PRANEITH RANGANATH

Boston, MA | +1 8574130507 | ranganth.p@northeastern.edu | LinedIn

EDUCATION

NORTHEASTERN UNIVERSITY Boston, MA
Master's in Data Analytics Expected Dec 2024
Relevant Coursework: Datawarehouse and Business Intelligence; Data Mining; Statistical Learning; MLOp's; NLP

Dr Ambedkar Institute of Technology Bengaluru, India
Bachelor's in Information Science and Engineering Aug 2014 - Jun 2018

SKILLS

Machine Learning Frameworks: Pandas, NumPy, TensorFlow, PyTorch, Scikit-learn, BeautifulSoup, Scrapy
Databases and Datawarehouse Tools: MySQL, PostgreSQL, MongoDB, Snowflake, Talend
Cloud and Visualization Tools: GCP, AWS, Tableau, Power BI
MLOps and DevOps: Airflow, MLflow, TFX, DVC, Docker, Kubernetes, GitHub Actions
Automation Tools: Automation Anywhere, UI Path, Microsoft Power Automate, RulAI

WORK EXPERIENCE

Business Development Analyst Intern, FHI 360, Washington Jun 2023 – Nov 2023
- Integrated data from USASpending, foreignassistance.gov, Gates & Melinda Foundation into a centralized data warehouse for uncovering market trends and identifying potential competitors.
- Engineered a Python framework to extract Congressional Budget Justification data which provided a strategic forecast of project outlook and funding allocations for upcoming fiscal years in specific regions and countries through dashboards.

Senior Automation Developer, Acronotics, Bengaluru, India Jul 2019 – Sep 2022
- Developed bots to automate data collection, transformation and creation of custom dashboards which ensured constant and real time update of data thereby decreasing turnaround time by 75% using tools such as Automation Anywhere, Python, MySQL, Tableau, Power BI and Excel Power Pivots.
- Used OCR to capture invoices generated and received in real time and automated financial analysis report generation on a day-to-day basis thereby eliminating any dependency

System Engineer, Tata Consultancy Services, Bengaluru, India Jun 2018 – Jul 2019
- Automated mundane repetitive tasks using Automation Anywhere, UiPath, Python and VB Scripts increasing efficiency and accuracy of process by 75% and 90% respectively.
- Designed Pipes to seamlessly ingest data from AWS S3 buckets into Snowflake and leveraged Streams and Tasks to automate ETL pipelines, ensuring efficient processing for analysis.

ACADEMIC PROJECTS

Brain Tumor Classification – Deployment ready product Jun 2024
- Implemented CI/CD using Git Actions, created the latest docker image and stored in GCP artifact registry
- Kubernetes picked this docker image and hosted it on one of the pods in one of the nodes in the cluster
- Used FastAPI in backend for endpoint communication and Streamlit for frontend application
- Implemented DVC, model versioning, tracking using MLFlow and automated pipelines using Airflow DAG's
- Designed a monitoring dashboard from parsed logs to identify data drift, data skew, confidence of model

Room Occupancy Estimation Apr 2024
- Developed machine learning algorithms - Linear Regression with closed room solution and gradient descent, LDA, QDA and Support Vector Machines with Sequential Optimization technique utilizing NumPy pandas, scipy optimize in Python for a multi-class classification problem.
- Applied techniques such as square root, cube root, logarithmic transformations to ensure data follows normal distribution which is best suited to give optimal results when fitted to model.

Health Centre Database and Datawarehouse for Analysis Apr 2024
- Designed a multidimensional model with several facts such as Consultation, Tests Conducted, Operations performed to analyze metrics of the health center and identify potential hazards around certain diagnosis and geographical regions.
- Implemented the OLTP and OLAP databases using PostgreSQL, used Talend Jobs for ETL operations.

MBTA – Predicting the load on a bus Apr 2023

- Trained supervised machine learning models - Linear regression, KNN, Random Forest and Decision trees using Scikit-learn on MBTA data to predict the load currently on the bus.
- Based on the accuracy of the various ML models we choose Random Forest model to be the most accurate with an accuracy of 82% and with further tuning of hyperparameters we choose four decision trees to be the most efficient increasing the accuracy to 84.8%

Football/Soccer Database management system Apr 2023

- Architected and implemented a highly scalable and robust database system for the European football/soccer league, leveraging MySQL and MongoDB. Ensured data security, consistency, and validation, while designing and implementing efficient ETL pipelines using Snowflake for data analysis.
- Integrated MySQL database with Python using SQLAlchemy, enabling data analysis for extracting valuable insights.

Formula1 – Dashboard and Season Analysis Dec 2022

- Extracted and integrated data from various sources and API's, utilized Python for data wrangling and created a comprehensive dashboard displaying the 2021 season insights using Tableau.
- Utilized Exploratory Data Analysis techniques to uncover the various factors which played out in 2021 season resulting in one of the best battles for the championship.