Cambridge, Boston, USA

sbit@bu.edu | subhrangshubit@gmail.com
Webpage : https://subhrangshubit.github.io
Github : https://github.com/SubhrangshuBit
+1-857-639-0897

Subhrangshu Bit

COMPUTER SCIENCE PHD STUDENT

EDUCATION

Boston University, Boston
PhD, Computer Science   *Sep' 23 -*
Ramakrishna Mission Vivekananda Educational & Research Institute, West Bengal
Master of Big Data Analytics, Computer Science   *Jul' 19 - Jun' 21*
GPA: 9.93/10
St. Xavier's College, Kolkata, West Bengal
Bachelor of Statistics, Statistics   *Jul' 16 - Jun' 19*
GPA: 7.6/10

WORK EXPERIENCE

Data Scientist
Dr. Reddy's Laboratories   *Apr' 22 - Present*

- Developed a new application to predict the humaneness and immunogenic potential of a protein using *hierarchical cluster analysis* and *artificial neural networks*. Implemented on the basis of peer-reviewed articles and bio-scientist's domain expertise.
- Designed an end-to-end *generative model* for pruning chemical space, drug designing and screening. Leveraged *Monte Carlo Tree Search* and graph based algorithms on top of *Bayesian Optimization*.
- Built an application on yield optimization of products through *convex optimization* techniques.
- Built a dashboard on automatic integration of chromatographic signals leveraging *nearest neighbors*.

Assistant System Engineer Trainee (AI & ML)
Tata Consultancy Services   *Jul' 21 - Apr' 22*

- Built a *Recommendation Engine* for allocation of associates to projects across TCS.
- The engine was built on the basis of a similarity metric (cosine distance) and fuzzy matching techniques between associate and project requirements/competencies.

Data Analyst Intern
Dr. Reddy's Laboratories   *Feb '21 - Jul' 21*

- Developed a framework to provide statistical estimates of quantitative distribution of ingredients of a drug from its *hyperspectral image*.
- Incorporated Beer Lambert's law to perform multiple linear regression of acquired Raman spectra on pure spectral signatures of the components. The coefficients of regression were interpreted as the proportional quantitative estimates.
- The components with concentrations $\leq 1\%$ were difficult to predict owing to erroneous data acquisition. On the contrary for those in significantly larger quantities were predicted accurately with $\leq 4\%$ error.
- Github Link

RESEARCH INTERESTS

3-Dimensional Computer Vision, Statistical Methods, Generative Neural Networks, AI in medicine

RESEARCH WORK

Alzheimer's prediction and progression using mixture of class Restricted Boltzmann Machines: Working Paper
Supervisor : Prof. Swami Vidyapradananda and Dr. Tapan K. Khan   *Sep '21 - Present*

- We used 3D structural MRI scans to determine both the current stage of dementia (CN/MCI/AD) and the probabilistic progression to advanced stages.
- Addressed the challenge of high dimensionality of sMRI by extracting the reduced dimensional latent feature vector using a *Variational Autoencoder*.
- Resolved the class imbalance due to low sample size by multiple sampling from the latent distribution.
- The extracted feature vectors are then used as input conditions to a *mixture of class Restricted Boltzmann Machines* that captures distinctiveness of parameters of the three separate classes of dementia
- The model outputs a conditional transition probability matrix (cTPM).With a 70-30 split of the total cases, we achieve an overall prediction accuracy of higher than 80% with class specific accuracies 86%(CN), 82%(MCI) and 93%(AD). The AUC values of CN, MCI and AD are respectively 0.94, 0.89 and 0.99. The overall progression accuracy is higher than 75% with precision over 0.70 for each class.

ACADEMIC PROJECTS

Implementation of improved second order optimization algorithms
Supervisor : Prof. Swami Vidyapradananda — *Jul '20*

- Explored a Quasi-Newton optimization approach to solve a quadratic function using *Davidon-Fletcher-Powell* Method and *Fletcher-Reeves* Conjugate Gradient method
- With the same initialization we analytically and theoretically show that both the methods generate identical gradient directions.
- [Github Link](#)

Comparative Study of Bayesian Estimators & Maximum Likelihood Estimators
Supervisors : Prof. Surabhi Dasgupta — *Jan '19 - Mar' 19*

- We studied the behaviour of Maximum Likelihood Estimators and Bayesian Estimators of three standard theoretical distributions - *Binomial*, *Poisson* and *Normal* with increasing sample size.
- The prior information for bayesian estimators considered under this study were: *Jeffreys' Invariant prior* and *Natural Conjugate (NC) prior*
- