Shrimanta Satpati

Kolkata – West Bengal – India

☎ +91 6289982030 • ✉ satpatishrimanta@gmail.com

in www.linkedin.com/in/shrimanta-satpati • ⌗ shrimantasatpati

Education

St. Joseph's University	**Bangalore**
M.Sc. Big Data Analytics (BDA), CGPA – 8.5 (5/59)	*2021–2023*
University of Calcutta (Vivekananda College)	**Kolkata**
B.Sc. (Hon) Economics, CGPA – 8.1 (3/50)	*2017–2020*
National Gems Higher Secondary School	**Kolkata**
ISC 92% (CS), ICSE 95.6%	*2017*

Experience

KPMG Digital Lighthouse — **Bangalore**
Consultant Trainee, Hybrid — *Feb - May 2023*
- LeanX SAP databases web scraping. Tech Stack – BeautifulSoup, Python, Excel
- Hierarchical TS forecasting – Time Series modeling and ML methods like XGBoost, LightGBM, Random Forests, etc
- Univariate TS (FBProphet+ARIMA ensemble) and Multivariate TS (AutoARIMA) on client data - Minimum (**Mean Absolute Percentage Error**) MAPE achieved – 1.98 and 3.97 respectively. Forecasts plotted on **KPI dashboard** in Streamlit.

VisioNxt — **Chennai**
Artificial Intelligence (AI) Intern, On-Site — *Jun - Sep 2023*
- Joint collaborative project - **Ministry of Textiles, Govt. of India**, **National Institute of Fashion Technology (NIFT), India**
- Finetuned **SDXL 1.0** (DreamBooth via LoRA), Stable Diffusion 2 (Textual Inversion) on web scrapped images of Indian ethnic art to design new generation infused Indian garments
- Created **VisioNxt Prompter**, a prompt generation interface in Gradio using **GPT-2, DistilGPT** pretrained models focusing Indian trend parameters.
- Research paper writing (Would be published)

Dabler.co.uk — **United Kingdom**
Intern, Remote — *Jul - Aug 2023*
- Created FastAPI endpoints of **Langchain-based AI** chatbot for personalized multiple pdf/doc QA, Semantic search using AzureOpenAI API, Chroma vector database
- Deployed on Streamlit.io server

The Sparks Foundation — **Singapore**
Data Science and Business Analytics Intern, Remote — *Mar - May 2021*
- Prediction using Supervised and Unsupervised Machine Learning, Decision Trees.
- Exploratory Data Analysis – Retail and Terrorism datasets, COVID-19 Timeline Analysis
- Stock Market Prediction using Numerical and Textual Analysis

Think Tanks and Civil Societies (TTCSP) — **Lauder Institute, University of Pennsylvania**
Fall Research Intern, Remote — *Sep - Dec 2021*
Regional and Global Summit and Data Group:
- Collect qualitative and quantitative data on global economic think tanks to provide data-supported analysis
- Excel analysis to clean data for regional-based ranking of 10000 think tanks

Foundation to Aid Industrial Recovery (FAIR) — **New Delhi**
Macroeconomic Analysis & Policy Studies (MAPS) Intern, Remote — *Jun - Aug 2021*
- Conducted research on contemporary macroeconomic issues facing economies from across the world
- Importance of Post Keynesian macroeconomics, through Modern Monetary Theory (MMT)
- Research Paper: thefairindia.org/maps/wp-content/uploads/2021/07/Shrimanta-Gauri.pdf

Shining Tomorrow Foundation — **New Delhi**
Social Research Intern, Remote — *May - Jun 2021*
- Strategic social research to improve program design, development of a monitoring and evaluation framework
- Developed content on how AI model for connecting India's farmers across corporate-government partnerships

Projects

Stock-ers:
- Used ARIMA forecasting to predict stock prices in real-time for BSE 30 stocks with stock analyzer features
- Built an interactive web app for stock investors (R, R Shiny)

Audible Chess:
- Developed chessboard recognition system using PCA. Real-time piece movement via speech recognition.
- Enhanced usability for visually impaired chess tournaments (Python, TensorFlow for Neural Network).

Twitter Sentiment Analysis in Indian Politics:
- Conducted sentiment analysis on user tweets about Indian politicians (R, Python, Excel, **Twitter API**, Tweepy)
- Implemented Naïve Bayes and Logistic Regression. Understanding voter sentiments/election results

License Plate Recognition with super-resolution GAN:
- License plate recognition with an accuracy of **87%**, utilized (**SRGAN**) (Python, YoloV7, EasyOCR)

Master's Research 2022:
- Researched flood-hit area natural disaster recovery using YOLOv7 object detection and instance segmentation
- Trained on **FAPESP** dataset. Achieved a high mAP@0.50 value of **99.5%** (To be published in Springer-Nature)

Bachelor's Dissertation 2020:
- Investigated relevance of Phillips Curve in Indian context, post-liberalization, Excel statistical/regression analysis

Hackathons

Smart India Hackathon, Government of India   *2022*
Grand finalist out of 10,000 participants in SIH, 2022. Accomplished a supply chain project for Ministry of External Affairs, India.

Bajaj Allianz Hackathon with UpGrad (Kaggle)   *2022*
Customer segmentation for buying life insurance products online using **K-means clustering, DBSCAN** (**LB - 15/32**)

Skills

Programming:: Python, R, C++, SQL, Java, Julia, Scala
Data Analysis/DBMS: Pandas, Numpy, Matplotlib, Scipy, Seaborn, MS-Excel, SQL, MongoDB
Frameworks/API/Cloud: Streamlit, Flask, FastAPI, AzureOpenAI, AWS, Hadoop ecosystem
ML/DL/AI: Scikit-learn, Keras, Tensorflow, Pytorch, BERT, NLP, OpenCV, Langchain, LLM
Tools/Front-End Stack: Git, Github, Jupyter Notebook, VSCode, Spyder/ HTML, CSS, JavaScript, d3.js
Data Visualization: PowerBI, Tableau, Plotly, Dash, RShiny, DOMO

Scholarships

- Completed **The Coding School's Introduction to IBM Quantum Computing** with qiskit. Full **$2500** scholarship
- BuddyforStudy scholarship (TOP 50/5000), partnered by Coursera and **LOreal India**. LOreal executives' mentorship
- Department of Science And Technology, Government of India - JBNSTS Inspire Internship Science Camp, Kolkata

Workshops/Summer schools

- 2-day Data Science workshop by E-Summit cell, **IIT Bombay** in association with Edufabrica (2021)
- KLA and **IIT Madras** workshop on AI and HPC Technology in Semiconductor Manufacturing
- Data Science Summer School (**DS3** 2022), organized by CIVICA and Hertie School, Berlin, Germany
- **Bright Network** Virtual Internship Experience, UK - Business, Operations, Marketing - **Google, Enterprise** etc
- **Amazon ML Summer School 2023** - RNN, Sequential Learning and other ML lectures by Amazon data scientists

Management and Leadership Skills/Sports/Extracurricular Activities

- Secretary of SigmaSquared Association (BDA), Department of Advanced Computing, SJU, Bangalore
- Moderated, Anchored departmental webinars of keynote ISI/CMI professors. Spearheaded DataGram 4.0
- University Chess Champion (1/98) in the Pegasus Intra-murals competition held at SJU
- YouTuber, Founder - **@DataDuo** - Curating AI and latest technology development news
- Volunteering internship at **Youth Empowerment Foundation, New Delhi**, a national NGO