|  |  |  |
| --- | --- | --- |
| 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, Selenium. OpenCV, QGIS
* **Cloud:** lambda, S3, CloudWatch, Sage maker, IAM, EC2, Glue.
* **MS Office :** Powerpoint, Excel, Word

**Certifications**

* AWS Machine Learning Specialist
* AWS Cloud Practitioner

**EXPERIENCE**

|  |  |
| --- | --- |
| **Data Analyst, Rutgers University - Communications and Marketing** | ***Feb 2023 - Present, New Brunswick*** |

* Analyzed student enrollment CRM data with Python and summarized data with interactive visualizations using Tableau.
* Prepared well structured excel sheets of the data for ease of future access.
* Utilized Python to preprocess Rutgers Football matches ticket sales data from eVenue.
* Conducted correlation analysis between ticket sales and opponent teams using historical data, identifying key teams that drive sales and support.
* Developed Tableau dashboards to present insights, proposed and implemented targeted marketing strategies, leading to an 87% renewal in season ticket sales and a 34% increase in single game sales the following season.

|  |  |
| --- | --- |
| **Data Analyst, Deloitte Consulting** | ***June 2019 - Jan 2022, Bangalore*** |

**Banking**:

* Mitigated the COVID-19 lockdown-induced customer churn by constructing a multivariate logistic regression model. Identified churn-prone customers and contributing factors.
* Analyzed customer data using Python’s Matplotlib and communicated results to stakeholders, resulting in a 30% churn reduction in the next quarter.

**Restaurant chain**:

* Clustered restaurant customers based on survey data using DBSCAN during lockdown. Correlated the results with their risk tolerance and safety behaviors using Tableau charts.
* This analysis provided a deeper understanding of each customer group. Strategic marketing, along with the promotion of safety protocols, led to an increase in take-away orders, enabling the restaurant chain to adapt to the new market conditions and maintain profitability.

**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.
* Used Git for version control and maintaining code repository.

**Key Achievement:** 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*** |

* 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 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.

***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