DEBODEEP BANERJEE

Viale della primavera, 15, Roma- 00172, Roma, RM, Italy
Contact:+39-3510511586 **E-mail: debodeepbanerjee33@gmail.com**
Github: Debodeep94 **Website:** debodeep94.github.io

Objective
I want to explore new dimensions of the world of statistical inference and data science. Through the ages, especially in the regime of Statistical data analysis, the inferential methods have played a revolutionary role. On that note, my interest lies in the field of statistical inference, machine learning and artificial intelligence. My aim is to work on this field in a more rigorous way so that my contribution enriches the subject as well as its application in real-life cases. Apart from that, playing with various visualization techniques is another point of interest of mine.

Areas of interest
•Bayesian inference
•Data visualization
• Human computer interaction

Education
•Doctoral Education
University of Trento and University of Pisa (2022-)
•Laurea Magistrale in Data Science (Masters in Data Science)
Sapienza University of Rome (2022)
•Master of Science in Statistics
Department of Statistics, University of Madras (2018)
•Bachelor of Science, Statistics (Honors)
Asutosh College, University of Calcutta (2016)

Work experience
Organisation: Project Consulting SRL
Designation: Intern
Details: Artificial intelligence in video analysis
Duration: June, 2022- October, 2022

Organization: Baker Hughes
Designation: Artificial Intelligence intern (curricular internship)
Title: Deep learning on data acquired by turbomachinery
Duration: July, 2021-Jan, 2022

Organization: SPI Global Private Limited
Designation: Subject Matter Expert, Statistics
Duration: May, 2018-August, 2018

Internship
Institute: Indian Statistical Insitute
Topic: Data Analysis using Bootstrapping and Logistic Regression.(May, 2017- June, 2017)

Paper presentation
•Presented Paper in International Conference on Theory and Application of Statistics and Information Science, 2018
Topic- Comparative Analysis of Two 1×j contingency table under Bayesian perspective
•Presented paper in National Conference on Dimensions of Sustainable Economic and Business Development organised by Marian College, Kuttikkanam, 2018
Topic- CSR and Sustainable Development: An Analysis of Practices in India

Academic project(Click on the title for Github link)

1 Websraping, LDA and forecasting

• Scraped all the speeches of several banks' board member(USA and China) from the website and stored in textual format.
• Concept used are NLP, Latent Dirichlet Analysis and forecasting using LSTM and LSTM autoencoder.
• Important packages are BeautifulSoup, plotly, tensorflow.

2 Stock market prediction using twitter data

• ONGC (Oil and Natural Gas Corporation) data (2015-2021). LSTM model is used to predict the future data.
• Another LSTM model used after combining the scores obtained from the sentiment analysis of the tweets and the stock data.

3 Model interpretability using LIME

• Image dataset having archaeological building with 10 classes.
• Classification task using CNN.
• LIME library used to address the concept of model interpretability.

4 Fire Detection using convolutional neural network

• Image data collected from Kaggle and CVPR lab.
• Classification task performed using CNN (ResNet50 and Bayesian CNN).
• Packages used are tensorflow, PyTorch, kerastuner.

5 Performance of Knowledge Organisation and Human Action using Probability Graphical Model and Efficiency Analysis

• Literature review of the paper obtained from Scopus and WOS.
• Application of PRISMA model and Ontological Modelling using Eddy and Protege.
• Visualization using Biblioshiny, VOSviewer.
• Further applied Probability Graphical Model and Efficiency Analysis.

6 Linear regression using JAGS

• R programming project using 2 statistical models - multiple linear regression model and bayesian model using JAGS.

7 Covid situation visualization

• Self interested project. Visualization (static and animated).

M.Sc.(in Statistics) Research project

Title: A study on classical inference on single binomial proportion
Duration: October, 2017- May, 2018
Supervisor: DR. Subbiah M. (Former Associate Professor, Presidency College, Chennai, India)
Short description:
We revisited the problem of interval estimation of a binomial distribution.The focus has been fixed on the justification of choice of several confidence intervals based on their respective coverage probabilities. The performances of the confidence intervals have been analyzed with help of coverage probabilities and expected run length. Further, proposition of new interval has been placed and it has been compared with the help of coverage probability and expected run length.

Research project: Master in Data Science

Title: Nonparametric modal regression for directional responses
Duration: January, 2022-
Short description:
Regression technique has been a corner stone of studies in machine learning. In this thesis we essentially focus on a specific genre of regression analysis where the predictors are coming from a spherical plane and the response variable is linear. To allow more flexibility to our model, we chose to work on nonparametric regression. However, properties of spherical data have been studied recently and there are extensive studies on nonparametric regression as well. One very known problem of regression models lie in its robustness as the conditional mean can be affected with the presence of outliers. Median regression or quantile regressions can be better alternatives, but they do not stand for the best choice. Instead, a modal regression seems to be a much robust option in this case. Thus, in this study, we try to develop a nonparametric regression tool which performs based on directional predictors and linear responses. Further we propose an EM based modal regression algorithm for linear response and directional predictors.

Technical report

Title: Comparative Analysis of Two $1 \times j$ contingency table under Bayesian perspective
Authors: Deboddep Banerjee, Trina Sahoo, DR. Subbiah M. and DR. M.R. Srinivasan
Duration: March, 2017- February, 2018
Publication: University department library, Department of Statistics, University of Madras, February, 2018
Short description:
This study has exploited these aspects into categorical data of two one-dimensional variables each with same J levels of categories. The notion focussing on the ratio of independent Beta distribution is considered using closed form approach with Gauss hypergeometric function and Monte-Carlo techniques. The entire approach is illustrated with a primary data set that aims to study the impact of gender on perceived important social issues; essential computations are carried out using R program. The modelling advantages of Bayesian approach has been studied and the results are directly interpreted regarding the context of the problem.

Certificate program

Program title: Certificate program on Machine learning and Artificial intelligence
Institute: Madras School of Economics
Duration: June, 2018- August, 2018

Program title: Certificate course on Python
Institute:FITA academy
Duration: December, 2018- January, 2019(tentatively)

Software proficiency

•**Languages:** Statistical applications of C language, Python, R
•**Statistical analysis software:** Minitab, SPSS
•**MS Office suit:** MS Office (MS Word, MS Excel, MS PowerPoint, MS Access)
•**Other(s):** LaTeX, SQL, Power BI(beginner)

Awards and achievements

•Secured first place in the Inter-Departmental Quiz Competition, organized by the Department of Library and Information Science, University of Madras.
•Secured first place in quiz in the Inter-Collegiate competition, organized by the Department of Statistics, Madras Christian College.

Key skills

•Language proficiency in Bengali, Hindi, English, Italian(Basic).
•Eager to learn new concepts.

Hobbies

•Reading books (Mainly non-fiction)
•Writing stories, articles, poems