SURAJ SUVENDU CHATTERJEE

(+91) 8652455659 | surajchatterjee5111@gmail.com

https://www.linkedin.com/in/suraj-chatterjee/ | https://github.com/zigma51

OBJECTIVE

Seeking a responsible career opportunity to fully utilize my training and skills, while constantly learning and making a significant contribution to the success of the company. Equipped with astute observational capabilities and the ability to challenge unique problems and hypotheses with an organized mindset through the scientific process.

EDUCATION

University of Mumbai   Mumbai, India
Bachelor of Engineering   **July 2017 - June 2021**

- Major in Information Technology
- CGPA: 9.28 / 10
- Relevant Coursework: Data Structures and Algorithms, Database Management System, Artificial Intelligence, Cloud Computing and Services, Big Data Analytics, Internet Programming, Computer Networks, and Software Engineering with Project Management

EMPLOYMENT EXPERIENCE

COLLINSON   Mumbai, India
Machine Learning Engineer   **May 2024 - Present**

- Designed and implemented **Generative AI** solutions, including **LLM**-based intelligent chatbots using **RAG** (Retrieval-Augmented Generation) architecture with **Langchain,** LangGraph**, AWS Bedrock**, **Llama2**, **Claude**, **GPT** and vector databases (FAISS, **Qdrant**, ChromaDB, Pinecone).
- Developed scalable ML pipelines for traditional ML, NLP and GenAI models using Amazon **SageMaker**, **Bedrock**, **EC2, Lambda**, **ECS, ECR**, **S3, Terraform** and **Snowflake** adhering to **MLOps best practices** for versioning, CI/CD, and model monitoring.
- Integrated serverless backend APIs for AI services using **AWS Lambda** and deployed them using **API Gateway** and **Terraform**.
- Contributed to real-time AI monitoring dashboards and performance reporting using Python-based visualization libraries, **Streamlit** and **Amazon CloudWatch** logs.
- Explored and implemented **NLP, AI** & **LLM** techniques in **NLP** tasks like entity resolution, question answering, text summarization, and semantic search, with a focus on scalability, latency, and response accuracy.
- Embraced **MLOps workflows**, using Github Actions, **Docker** and **Terraform** for containerization and infrastructure provisioning, and monitored deployments through **CloudWatch**.

MONSOON CREDITTECH   Mumbai, India
Data Scientist   **July 2021 - April 2024**

- Developed and deployed end-to-end ML models for **credit risk scorecards** on **time series** data for Indian Banks & NBFCs (Application Scorecard, Behaviour Scorecard, Collection Scorecard, etc) with a strong focus on **predictive modeling, explainability**, and **automated deployment** on **AWS.**
- Achieved a **~90% AUC metric** for Behavior Scorecard predicting the probability of default in the next 3 months.
- Optimized and implemented **multiprocessing** in data processing scripts which increased **processing speed by 90%**
- **Led and managed a cross-functional team of 5 engineers**, overseeing end-to-end delivery of multiple ML and AI projects — from initial planning and architecture design to model development, deployment, and post-launch optimization.
- Automated data & model pipelines and model scoring systems using **AWS Lambda**, **S3**, and **SageMaker**, improving model turnaround and scaling capabilities.
- Processed and **analyzed** large-scale financial and behavioral datasets using **Pandas**, **NumPy**, and **SQL**, enabling robust feature engineering and anomaly detection and showcased the insights via visualizations for client and stakeholder's understanding.
- Gained deep familiarity and implemented statistical ML models using PyData Stack (Pandas, Numpy, Scikit-Learn), **SQL**, **Bayesian Hyperparameter Tuning** (Optuna), Ensemble Boosted Trees **(LGBM, XGBoost), Feature Selection** (RFE, Boruta) and deployment technologies such as RESTful **Django, FastAPI, Flask, Docker** & **AWS** and experimented with **TensorFlow** and **PyTorch** for advanced use cases.
- Led various experiments in **deep learning**, while continuously optimizing performance through **hyperparameter tuning**.
- Authored detailed internal documentation, model explainability reports, and stakeholder presentations showcasing **model insights**, risk stratification, and performance over time.
- Contributed to the adoption of **AI-first approaches** within the organization by identifying use cases for automation and Generative AI experimentation.

