DEGA BALAJI VARA PRASAD

Email dbalajivaraprasad@gmail.com
Mobile +91 7386391303
+91 9441051303

Website https://balaji24092001.github.io/portfolio/
LinkedIn https://www.linkedin.com/in/balaji24092001/
GitHub https://github.com/BALAJI24092001
Kaggle https://www.kaggle.com/balajivaraprasad

EDUCATION

Integrated Masters in Statistics	Pondicherry University	8.38 CGPA	2019-24
Higher Secondary Education (XII)	Sri Chaitanya Jr. College	9.61 CGPA	2017-19
Secondary School Education (X)	Ravindra Bharathi	9.70 CGPA	2016-17

SKILLS

- **Programming Languages:** Python (Plotly, PyTorch, SkLearn, SciPy), C++ (STL), R (dplyr, R Shiny)
- **Data Analysis & Visualization:** SQL, PowerBI, SPSS
- **Version Control & Tools:** Git, GitHub, Docker
- **Area of Interests:** Deep Learning, Causal Inference, Natural Language Processing
- **Additional Skills:** Problem Solving, Self-learning, Presentation, Adaptability

EXPERIENCE

Graduate Intern **GSK -** Bangalore (Jan 2024 to Jul 2024)

- I gained a comprehensive understanding of Survival Analysis, particularly in the context of pharmaceutical research. I learned how to model and interpret time-to-event data, which is crucial for analysing patient outcomes.
- I delved into the complexities of Causal Inference, focusing on how to establish causal relationships between treatment and outcomes in the presence of competing risks. This has been essential in improving the reliability of conclusions drawn from clinical data.
- I researched and applied methods to address the challenges posed by competing risks in Survival Analysis. This involved learning advanced statistical techniques to ensure accurate interpretation of survival outcomes, even when multiple potential outcomes are possible.

Statistical Assistant **Pondicherry University** - Dr R Vishnu Vardhan (Nov 2022 to Dec 2023)

- I worked on several projects that deepened my knowledge of biostatistics, applying statistical methods to real-world biological data. This experience strengthened my ability to analyse and interpret complex datasets effectively.
- I contributed to projects involving comparative analysis, honing my skills in comparing statistical methods and outcomes across different studies. This work enhanced my ability to draw meaningful conclusions from diverse data sources.

PROJECTS

Thoracic Disease Detection using Deep Learning

- This is a Python3 (Pytorch) reimplementation of CheXNet. The model takes a chest X-ray image as input and outputs the probability of each thoracic disease along with a likelihood map of pathologies.
- We have also proposed a slightly-improved model which achieves a mean AUROC of 0.847 (v.s. 0.841 of the original CheXNet).

Mushroom Classification

- I developed a classification model using the XGBoost algorithm, which is known for its efficiency and high performance. To optimize the model, I implemented a function to search for the best hyperparameters, leading to an exceptional classification accuracy of 98.530%. This high accuracy underscores the model's robustness in distinguishing between classes.
- The performance of the classification model was evaluated using the Matthews Correlation Coefficient (MCC), a metric that takes into account true and false positives and negatives, providing a balanced measure of accuracy.

ACHIEVEMENTS

- Selected for final round in the competition "Data Quest" conducted by Novo Nordisk
- Stood 1st in FACTISTICS in the Inference 2022, an online statistical fest conducted by the Christ University, Bangalore.
- Selected for "Mathematics Training and Talent Search (MTTS) program"

CERTIFICATION

- Tableau – Amity University
- Python – HackerRank
- Machine Learning Specialization – Coursera
- Causal Inference – Coursera

EXTRA CURRICULAR ACTIVITIES

- Hobbies include listening to music, playing Badminton, Reading Poetry, Novels and Non-Fiction.
- Experience with Manjaro Linux for a year, with basic knowledge and most interest in the prompt-based system handling.
- Developed interest and self-taught how to use a prompt-based text editor called NeoVim.