

Sulakshana Chakraborty

Data Scientist

+44 7765720314 | sulakshana.ch2014@gmail.com | [linkedin.com/in/sulakshana-chakraborty/](https://www.linkedin.com/in/sulakshana-chakraborty/) | github.com/SulakshanaChakraborty | E3 4HF, London

SUMMARY

With a record of academic excellence, a solid foundation in theoretical machine learning, coupled with over 5 years of hands-on experience, I excel at leveraging data to drive informed decision-making and creating production grade ML solutions.

EXPERIENCE

Data Scientist

Jan 2023 – Present

Argus Media – Provider of price benchmarks and financial analysis of commodity markets

London, UK

- Built and deployed a sub-editing tool for the editorial team using LLMs and NLP techniques, expediting the news editing process through text evaluation, suggestions, edits, and detailed sentence-by-sentence insights.
- Created an analytics dashboard to monitor sub-editing tool feedback and sentiment, continuously refining the algorithm based on user input.
- Implemented a retrieval-augmented generation system using Vector Databases to facilitate news stories retrieval, empowering reporters with extractive summarisation, timeline summaries, and relevant price time series searches.
- Demonstrated the AI sub-editing tool to multiple global editorial teams, conducted detailed feedback sessions, and developed a comprehensive rollout plan and analytics, enhancing tool adoption and integration across editorial groups.
- Extracted sentiment from published news data for crude oil, generating sentiment time series as features for crude oil price forecasting models, adding a qualitative dimension to quantitative analysis.
- Created interactive visualizations for historical fundamental time series and developed real-time scenario analysis functionality for forecasting models on the Argus Risk Management Platform, enabling dynamic risk assessment.
- Modeled front-month gasoline contract prices using Generalized Additive Models for Location Scale and Shape (GAMLS), and deployed the probabilistic forecasting model to production.

Machine Learning Engineer, Intern

Sep 2022 – Dec 2022

GreenSwapp – Climate Impact Tracker API

London, UK

- Web scraped data, collated training sets, built, and deployed a semantic search model to estimate carbon footprint values for food items and recipes.
- Developed techniques using LLMs to facilitate automated extraction of carbon footprint values from research papers.

Senior Business Technology Analyst

Jan 2021 - Jun 2021

Business Technology Analyst

May 2018 - Dec 2020

ZS Associates - Consulting, technology and analytics focusing on healthcare.

New Delhi, India

- Led the end-to-end development of a sales reporting dashboard for a rare disease drug across 30+ countries, featuring interactive visualizations, KPI tracking, and performance summaries for new region launches.
- Oversaw the development of big data pipelines on AWS, designing data models for efficient analytics storage and ensuring seamless data ingestion, transformation, filtering, and aggregation.
- Managed a DevOps team for an analytics dashboard, handling Airflow jobs for weekly data refreshes, addressing client requests, and ensuring smooth operational workflow.
- Built a sales rep evaluation platform, enabling valuation score collection, hierarchical approval logic, incentive data creation, and incentive payouts for approximately 2,000 employees.
- Prepared project plans, finalized requirements and timelines with key stakeholders, led weekly prioritization discussions, tracked team bandwidth, and facilitated client demos.

TECHNICAL SKILLS

Languages: Python, R, SQL, HTML/CSS, Bash, C, MATLAB

Tools: Git, VS Code, Jupyter, RStudio, Airflow, GitHub Actions, Docker, ChromaDB, MySQL, Spark, MLFlow

Libraries: LangChain, SpaCy, Data.Table, Plotly, PyTorch, Pandas, Numpy, TensorFlow, Keras, Sklearn, Xgboost, Huggingface, LangChain, FastAPI, RShiny, PySpark, gamlss, statsmodels

Cloud: AWS - S3, EC2, EMR, Lambda, RDS, Sagemaker

Expertise: NLP, Recommender Systems, Deep Learning, Statistical Modeling, Timeseries Forecasting, Large Language Models - (GPT, Llama), Prompt Engineering, API development, ETL, CI/CD, MLOps, Agile

EDUCATION

University College London

MSC Data Science and Machine Learning — 76.58 (Distinction)

London, UK

2021 - 2022

Vellore Institute of Technology

BTech Electrical and Electronics Engineering — 8.67 GPA

Vellore, India

2014 - 2018

PROJECTS

Generating Expressive Animated Avatars - Emotech | *Python, TensorFlow, OpenCV, Librosa, Praat* 2021

- Engineered pipelines to process large-scale video datasets to efficiently extract facial key points and acoustic features per video frame
- Implemented a generative model which outputs a sequence of animated facial key points aligned with a given audio sequence.
- Improved baseline results by using acoustic novel features that capture visual prosody (emotions). Achieved state-of-the-art results in the synthesis of expressive digital heads.

Neutralizing Sexist Language | *Python, PyTorch, NLTK* 2021

- Contributed to a new sexist dataset by incorporating 2500 new sexist, non-sexist pairs of tweets
- Finetuned bias neutralization model with BERT tagger and bi-LSTM Seq2Seq autoencoder
- Found 2 novel ways of evaluating the similarity between pairs of sentences

Multi-task Learning for Image Segmentation | *Python, PyTorch* 2018

- Built multi-task learning with attention masks (MTAN) architecture with VGG SegNet as the main task and bounding box detection and classification as auxiliary tasks.
- To further improve performance, experimented with three additional self-supervised auxiliary tasks and conducted ablation studies. Achieved 0.88 IOU.

PUBLICATIONS

Implementation of Facial Recognition for Home Security Systems

Published in International Journal of Engineering and Technology (IJET), 2018

CERTIFICATIONS

Deep Learning Specialization – DeepLearning.AI Issued in 2021

Data Science Professional Certificate – IBM Issued in 2021