MAHMUDA YASMIN

mahmuda.yasmin@gmail.com | (248)-752-7781, (226) 246-0226 | LinkedIn | Github | Portfolio

PROFESSIONAL SUMMARY:

Results-driven Data Scientist with over 6 years of experience specializing in data visualization, machine learning, and predictive analytics. Proven track record of transforming complex datasets into actionable insights to drive decision-making and optimize performance. Skilled in developing machine learning models and visual dashboards that reduce risks and improve efficiency. Passionate about leveraging data to solve real-world challenges in dynamic, fast-paced environments.

SKILLS:

Data Manipulation and Analysis Tools:	SQL, Databricks, Power Query, MS Excel (Solver, VLOOKUP, PowerPivot), Apache Spark (PySpark)
Data Visualization Tools:	Power BI (DAX/M), Tableau, Python (Plotly, Dash, Seaborn, Matplotlib), R (ggplot2)
Programming Languages and Libraries:	Python (Pandas, NumPy, Scikit-learn, Keras, PyTorch, TensorFlow, sklearn scipy), R
Machine Learning and Deep Learning Frameworks:	Supervised (Random Forest, SVM), Unsupervised (K-means, PCA), Neural Networks (CNN, RNN)
Deep Learning & NLP:	TensorFlow, Keras, PyTorch, Sentiment Analysis, Topic Modeling
Cloud & Pipeline Tools:	Azure ML Studio, Amazon SageMaker, Git.

EXPERIENCE:

RIIPEN

Data Visualization Specialist (Intern), Client - Zoptic

July 2024 – Oct 2024

• Led the data analysis and statistical modeling of sports performance data, optimizing athlete performance metrics for injury prevention and risk mitigation.

• Created interactive dashboards for coaches using Power BI, Excel, and Python, summarizing key metrics such as speed, acceleration, and deceleration for real-time athlete assessment.

• Automated data processing workflows and visualizations using Python and SQL, reducing manual reporting time by 40% and improving the accuracy of performance assessments.

• Designed and implemented machine learning pipelines for predictive modeling (regression, classification) and anomaly detection to flag deviations in performance and reduce injury risks.

SPRINGBOARD

Data Science Trainee

Jan 2023 – Feb 2024

• Developed proficiency in Python, SQL Programming and Git; focused on supervised and unsupervised learning, and data visualization using Python libraries, PowerBI and Tableau.

• Completed 600+ hours of hands-on coursework with 1:1 expert mentor oversight.

• Executed two in-depth portfolio projects: predictive modeling on Consumer Churn data for a Telecom Company and Sentiment Analysis on Restaurant Reviews from TripAdvisor.

• Topics Extensively covered - Python, Software Engineering Principles, Data Wrangling, Inferential Statistics, Time Series Analysis, Recommender Systems, Customer Segmentation using Clustering, MapReduce with Spark, Social Network Analysis etc.

• Capstone 1: A project on developing ML Model on Predicting Customer Churn from a Telecom Dataset.

Developed a Machine Learning model for predicting Customer Churn from a Telecom Dataset. Analyzed factors contributing to customer churn and provided actionable business insights such as- Talk time, Internet, Competitor offers etc. contributes on predicting Customer Churn.

• Capstone 2: An NLP project on Sentiment Analysis on Reviews on Restaurants from TripAdvisor Dataset.

Conducted an NLP project on Sentiment Analysis of Restaurant Reviews from TripAdvisor. Implemented Topic Modelling to identify key topics and predict sentiments. Findings indicate that - reviewers use more words to leave negative comments and the mostly bad service-related issues trigger negative reviews.

PTI QCS, DETROIT, MI
Data Visualization Analyst
Jan 2016 – Dec 2022

• Spearheaded business forecasting, dashboard development, and report generation, enhancing decision-making capabilities across workforce and inventory management.
• Designed and optimized workforce and inventory management plans, resulting in a 15% improvement in resource utilization through data-driven insights.
• Conducted rigorous statistical tests (A/B testing, hypothesis testing) to assess service processes, driving the identification and optimization of key performance indicators (KPIs).
• Utilized time series analysis to accurately forecast future workload and optimize inventory levels, reducing stock outages by 20%.
• Leveraged advanced data tools (R, Python, Power BI, Tableau, Apache Spark) to streamline data processing and automate dashboard updates, improving reporting efficiency by 30%.

EDUCATION:

SPRINGBOARD
March 2024
Data Science Career Professional Certification

University of Windsor - Windsor, ON Canada
Master of Science: Statistics
Aug 2014

• M.Sc. Thesis Topic: Efficiency and Coverage Probability of the Over-Dispersion Parameter in Clustered Binomial Data", 2014.
• Master's International Entrance Scholarship Issued by University of Windsor, Jan 2013.
• Former Co-Ordinator of Centre for Teaching and Learning (CTL) resources for Graduate and Teaching Assistant, Dept of Math and Stat, UofWindsor.

INTERESTS:

Social Activism, Modern Art and Photography, Healthy Living enthusiast.