Shemonto Das

St. John's, NL, Canada | shemonto21@gmail.com | +17099867573 | shemontodas.com | linkedin - shemontodas |

Highlights of Qualification

Skilled Data Scientist with an MSc in Computer Science and 2+ years of experience in research, data analytics, statistical and machine learning model development, and automation at Instrumar. Proficient in deriving data insights, presenting complex findings to stakeholders, and delivering solutions aligned with business goals.

EDUCATION

MSc. in Computer Science, Memorial University of Newfoundland, St.John's, NL — May 2024
BSc. in Computer Science & Engineering, BRAC University, Dhaka, Bangladesh — Jan 2021

SKILLS

- **Language and Frameworks**: Python, Java, HTML, SQL, Pycaret, Tensorflow, Keras, Pandas, Scikit-learn, JSON.
- **Soft skills**: Scientific Research,Academic writing, Presentation, Critical thinker, Team player, Leadership.

Experience

Data Quality Specialist, **SiftMed** – St. John's, NL, Canada — October 2024 – Present

- Streamlined digital document organization, labeling accuracy and proactively resolved data quality issues.

Data Scientist, **Instrumar Limited** – St. John's, NL, Canada — June 2024 – September 2024

- Analyzing and exploring time series to develop, optimize, and interpret machine learning models.
- Developing and automating scalable pipeline for data querying, cleaning and feature engineering.
- Reduced data annotation cost and time for Instrumar by 40%, making fault classification efficient.

Data Science Intern, **Instrumar Limited** – St. John's, NL, Canada — May 2022 – April 2024

- Interpreted complex manufacturing data into actional insights, improving data-driven decision-making.
- Optimized data preprocessing and feature engineering to enhance anomaly detection and classification.

Projects

Predicting Cross-sell Insurance: Responsible AI Project (Python, Scikit-learn, Imbalanced-learn) (Link)

- Developed a predictive model for cross-selling health & vehicle insurance, earning 2nd place in the competition.
- Enhanced model interpretability by integrating Responsible AI tools like LIME and SHAP, ensuring transparency

Preprocessing & Data Labeling Pipeline (Python, Pandas, Matplotlib, Tkinter) (Link)

- Boosted event retrieval accuracy and annotation rate by 50% via streamlined sensor data preprocessing.
- Created an interactive interface with Matplotlib and Tkinter, enhancing event annotation and analysis efficiency.

Instrumar's Machine Learning Pipeline for Fault Classification (Python, Pycaret, Scikit-learn, Graphviz) (Link)

- Reduced quality downgrades by 30% through automated fault classification, improving detection efficiency.
- Developed a novel frequency-based algorithm to tackle data imbalance, boosting accuracy for Instrumar.
- Ensembled and stacked ML models to enhance performance and predictive accuracy.

Active Learning Framework for Texture Classification (Python, Scikit-learn, Pandas, Scipy) (Link)

- Preprocessed tactile sensing data and conducted cluster and regression analyses to enhance predictive insights.
- Validated machine learning model results for texture classification using Wilcoxon hypothesis testing.

AWARDS

- Awarded "Fellow of the School of Graduate Studies" title by Memorial University, 2024.
- Graduate Research Award 2023 for strong research contribution, Computer Science, Memorial University.
- Presented accepted research paper at IEEE Systems Conference 2024, Montreal, Canada.