



Paramartha Sengupta

Data Scientist

Focused and detail-oriented Data Scientist with ~4 years of experience in working with vast datasets and delivering business-enhancing insights for Fortune 500 corporations. Currently seeking an opportunity in the Data Science/Machine Learning domain to leverage my skills and expertise effectively. I am a highly motivated individual, committed on ongoing learning and collaborative efforts to solve complex business challenges

Contact Information:

Phone: +91-7980458439

[Mail](#)

[LinkedIn](#)

[Kaggle](#)

[HackerRank](#)

[GitHub](#)

WORK EXPERIENCE

Data Scientist

ZS Associates

10/2021 - Present

Bengaluru, India

Achievements/Tasks

- **Potential User Identification:** Developed a semi-supervised ML algorithm (PU learning) to identify patients requiring critical healthcare treatment based on their medical history. This model proved effective in **identifying 60% patients** in urgent need of treatment, **compared to a baseline of 25%**
- **Web Usage Optimization:** Achieved **~17.2% increase in overall website utility** by developing a causal relationship model that establishes the correlation between user interaction on a website to product purchase. Recommendations for website enhancement were made based on this model
- **Data Preparation Tool:** Designed and developed an end-to-end application hosted on Streamlit that provides augmented insights on patients' journeys and milestones using raw claims data as inputs. **Implemented a GPT-3 module API** that enables users to control the application using plain English text, without requiring any coding experience
- **Prescriber Prediction Model:** Currently engaged in the development of a model that **predicts the most suitable adopters of a specific drug** in the market, utilizing cross-matrices of their potential and propensity

Data Science Analyst

Novartis

08/2019 - 10/2021

Hyderabad, India

Achievements/Tasks

- **Data Modelling Pipeline Development:** Contributed to the development of a comprehensive data processing and **modelling pipeline that improved global clinical trials monitoring**, and designed and implemented a dashboard for continuous performance tracking
- **Field Distance Optimizer:** Developed a field distance optimization algorithm for United States and China, which **resulted in a 27.1% and 18.7% reduction in overall travel distance for monitoring purposes** in the respective countries
- **Risk Analysis:** Developed a risk estimation indicator that facilitates the identification of the most promising sites for conducting clinical trials, **resulting in ~24% optimization of global monitoring activities**

EDUCATION

Bachelor of Technology

National Institute of Technology Durgapur

08/2015 - 06/2019

8.98/10.00

SKILLS

Data Science: Machine Learning

(Supervised/Unsupervised/Semi-Supervised), Statistical analysis, Deep Learning, NLP

Databases: MySQL, PostgreSQL, Oracle

Python: NumPy, Pandas, Scikit-Learn, Statsmodels, Scipy

ACHIEVEMENTS

"Rising Star of the Profession" Award (09/2022)

Recognized for excellent performance in delivering dynamic client projects, and parallelly working on the development of the Streamlit based data processing tool for an internal POC

Novartis Star Impact Award (01/2020 - 09/2021)

Have been awarded with the award 5 times for excellent performance in stakeholder management and designing and developing the backend data modelling pipeline for the Global Trails monitoring Dashboard

Kaggle Master (12/2020 - Present)

Have achieved the title of Kaggle Notebooks Master and ranking globally in the top 0.2% of the users

PERSONAL PROJECTS

KaggleX BIPOC Mentor (12/2022 - 03/2023)

- **Provided mentorship and guidance to mentees on the implementation of ML and Data Science concepts at an Industry level**, resulting in the completion of their first end-to-end ML project within three months [Link](#)
- Developed a **feature evaluation module that enables the assessment and evaluation of features' influence on model's learning**
- **Designed and developed a customized food recommendation engine** that leverages user preference and location choice to predict the best restaurants and dishes

Voice Assistant Automation (12/2022 - Present)

- Currently **leading the development of a voice-controlled personal-assistant project** that is designed to transform user commands into actionable responses by performing a range of tasks including searching Wikipedia, playing YouTube videos etc.
- As this project is currently in development, I am focused on implementing performance fixes, including search functionality and usability

CERTIFICATES

Python3 Programming Specialization (08/2020 - Present)

Specialization course comprising of basic to Intermediate python programming and two major milestone projects in the domain of NLP [Link](#)

Applied Machine Learning in Python (10/2020 - Present)

Machine Learning course from the University of Michigan, focusing primarily on industry level ML use cases [Link](#)

ORGANIZATIONS

Kaggle (12/2022 - 03/2023)

Data Science Mentor