system performance.

**PROJECTS**

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

* Developed a generative AI (Gen AI) seq2seq model to replicate Chandler Bing’s dialogue style from “Friends,” utilizing an extensive dataset of 8,700 dialogues. The model, featuring a 2-layer LSTM network with a dropout layer, achieved a BLEU score of 0.63.
* Enhanced Google’s T5-small model, setting output size to 50, for benchmarking against the baseline.

***2023 Travelers Insurance Analytics University Contest:***

* Performed Tweedie regression for a zero-inflated dataset with 29,000+ records, optimizing parameters via grid search and evaluating with the Gini index.
* Achieved third place among 200+ teams and campus recognition for 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.