

SULAKSHANA CHAKRABORTY

+44 7765720314 | sulakshana.ch2014@gmail.com | <https://github.com/SulakshanaChakraborty> | <https://www.linkedin.com/in/sulakshana-chakraborty/> | F6 Oak house, N22 8YT, London

OVERVIEW

Data Scientist with 4 years of experience in advanced analytics, data engineering and machine learning. UCL Master's graduate with a strong theoretical background in data science and research experience. Proven track record in developing technological solutions for fortune 500 companies. Capacity to work in multidisciplinary teams efficiently in fast-paced, intensive environments.

EXPERIENCE

ARGUS MEDIA

Data Scientist, Analytics

London, United Kingdom

January 2023 – Present

- Leveraged **cutting-edge forecasting models** to accurately predict oil commodity prices. Utilized the **GAMLSS framework** for distributional regression.
- Strengthened model performance evaluation by enhancing existing diagnostic reports with **statistical tests and informative visualizations**. Communicated results effectively to stakeholders.
- **Streamlined and optimised in-house forecasting codebase** to improve functionality and increase production readiness. Implemented feature enhancements and speed improvements.
- Extended the current forecasting framework and **spearheaded the development of ML pipelines** (data ingestion, model selection and inference) for new commodities.

Tech Stack: R, AWS, MySQL, Oracle, Python, RShiny

GREENSWAPP

Machine Learning Engineer, Intern

London, United Kingdom

September 2022 – Dec 2022

- **Architected and led the development of a state-of-the-art information retrieval model** for accurately predicting CO2 emission values for food recipes. Implemented cutting-edge natural language processing techniques and machine learning algorithms to achieve highly accurate predictions.
- **Designed and developed an endpoint for model inference** and integrated with the web API framework.
- **Supervised an analyst on data collection, cleaning, and model training**, overseeing the creation of a high-quality dataset for model development. Provided guidance on data pre-processing and training models.

Tech Stack: Python, AWS, MongoDB, FastAPI

ZS ASSOCIATES

Business Technology Associate Consultant

New Delhi, India

January 2021 - June 2021

Business Technology Associate

May 2018 - December 2020

- **Designed and developed BI solution** leveraged by 3000+ sales representatives from a US fortune 500 pharma firm, for account planning, market segmentation and targeting customers to efficiently launch new drug to the market.
- **Led daily operations of a cloud data warehouse solution**, containing commercial healthcare data, with a 2-member team and supported multiple service requests & incidents every month.
- Oversaw the **development of scalable big data pipelines**, including designing optimized data models for efficient storage of large-scale data sets. The pipelines involved data ingestion, transformations, filtering, and aggregations for BI application
- Set up an **orchestration framework on Airflow**, automating the data pipelines and reducing manual oversight. The framework included advanced features such as error handling, retry mechanisms, and email alerts for pipeline failures.
-

Tech Stack: Python, SQL, PySpark, Airflow, AWS, Bash

SKILLS

Languages & Libraries: Python (TensorFlow, PyTorch, PySpark, NLTK, Scikit-learn, NumPy, Pandas), SQL (MySQL), R, Bash

Software / Tools / OS: AWS (Lambda, EC2, EMR, S3), Airflow, Apache Spark, Linux, Git, Docker

Area of Expertise: Data Science, Machine Learning, Deep Learning, Statistical Modelling, Computer Vision, Natural Language processing, Recommendation Systems, Object Oriented Programming, Data Engineering, Agile development

EDUCATION

UNIVERSITY COLLEGE LONDON (UCL), London, UK

MSC Data Science and Machine Learning

2021 – 2022

Grade: 76.58% (Distinction)

Thesis: Synthesizing visual prosody in digital humans (Research project with Emotech LTD)

Relevant Modules: Natural Language Processing, Statistical Data Science, Information Retrieval and Data Mining, Deep Learning, Machine Vision, Applied Machine Learning

VELLORE INSTITUTE OF TECHNOLOGY (VIT), Tamil Nadu, India

Btech Electrical and Electronics

2014 - 2018

CGPA: 8.67

Thesis: Facial Recognition for Home Security Systems. Implemented a facial recognition system using PCA & Eigenfaces that was integrated with a door access system controlled by microcontrollers. (Published in an IJET journal)

Relevant Modules: Data Structure and Algorithms, Microcontroller and Applications, Digital Signal Processing, C programming, Neural Networks and Fuzzy Logic.

PROJECTS

Neutralizing sexist language

- 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

- 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.

CERTIFICATIONS

Deep Learning Specialization by DeepLearning.AI
IBM Data Science Professional Certificate

LANGUAGES

English (Bilingual Proficiency)
Hindi, Bengali (Native Proficiency)

EMOTECH LTD

London, UK

Research Intern (MSc Project)

June 2022 – September 2022

- 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.

Tech Stack: Python, TensorFlow, Linux, OpenCV, Librosa, NumPy, OpenSmile, Praat