Sushmita Das

✉ sushmitd@alumni.cmu.edu   in linkedin.com/in/sushmita-das09   🌐 sushmitadas.github.io   ☎+1-412-641-0659

EDUCATION

Carnegie Mellon University (CMU)   3.96/4.0
MS in Electrical and Computer Engineering   Jan. 2021 – May. 2022

Assam Engineering College   81.7/100
BE in Electrical Engineering   Aug. 2016 – Sep. 2020

CURRENT EXPERIENCE

Data Scientist, Meltwater | *Python, Time Series Analysis, Forecasting, Anomaly Detection, DevOps*   Jun. 2022 – Present
- Developed ML models for Meltwater's media intelligence platform to understand social media engagement and mention trends, customer behavior, identify and explain anomalies and make predictions about future trends.
- Helped deploy scalable ML models using FastAPI and AWS SageMaker to serve 27,000 corporate clients across 120 countries.

PROJECTS

Real-Time Spike Detection and Analysis (In Production) | *Python, Anomaly Detection, LLM*   Aug. 2023 - Oct. 2023
- Developed and deployed algorithms for anomaly detection and for identifying root causes of spike in social media activity.
- Additionally, utilized GPT 3.5 through OpenAI API to offer concise explanations and summaries for spike-driving articles and tweets, enhancing user insights for decision-making.

Forecasting Social Media Engagement (In Production) | *Keras, Scikit-learn, Statsmodels*   Jan. 2023 - March. 2023
- Retrieved and analyzed real-life large datasets from various social media engagement sources, capturing significant global events.
- Evaluated LSTM, Prophet, MLP, Linear regression and EWMA. Deployed ARIMA for its balanced performance and simplity. It achieved desirable MASE for 80% of the cases over a 2-day forecast horizon, which is in production for active use and feedback.

Predicting Customer Retention | *Python, Pandas, Numpy, Scikit-learn, SQL*   Nov. 2022 - Nov. 2022
- Collaborated with cross-functional teams to collect and clean real-life company data, incorporating 30 user analytics features.
- Built a Logistic Regression model with 78% test accuracy to predict customer churn, reducing the churn rate and enhancing customer retention strategies.

Weather vs. Electricity Consumption Data Analysis | *Python, Pandas, Numpy, Scikit-learn, Excel* Jan. 2022 - Jan. 2022
- Performed data cleaning and analyzed one year of weather and electricity consumption data from France.
- Used Multiple Linear Regression for forecasting energy consumption, facilitating informed resource allocation and strategic decision-making.

PUBLICATION

Sushmita Das, Ankur Deka, Yuji Iwahori, M. K. Bhuyan, Takashi Iwamoto, Jun Ueda, *Contour-Aware Residual W-Net for Nuclei Segmentation*, KES 2019 special session on Computational Intelligence System and Applications.

PREVIOUS EXPERIENCES

Research Assistant, Carnegie Mellon University | *Python, SQL, Tableau*   Jan. 2022 – May. 2022
- Analyzed real data on Nigeria's electricity consumption to aid in the regulation of power-cuts.
- Conducted data cleaning, analyzed continuous time periods, and provided comprehensive summary statistics.

Research Assistant, Carnegie Mellon University | *Image Segmentation, Computer Vision, Deep Learning* Sep. 2021 – Dec. 2021
- Helped in creating a new dataset to perform Bird Eye View (BEV) Satellite Image Segmentation to enhance drone navigation.
- Modified the W-Net model for multiclass road segmentation, surpassing the performance of the baseline U-Net model.

Summer Intern, Carnegie Mellon University | *Deep Learning, Computer Vision*   Jun. 2021 – Aug. 2021
- Implemented a driver alert system on Nvidia Jetson Xavier NX, integrating Mask RCNN, YoloNet, and MobileNet object detection models.

Research Intern, Chubu University | *Medical Image Segmentation, Computer Vision*   Sep. 2018 – Mar. 2019
- Developed a deep learning based model, Contour-Aware Residual W-Net (WRC-Net), consisting of double U-Net for nuclei segmentation and evaluated it on real Hematoxylin and Eosin stained cell images.
- It showed better overall performance (higher dice coefficient) against previous nuclei segmentation methods.

TECHNICAL SKILLS

Languages: Python, R, SQL, C/C++.
Data Analytics: SAS Enterprise Miner, Excel.
Visualization Tools: Tableau, Streamlit.
DevOps: Docker, Kubernetes, FastAPI, Flask.

Developer Tools: Jupyter Notebook, R Studio, Matlab, Octave.
Libraries: Scikit-learn, Pandas, NumPy, Matplotlib, Keras, PyTorch.
Cloud: AWS, GCP, Heroku.
Operating Systems: Linux, MacOS, Windows.

COURSES

Intro to ML (Python)	Pattern Recognition Theory (Python)	Neural Signal Processing (Matlab)
A/B Testing (R)	BI and Data Mining (SAS Enterprise Miner)	Google Professional Data Analytics (SQL, Tableau, Excel, R)
Data Analytics (Py)	Big Data Science (Python)	Data Science for Technology, Innovation and Policy (R)