GANESH RAJ K

ganesh_012@outlook.com +1-848-313-8525 linkedin.com/in/ganeshrajk/ github.com/ganeshraj-k

EDUCATION

- ***Rutgers University***, master's in data science-Statistics (GPA – 3.75)
  Data Mining, Regression and Time Series Analysis, Statistical Modeling, Data Wrangling, Natural Language Processing
- ***Indian Institute of Technology Indore***, Bachelor of Technology in Computer Science and Engineering

SKILLS AND CERTIFICATIONS

Programming: Python, R, C++ ***Databases:*** Postgres, tSQL, MongoDB ***Data Visualization:*** Tableau, PowerBI, MS Excel
Libraries: scikitlearn, pandas, numpy, seaborn, matlplotlib ***Frameworks***: PyTorch, Tensorflow, Flask ***Cloud:*** Sage Maker, S3, EC2, Glue, Lambda, Github

- AWS Machine learning Specialist Certification
- AWS Cloud Practitioner Certification

PROFESSIONAL EXPERIENCE

Data Analyst, Deloitte *June 2019 – Jan 2022, Bangalore*

- Identified banking customers at risk of churn during the lockdown using a Logistic regression model. Carried out in-depth analysis using Tableau dashboards paired with model findings to understand demographics likely causing the churn. Achieved an accuracy of 75% and recall of 78%. Communicated insights to clients and helped them reduce churn in the following quarter by 43%.
- Automated boat detection in a private harbor using NNU-Net for image segmentation and OpenCV for image differencing on segmented masks. Procured high-res GEOTIFF satellite image using LandViewer API. Cleaned, dehazed the images to remove fog. Extracted geo-location of boats using geotiff. Hosted and scheduled end to end process on AWS to be executed hourly. Reduced 30 hours of work per week to less than 3 hours by eliminating the need for manual tracking.
- Segmented restaurant chain customers using K means clustering, conducting RFM analysis on transaction records and the cluster centroids. Analyzed customer feedback on dining preferences and lockdown caution using Tableau dashboards and correlated the observed patterns with cluster data. Presented clients with actionable insights for developing targeted marketing strategies tailored to each customer segment.
- Developed an Employee Search app that extracts text from employee resumes, cleans it, and enables text-based search using R Shiny for the UI as a personal proof of concept (POC).
- Processed medical records with AWS Glue and applied AWS Comprehend for Named Entity Recognition, enabled more efficient record query speed and medical case segmentation.

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

- Created and implemented an ETL pipeline in SQL Server Management Studio and SSIS for cleaning, structuring, and consolidating three Data Marts into one. Also wrote triggers in place and stored procedures to identify and resolve errors and inconsistencies in tables while transferring.

UCM Rutgers, Unit Computing Specialist **Feb 2023- Present, New Brunswick**

- Developed Python scripts to automate text dataset cleaning, reducing manual efforts.
- Administered users in the workgroup using Active Directory Software. Managed over the internet software installations via Kace.

PROJECTS

Chatbot model with a personality:

- Developed a conversational model mimicking Chandler Bing's dialogue style from "Friends" using a seq2seq model with LSTM.
- Concurrently, fine-tuned Google's T5-small model on the same dataset for benchmarking and comparison purposes.

2023 Travelers Insurance Analytics University Contest:

- Developed a **Tweedie regressor model** for a zero-inflated claim cost dataset. To evaluate its performance, used the **Gini index metric** and performed **grid search** for parameter tuning.
- As a result, secured **third place** in the contest and was recognized as the **campus winner** for accurately predicting insurance claims.

Music Sentiment Analysis:

- Performed sentiment analysis with R on four artists' albums to identify key emotions in their music, to get a deeper understanding of their style.
- Also used metadata from the songs to create charts depicting danceability, valence, tempo, and major key notes for each album. Extracted lyrics for each song using GeniusR and Spotify APIs.

S&P 500 Time Series Analysis:

- Generated visualizations in R to analyze sector-by-sector trends of the S&P 500 during 2019-2020. Also graphically visualized weekly and monthly closing trends, along with inter-sector comparisons.
- Applied an auto ARIMA model for deeper seasonal analysis and future index value forecasting.

Twitter Search:

- Created a web application to search tweets, users, and hashtags in a tweet database. Hosted tweets on MongoDB for rapid querying and user profiles on PostgreSQL.
- Integrated synonym and minimum edit distance search methods. Utilized Flask for API requests and web app management.

NCAA Player Database analysis:

- Utilized BigQuery to query and analyze NCAA basketball data. Extracted insights on team performance and player trends, presenting them effectively.

College Leave Application:

- Developed a Web Application using RESTful API that had features to submit, track, approve leave applications.

Bounce 3D game:

- Developed a basic version of the classic bounce game using C++ and Cgraphics for a course project.

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:** Reached out to hospitals, clinics, and medical suppliers to collect information on available beds, nebulizers, and medicines by making calls every day during the COVID-19 second wave. Kept this data updated on our website so people could easily find what they needed based on where they lived. Also helped people by quickly connecting them to the closest medical resources, making sure they got the help they needed.