|  |  |  |
| --- | --- | --- |
| New Brunswick, NJ  (848) 313 8525  [ganesh\_012@outlook.com](mailto:ganesh_012@outlook.com) | **GANESH RAJ K** | [Linkedin.com/ganeshrajk](https://www.linkedin.com/in/ganeshrajk/)  [Github.com/ganeshraj-k](https://github.com/ganeshraj-k) |

**E****DUCATION**

* **Master’s in data science,** Rutgers University New Brunswick May 2024
* **Bachelor’s in computer science,** Indian Institute of Technology Indore May 2019

**TECHNICAL SKILLS**

* **Programming Languages:** Python, R
* **Libraries and Frameworks:** Git, Spark, PyTorch, TensorFlow, MongoDB, Postgres SQL, scikit-learn
* **Cloud:** lambda, S3, CloudWatch, Sage maker, IAM, EC2
* AWS certified Cloud practitioner
* AWS certified Machine Learning specialist

**EXPERIENCE**

|  |  |
| --- | --- |
| **Freelance Data Analyst, Potlock** | ***May 2024 – Present, Remote*** |

* Extracted data from over 20,000 blockchain transactions of quadratic funding rounds using SQL. Conducted a thorough analysis of application rates and approval rates.
* Prepared comprehensive charts to visually represent the observations based on the findings.

|  |  |
| --- | --- |
| **Unit Computing Specialist (student worker), Rutgers UCM** | ***Feb 2023 - Present, New Brunswick*** |

* Wrote Python programs to clean Excel data, eliminating over 50 hours of manual labor every quarter.
* Created Tableau dashboards for budget analysis and IT ticket trends, helping identify and decrease recurrent issues by 70%.

|  |  |
| --- | --- |
| **Data Scientist, Deloitte** | ***June 2019 - Jan 2022, Bangalore*** |

* Reduced banking customer churn by 33% during lockdown by identifying churn customers and key factors using logistic regression in python scikit-learn, while helping the client target strategies toward online features.
* Cut down on over 35 hours of manual labor per week to identify boats on a dock. Used AWS Lambda and CloudWatch to automate the development of a change detection system with OpenCV and Mask R-CNN.
* Boosted a restaurant chain’s take-away orders by over 80% during COVID-19 through customer segmentation using DBSCAN clustering that drove a tailored marketing approach. Developed a Tableau dashboard to present findings.
* Accelerated medical record query speed by over 110% for 800k+ records by establishing an ETL pipeline using AWS Glue and applying Amazon Comprehend for Named Entity Recognition. Leveraged identified entities as tags for superior indexing.
* Mentored interns, reviewed code of peers and followed Agile methodology for project management.
* Was recognized with applause award twice for my client centric work approach and timely deliverables.

|  |  |
| --- | --- |
| **Database Management Intern, MAQ Software** | ***May 2018 - July 2018, Hyderabad*** |

* Implemented an ETL pipeline using SQL Server Management Studio and SSIS, consolidating three large data marts with over 2 million records into one structured format. Developed triggers and stored procedures to identify and rectify table errors and inconsistencies during transfer.

**PROJECTS**

***Chatbot model with a personality:*** [**[ Github ]**](https://github.com/ganeshraj-k/Conversational-model-with-a-personality)

* 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.
* Parallelly fine-tuned Google’s T5 small model with the same dataset for benchmarking purposes.

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

* Conducted Tweedie regression 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.

***Twitter Search****:* [**[ Github ]**](https://github.com/ganeshraj-k/Twitter_search_engine)

* 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 and the web app using Flask.

***Ear image recognition:***

* Achieved 91% accuracy and 73% precision by utilizing a CNN model with SIFT features to implement image recognition on ear pictures for biometric verification. Prepared high-resolution photos of the pupils for the dataset.

**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:** Facilitated access to medical resources during the COVID-19 second wave by updating our website daily with information on available beds, nebulizers, and medicines, and promptly connecting individuals to necessary healthcare services.