

SAMIR SENGUPTA

Data Scientist

Mumbai, India | Phone: 8356075699 | samir843301003@gmail.com | [linkedin.com/in/samirsengupta/](https://www.linkedin.com/in/samirsengupta/) | github.com/SamirSengupta | Samir-Sengupta.com

SUMMARY

Experienced Data Scientist proficient in Large Language Models (LLM), Machine Learning (ML), Artificial Intelligence (AI), and Deep Learning techniques. Skilled in Python libraries such as NumPy, Pandas, Scikit-learn, TensorFlow, and Keras, with expertise in advanced analytics and model development. Possesses a strong foundation in SQL databases, Tableau, and Power BI for effective data visualization and analysis. Capable of implementing cutting-edge algorithms and models to efficiently address complex challenges in the data science industry. Dedicated to leveraging AI, LLM, and data visualization skills to drive innovation and continuous improvement in solving real-world problems.

EDUCATION

Bachelor's in Data Science

March 2023

University of Mumbai

Mumbai, India

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

Higher Secondary Education (Science)

July 2020

University of Mumbai

Mumbai, India

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

WORKEXPERIENCE

Associate Software Developer Python and Machine Learning, Synradar

Aug 2023 – Present

- Led the comprehensive migration of the entire code base from PHP to Python, yielding a notable 160% enhancement in overall system speed and performance.
- Engineered a resilient Python application employing Flask to preemptively fortify clients against cyber-attacks, showcasing dedication to bolstering security protocols.
- Utilized Key Performance Indicators (KPIs) for elucidating pertinent information and insights to both clients and the team using Power BI and Tableau.
- Employed artificial intelligence and machine learning methodologies to refine the code base, resulting in substantial time efficiencies and the automation of pivotal procedures.

Data Analyst & Backend Developer, Profit Maxima

Aug 2022 – July 2023

- Utilized Python and SQL for comprehensive data analysis on stock market data, culminating in a notable 25% surge in client satisfaction regarding investment recommendations.
- Developed interactive data visualizations using Tableau, facilitating enhanced comprehension for clients regarding optimal investment choices and contributing to a 20% decrease in response time for client queries.
- Enhanced backend infrastructure leveraging MySQL, thereby realizing a significant 40% reduction in data retrieval time and advancing operational efficiency.

Data Scientist, Neural Thread (Internship)

Jan 2021 – July 2022

- Engineered ML models for precise customer churn prediction, resulting in a significant increase in customer retention rates from 71% to 93% annually.
- Applied cutting-edge deep learning algorithms to enhance the precision of Predictive Analytics Solutions, achieving a better prediction accuracy of 42%.
- Led K-means clustering initiatives for dynamic customer segmentation, optimizing targeted marketing strategies.
- Developed and implemented tailored retention models, resulting in a substantial increase in customer loyalty from 47% to 83%.

PROJECTS

Movie Recommendation System

March 2023

- Created a movie recommendation system using the TMDB API with content-based filtering.
- Deployed and hosted the system on a Streamlit server.
- Performed sentiment analysis using NLP techniques and deployed the application with Streamlit for user interaction and visualization.

Music Mate: Song Downloading System

May 2023

- Developed a music downloader using the Spotify API and Python Tube (PyTube) library to download songs from Spotify playlists or YouTube videos.
- Built a Flask backend to handle API requests for downloading songs based on user input.
- Designed front-end interface with HTML, CSS, and JavaScript for user interaction and input of Spotify playlist or YouTube video URLs.

Customer Segmentation

Nov 2023

- Executed K-means clustering to segment customers based on their Annual Income and Spending Score, facilitating a deeper understanding of consumer behavior.
- Employed visualizations to clearly delineate clustering results, identifying 5 discernible customer segments with distinct characteristics.
- Delivered actionable insights tailored to inform targeted marketing campaigns and strategic business decisions, leveraging the nuanced understanding derived from customer segmentation analysis.

Power BI Sales Forecasting Dashboard

Jan 2024

- Engineered a dynamic Power BI dashboard dedicated to sales forecasting, seamlessly integrating sophisticated data analysis and visualization techniques to deliver actionable insights.
- Deployed advanced forecasting models adept at predicting sales trends, thereby enabling strategic optimization of business approaches.
- Enhanced decision-making efficiency by presenting comprehensive sales analytics through an interactive and user-friendly interface, streamlining the interpretation of critical insights for stakeholders.

Gemini Pro

Feb 2024

- Engineered a Flask application leveraging Google's Gemini Large Language Model (LLM) to proficiently summarize documents such as CVs and job descriptions.
- Implemented robust functionality enabling the generation of concise summaries for both job descriptions and CVs, offering valuable insights into candidate suitability for hiring decisions.
- Integrated feedback generation features aimed at recommending enhancements for candidates' professional profiles, thereby facilitating their career advancement.

Gemini Vision Pro

March 2024

- Engineered an image recognition application employing Google's Gemini Pro Vision model to analyze user-provided images.
- Implemented sophisticated functionality to interpret images and generate descriptive prompts, aiding users in creating similar imagery.
- Integrated the Gemini Pro Vision model to discern image content accurately, ensuring precise responses.
- Designed an interactive user interface facilitating image input and result display, thereby elevating user experience, and fostering engagement.

ADDITIONAL SKILLS & CERTIFICATIONS

- Technical Skills: Python (NumPy, Pandas, Scikit-learn, TensorFlow, Keras), SQL (MySQL, PostgreSQL), Tableau, Power BI, R programming, Machine Learning, Deep Learning, Neural Networks, Data Processing, Artificial Intelligence, LLM (Large Language Models), Generative AI.
- Soft Skills: Collaborative teamwork, Communication, Problem-solving, Attention to detail, Time management, Adaptability, Continuous learning, Critical thinking, Leadership, Analytical thinking.
- Certifications: Google Certified Data Analyst, Blockchain Technology, Power BI, MySQL, Machine Learning, Recommender System, Adobe Photoshop, Advanced Excel, Python.