

# SAMIR SENGUPTA

## Data Scientist

Mumbai, India | 8356075699 | samir843301003@gmail.com | [github.com/SamirSengupta](https://github.com/SamirSengupta) | [linkedin.com/in/samirsengupta/](https://linkedin.com/in/samirsengupta/) | [neuralthread.cloud/samir](https://neuralthread.cloud/samir)

### SUMMARY

Data Scientist with a Bachelor's in Data Science from the University of Mumbai and extensive experience in machine learning, Artificial Intelligence (AI), and cybersecurity. Achieved a 40% increase in threat detection accuracy and a 35% improvement in incident response times at Synradar. At Neural Thread, boosted forecast accuracy by 25% and reduced error rates by 30% with advanced machine learning models. Proficient in Python, SQL, Power BI, and various ML frameworks, with certifications in Data Analysis, Generative AI, and Machine Learning. Skilled in deploying innovative AI solutions and optimizing code bases for enhanced performance and efficiency. Adept at translating complex data into actionable insights to drive strategic decision-making.

### EDUCATION

#### BACHELORS IN DATA SCIENCE.

April 2023

University of Mumbai.

Mumbai, India

Acquired Skills: Deep Learning, Machine Learning, Artificial Intelligence, Databases, Neural Networks, Large Language Model.

#### Higher Secondary Education (Science).

April 2020

University of Mumbai.

Mumbai, India

Acquired Skills: Chemistry, Information Technology, Physics, Mathematics, Statistics.

### WORK EXPERIENCE

#### Software Developer, Synradar.

Aug 2023 – July 2024

- Developed machine learning-based intrusion detection systems with Python, increasing threat detection accuracy by 40% and reducing response times by 25%.
- Automated threat analysis using ML algorithms, enhancing breach identification efficiency by 50%.
- Designed interactive security analytics dashboards with Python, leading to a 35% improvement in incident response times.
- Created and optimized ML models for anomaly detection, boosting the identification of unusual patterns and potential threats by 45%.
- Integrated LLaMA 3 for code generation, optimization, and security evaluation, improving code quality and assessment efficiency by 30%.

### PROJECTS

#### Music Mate: Song Downloading System.

May 2023

- Created a music downloader utilizing the Spotify API and Python Tube library to extract songs from Spotify playlists and YouTube videos.
- Constructed a Flask backend to manage API requests, enabling the downloading of songs based on user input.
- Developed a front-end interface using HTML, CSS, and JavaScript for user interaction and input of Spotify playlist or YouTube video URLs.

### **Power BI: Sales Forecasting Dashboard.**

Jan 2024

- Created an interactive Power BI dashboard for sales forecasting, using advanced data analysis and visualization techniques to provide useful insights.
- Implemented powerful forecasting models to predict sales trends accurately, helping businesses optimize their strategies.
- Improved decision-making by presenting detailed sales analytics in an easy-to-use interface, making it simple for stakeholders to understand important insights.

### **Resume Evaluator: Gemini LLM based Candidate Shortlisting.**

Feb 2024

- Developed a Flask application utilizing Google's Gemini Large Language Model (LLM) to effectively summarize documents like CVs and job descriptions.
- Created robust functionality allowing the generation of concise summaries for both job descriptions and CVs, providing valuable insights into candidate suitability for hiring decisions.
- Integrated feedback generation features to recommend enhancements for candidates' professional profiles, thereby facilitating their career advancement.

### **MedScan.ai: Medical Documents Scanner.**

March 2024

- Scans medical images to provide accurate and timely diagnoses, enhancing clinical decision-making.
- Combines visual and textual data inputs, offering comprehensive medical insights and a holistic view of patient health.
- Designed for seamless operation, making it easy for healthcare professionals to integrate into their workflows and improve patient care.

### **Jake.ai: Conversational AI.**

July 2024

- Jake.AI leverages advanced open-source large language models like LLaMA 3, Gemma 2, and Mistral for enhanced conversational and generative capabilities.
- It is locally hosted on your server using LM Studio, ensuring that all data remains private and secure.
- The platform functions as an AI companion, offering robust conversational abilities and innovative generative features.
- As an open-source solution, Jake.AI promotes transparency and flexibility in its AI interactions and implementations.

## **SKILLS & CERTIFICATIONS**

- Technical Skills: Proficient in Python (NumPy, Pandas, Scikit-learn, TensorFlow, Keras), SQL (MySQL, PostgreSQL), Tableau, Power BI, R programming, with expertise in Machine Learning, Deep Learning, Neural Networks, Data Processing, Artificial Intelligence, LLM (Large Language Models), Generative AI, and Langchain, RAG (Retrieval-Augmented Generation).
- Soft Skills: Strong collaborative teamwork, effective communication, adept problem-solving, meticulous attention to detail, excellent time management, adaptable to change, committed to continuous learning, skilled in critical thinking, capable leadership, and analytical thinking.
- Certifications: Google Certified Data Analyst, Generative AI from Google, Power BI, MySQL, Machine Learning, Large Language Models (LLMs) from DeepLearning.AI, Adobe Photoshop, Advanced Excel, Python, Machine Learning from Stanford University.