Statement of Purpose

Maitreya Sameer Ganu
maitreyaganu@gmail.com
Personal Website

1 Introduction:

I am deeply interested in pursuing research at the intersection of data science, AI, and applied mathematics. My academic journey has been shaped by my diverse experiences in machine learning, data analysis, and computational modeling, which have sparked my passion for tackling real-world problems through innovative methods. In particular, I aim to explore uncharted territories in clinical data science, mental health diagnosis, fraud detection, and prediction. Through rigorous research and an interdisciplinary approach, I hope to contribute to fields that have significant societal impact.

2 Education:

I am currently a Data Science Major at IISER Thiruvananthapuram. My coursework includes Data Science. Through various research projects and coursework, I have gained proficiency in machine learning, statistical modeling, and algorithm design.

3 Projects :

3.1 Heart Disease Prediction

This project aimed to predict the likelihood of heart disease in individuals using medical data. The dataset included features such as age, blood pressure, cholesterol levels, and exercise habits. The model was developed using Logistic Regression, which achieved an accuracy of 81.463%. I focused on feature selection and data preprocessing to improve model performance, applying techniques like normalization and handling missing values. The model's predictive capability can help healthcare professionals identify individuals at high risk for cardiovascular diseases, allowing for early diagnosis and intervention, thereby saving lives.

3.2 Diabetes Prediction

In this project, I developed a model to predict the likelihood of diabetes based on various health parameters like age, BMI, family history, and blood sugar levels. The model was built using machine learning techniques, without utilizing TensorFlow, and achieved an accuracy of 80.5%. I focused on preprocessing the dataset, including feature scaling and handling imbalanced data using oversampling techniques. The goal was to provide a tool for early detection, enabling healthcare providers to intervene before the onset of the disease. This work demonstrated the potential of predictive modeling in improving patient care and health outcomes.

3.3 Stock Prices Prediction

This project involved applying neural networks and deep learning techniques to predict future stock prices using historical data. By analyzing stock market trends, the model aimed to forecast price movements of specific stocks based on previous performance. I implemented a neural network model to capture the complex non-linear relationships in the stock price data. The project provided valuable insights into the use of deep learning models in finance, helping me understand the practical challenges and benefits of applying AI in predicting financial markets. It also improved my understanding of time-series forecasting and data preprocessing for financial datasets.

3.4 Deep Learning Approach on the MNIST Dataset

In this project, I applied deep neural networks to classify handwritten digits from the MNIST dataset, achieving an accuracy of around 99%. The project allowed me to explore various neural network architectures and experiment with techniques such as dropout and batch normalization to enhance the model's performance. I also fine-tuned hyperparameters and employed data augmentation to avoid overfitting.

3.5 Temperature Prediction

This project utilized Long Short-Term Memory (LSTM) networks to predict the maximum and minimum temperatures for future days. The dataset consisted of historical temperature data, and the model was trained to forecast future temperature trends based on past observations. By using LSTMs, a type of recurrent neural network (RNN), I was able to capture the temporal dependencies in the data. The project demonstrated the effectiveness of LSTM networks in time-series forecasting tasks, and it highlighted their ability to model sequential data, which can be applied to various domains such as weather forecasting, finance, and traffic prediction.

3.6 IPL Score Prediction

In this project, I built a neural network model to predict the total score of any match in the Indian Premier League (IPL) based on various features such as team statistics, player performance metrics, and venue-specific data. The model was developed using historical match data and team/player statistics. An interactive user interface was created in Jupyter Notebook, allowing users to input parameters and predict scores for different IPL matches. This project not only helped me apply neural networks to sports analytics but also gave me experience in building interactive data-driven applications, demonstrating how machine learning can be used for real-time predictions in sports analytics.

4 Internships:

4.1 Cryptography Internship, IISER Pune

During my reading project at IISER Pune, I worked on cryptographic protocols such as Data Encryption Standard and Group theory. I focused on the mathematical foundations behind these protocols and explored their applications in securing digital communications. Additionally, I worked on optimizing encryption algorithms and assessing their resilience against various types of attacks, gaining hands-on experience in both theoretical and applied cryptography.

4.2 Community Detection and Network Analysis using Modularity, IISER Thiruvananthapuram

At IISER Thiruvananthapuram, I worked on a project focused on modularity analysis of complex graphs, where I aimed to identify community structures within different real-world networks such as road networks, protein interaction networks, and neural networks. The project involved constructing graphs by representing entities as nodes and their interactions as edges. I then applied community detection algorithms like Louvain and Girvan-Newman to optimize modularity and uncover clusters or subgroups within the networks. By analyzing the modularity values, I gained insights into the organization of these systems, helping to understand their functional and structural properties, which has applications in optimization, biological research, and neural network design.

5 Concluding Remarks :

With my diverse research experience, technical skills, and passion for solving complex problems through data science, I am eager to continue my academic journey in AI and machine learning. I am particularly interested in exploring uncharted areas in clinical data science, mental health diagnosis, and fraud detection. I am confident that the skills and experiences I have gained will enable me to contribute meaningfully to cutting-edge research in these fields and make a tangible impact.