Farah T. Ahmed

farahtasnimahmed800@gmail.com | +1(213)352-2229 | GitHub
linkedin.com/Farah-Ahmed | US Citizen | Brookfield, WI

Work Experience

Data Science Intern | *Spring Venture Group, Kansas City, MO* June 2024 – August 2024

- Enhanced an agent-matching model that initially relied on a single feature, adding 20+ features to better align customers with agents, improving accuracy by 3.6% over the baseline.
- Defined the scope of the binary classification model and analyzed its potential business impacts, ensuring the training data closely mirrored production data to optimize performance.
- Applied gradient boosting models to capture complex relationships in large datasets, tuning hyperparameters using SageMaker's distributed computing resources for better model performance.
- Selected appropriate metrics like Brier score and log loss for calibration, and ROCAUC and PRAUC for discrimination.
- Employed Redis to cache historical data for fast production deployment. Wrote scripts to fill redis cache, using pydantic for data validation to ensure data integrity and reliability.

Technical Skills

Data Science & Analytics: Data Analysis, Data Visualization, Data Wrangling, Predictive Modeling, Statistical Analysis, AWS, Machine/Statistical Learning, Neural Networks.

Programming Languages: Python, SQL, R.

Skills/Packages/libraries: XGBoost, mySQL, Sagemaker, AWS s3, redis, pydantic, Jupyter, Git(version control tool), matplotlib, PyTorch, scikit-learn.

Education

Master of Science in Data Science Aug. 2022 – Dec. 2023
The University of Iowa; Iowa city, IA

Bachelor of Science in Mathematics Aug. 2019 – Dec. 2021
California State University of Northridge; Northridge, CA

Relevant Projects

Predicting Used Car Price Using Machine Learning | GitHub

- Worked as a team player to implemented diverse regression models, evaluated their performance, and identified the optimal predictive model to estimate used car prices accurately.
- Achieved an R-square score of 0.89, Tackled data inconsistencies and missing values through thorough data cleaning techniques, demonstrating problem-solving skills in handling real-world data challenges.

Travelers Modeling Competition (Kaggle) | GitHub

- Explored various predictive models including Linear, Ridge/Lasso, and Tree-based models to predict auto insurance claim costs, evaluating model efficacy and performance.
- Performed data cleansing, preprocessing, transformation, and and feature engineering with careful attention to detail to ensure data quality and relevance for analysis, Achieved a Mean Squared Error (MSE) of 0.007.

Infographic Analysis: Criminal Behavior Patterns and Trends in Iowa | GitHub

- Conducted exploratory data analysis (EDA) on *Iowa Prison Admissions* and *Offenders Released* datasets data to identify trends and patterns using data analysis tools, demonstrating strong analytical and quantitative skills.
- Performed data manipulation, preprocessing, and statistical techniques and data visualizations to extract reliable insights, prepared reports and present the findings through communication.

T1w-T2w MRI image Translation Using GANs | GitHub

- Implemented a Conditional Generative Adversarial Network to synthesize paired MRI images, specifically targeting the T1w-T2w modality, achieving a SSIM of 80-81% between ground truth and generated images.
- Utilized Python, PyTorch, and Torchvision for model development, alongside extensive data manipulation and visualization using Numpy, Pandas, and Jupyter Notebook.