FAZAL HYDER SHAIK
fazalhyders@gmail.com | Atlanta, GA | 773-565-5665 | LinkedIn

SUMMARY

Highly self-motivated and detail-oriented individual who can multi-task and communicate effectively Communicates business insights through data visualization, verbal, and written communication for decision making. Demonstrates leadership, Interpersonal skills, Presentation skills, problem-solving skills, and ability to work in fast-paced, dynamic environment.

EDUCATION

Mercer University — Atlanta, GA
Master of science, Business Analytics — Jan 2023 - May 2024

- Courses: Data Mining, Data Visualization (Tableau, Power BI), Advanced Statistics (R studio, Python, SQL), Data-Driven insights, Management Analytics, Marketing Analytics, and Financial Analytics.

JNTUH — Hyderabad, TG
Bachelor of Technology, Computer Science Engineering — Jul 2018 - Aug 2022

- Courses: C programming, C++, Java, Machine Learning Algorithms, Data Science, Data Mining, Cloud Computing, Database Management (SQL), Data Structures, SDLC etc.

SKILLS

Competencies: Advanced Statistical Analysis, Data Mining, Data Models, Advanced Statistical Skills, Data Analysis, Predictive Modeling, Model Building, Model Evaluation,ETL, GenAI, NLP,OpenAI Data Warehouse, ChatGPT 4o, ML and AI, GPT Models Creation ,Data Engineering, Ad- hoc Analysis, A/B Testing.

Programming Languages: Python, R Programming, SQL, Java, C, C++, HTML, Linux.

Tools/Frameworks: PyCharm, Anaconda, Jupyter , Google Colab, Looker, CPLEX, VS Code, AWS, Azure Databricks, SciKit-Learn, NumPy, Pandas, TensorFlow, PyTorch, MS Office , GitHub, CICD, SAS,VBA.

EXPERIENCE

Data Science and Machine Learning Intern — REGex Software Services
May 2021 – Aug 2021

- Leveraged R and SQL within Microsoft SQL Server and R Studio to refine and correct inconsistencies across extensive sales datasets, ensuring 95% data accuracy.
- Applied statistical methods and machine learning techniques, including classification algorithms, to identify and analyze sales trends, increasing forecast accuracy by 30%.
- Developed data pipelines using SQL and Python (pandas, SciPy) to efficiently monitor, extract, and manage sales data, reducing data retrieval times by 25%.
- Designed and created data visualizations, models, and infographics, which facilitated strategic decision-making by presenting complex insights in an accessible format to non-technical stakeholders.

PROJECTS

Hospital Readmission Prediction using Machine Learning — Mercer University
Jan 2024 – May 2024

- Achieved 77% accuracy in predicting hospital readmissions using logistic regression, random forests, k-means clustering and Bayesian methods, and neural networks on Diabetes Dataset from healthcare systems, reducing readmissions by 15%.
- Improved model performance by 10% through advanced feature engineering and PCA for dimensionality reduction and SMOTE-TOMEK for sampling the minority class.
- Developed a Python and SQL-based data pipeline, processing over 130,000 patient records and reducing data prep time by 30%. Used ggplot, ggplot2, plotly, matplotlib and seaborn for data visualization.
- Created a generative LLM integrating Chat GPT & a web page using HTML/CSS/JS which would predict patient readmission and advise precautions to be taken to avoid patient readmission using coefficient of variables which significantly reduced readmissions up to 30%.