UNESCO CHAIR, INSTITUTE OF TECHNOLOGY, TRALEE IRELAND   Mumbai (Remote), India
Data Scientist (Freelance)   **Feb 2021 - June 2021**
- Obtained data through **Web Scraping** using Selenium, Scrapy, and BeautifulSoup from multinational banks' websites
- Maintained and preprocessed the data to remove anomalies, treat outliers, etc for further analysis using Python
- Performed Statistical, Contextual (using **TFIDF, IDF & Glove**), and Sentiment analysis (using **AFINN**, **Vader**) through Natural Language Processing (NLP) to see how Multinational Banks are supporting the Sports for development sector
- Created interesting visualizations on the processed data like Wordclouds, Pie Charts, etc. using **Seaborn** and **Matplotlib** libraries
- Designed a Bitcoin price prediction model using **Deep Learning** techniques such as **LSTM, RNN**, etc, and statistical modeling techniques like **ARIMA**
- Performed sentiment analysis on Bitcoin news articles using unsupervised Sentiment analysis techniques such as Vader, Textblob, AFINN, etc., and incorporated the same into the Bitcoin deep learning model

CAPSTONE PROJECT

ADVANCED RECOMMENDATION SYSTEM USING SENTIMENT ANALYSIS - ML, WEB   Mumbai, India
Undergrad Final Year Project   **Aug 2020 - May 2021**
- Developed an Advanced product recommendation system using Machine Learning clustering techniques like KNN and NLP
- Processed customer reviews using NLP, generated sentiments of reviews using Vader, trained an ML model to recommend products, and finally used Flask to provide a web application for the same and achieved a precision of 85.75% on the model
- Wrote a research paper based on the project which is available here: https://easychair.org/publications/preprint/LHFb

PROJECTS

HANDWRITTEN DIGIT RECOGNITION - DEEP LEARNING   **Jan 2020 - Mar 2020**
- An ML application made using python language for recognizing handwritten digits using CNN by using the MNIST dataset

NEURAL STYLE TRANSFER - DEEP LEARNING   **Apr 2020 - June 2020**
- Performing neural style transfer of an artistic image to a regular image using transfer learning by using the VGG-19 model

SENTIMENT CLASSIFIER - DEEP LEARNING   **Aug 2020 - Oct 2020**
- A project to implement sentiment analysis using the IMDb review dataset with the help of TensorFlow Keras library and Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM)

IMAGE RECOGNITION AND VERIFICATION - DEEP LEARNING   **Feb 2021 - Apr 2021**
- Created a project to implement face recognition and verification using Convolutional Neural Network (CNN)

TECHNICAL SKILLS

- *PROGRAMMING:* R, Python, SQL, C, C++, Java, HTML5, CSS3, Javascript
- *DATABASE SYSTEMS:* MySQL, PostgreSQL, Oracle
- *SKILLS:* Android Development, Web Development, Data Structures & Algorithms, Machine Learning, AI
- *TECHNOLOGIES:* Django, Docker, Flask, Jupyter Notebook, Excel, Powerpoint, AWS - Bedrock, Sagemaker, Lambda, S3, NLP, Time Series Forecasting and analysis, Terraform, Generative AI (GenAI), LLM, Large Language Model, Snowflake

EXTRA CURRICULAR ACTIVITIES

CERTIFICATION COURSES
- Oracle Cloud Infrastructure 2024 Generative AI Certified Professional – Oracle
- Deep Learning Specialization – deeplearning.ai, Coursera
- Tensorflow in Practice Specialization – deeplearning.ai, Coursera
- Data Visualization with python – Cognitive Class
- Sentiment Analysis with Deep Learning using BERT – Coursera
- Accelerating Deep Learning with GPU – Cognitive Class

AWARDS AND PARTICIPATIONS
- Winner of the Coding competition held at St. Francis Institute of Technology (SFIT) for IT Colloquium, 2019
- Technical Executive at Information Technology Students' Association (ITSA) in SFIT: 2018 - 2019
- Participated in Technical & Business Event (Pragati 2019 and 2020) held at SFIT with topics such as Smart Helmet (IOT project for multimedia helmet) & Campus Live (intercampus website for event and job registration) respectively
- Participated in Thadomal Shahani Engineering College (TSEC) hackathons in 2019 and 2020 with topics such as Fire Buzzer (web app for prevention and mitigation of fire hazards) and Shopify (an e-commerce website with an advanced recommendation system using NLP of customer reviews and ML) respectively
- Participated in Smart India Hackathon (SIH) 2020 with Crime chatbot (an AI chatbot for quick crime registration and processing)