MUSTAAQ AHMED KHAN

New Haven, CT, 06519 | +1 (203)627-1083 | muskhan@my.bridgeport.edu

EDUCATION

University of Bridgeport | Bridgeport, CT — Expected Dec 2025
Master of Science in Computer Science
Coursework: C++, Python for Data Science, Big Data, Machine Learning, Deep Learning, Cloud Computing, Advanced Databases, Operating Systems, Web Based Application Development

Osmania University | Hyderabad, India — Jun 2023
Bachelor of Engineering, Electronics and Communication Engineering

TECHNICAL SKILLS

Programming Languages: Python; C; C++; R programming; HTML; CSS; SQL
Tools: Power BI/Tableau; MS-Excel; Scikit-learn; TensorFlow; Keras; OpenCV; Pandas; MATLAB

WORK EXPERIENCE

Delta Sigma Technologies | Hyderabad, India — Jun 2022 - Nov 2022
Data Science Intern

- Worked on tasks typical of a Data Scientist, including optimizing data processing workflows, reducing processing time by 90%, from 10 hours to just 1 hour, by implementing efficient algorithms and leveraging parallel computing across multiple systems.
- Analyzed 500K+ data points weekly to enhance machine learning models for predictive analytics.
- Enhanced decision-making processes by 30%, presenting findings with 90% accuracy using advanced statistical techniques and interactive data visualization tools, directly impacting business-critical outcomes.

ACADEMIC PROJECTS

Social Media Sentiment Analysis for Brand Reputation Management

- Developed and optimized machine learning models (BERT, Random Forest) to classify social media sentiment with 89% accuracy, enabling real-time insights into public opinion for over 1 million tweets on brand mentions and competitors.
- Built scalable NLP pipelines using NLTK and TensorFlow, automating data cleaning, pre-processing, and model deployment, resulting in a 30% increase in processing efficiency and faster response times.
- Enhanced model precision through advanced hyperparameter tuning and k-fold cross-validation, improving F1 score from 0.87 to 0.91, ensuring reliable sentiment insights across diverse social media platforms.

New York City Taxi Fare Prediction

- Developed and fine-tuned machine learning models (XGBoost, SVM), boosting prediction accuracy by 25% and reducing error margins by 15%, ensuring precise fare estimates for over 500,000 taxi trips across varying conditions.
- Designed scalable machine learning pipelines using Scikit-learn and TensorFlow, streamlining data preprocessing, model training, and deployment, cutting processing time by 20%.
- Improved model robustness by applying cross-validation and hyperparameter tuning, reducing error margins from 5.22 to 5.05, ensuring consistent results across diverse datasets.

Risk Analytics for Loan Default Prediction

- Conducted EDA on financial loan data to identify key factors influencing defaults, reducing approval errors by 30% and enhancing risk assessment. Applied correlation analysis (CORREL, heatmaps) to determine top predictors, improving classification accuracy by 15%
- Handled missing data using Excel functions (COUNT, AVERAGE, MEDIAN) and identified outliers with IQR, quartiles, and conditional formatting, improving data integrity by 25% and refining loan approval criteria.
- Assessed data imbalance using COUNTIF, SUM, and bar charts, providing key insights into credit risk modeling and enabling better decision-making for high-risk applicants.

CERTIFICATIONS

- **Trainity Data Analytics Virtual Internship (Feb 2025) –** Gained hands-on experience in EDA, visualization, and business insights using Excel, SQL, and Power BI.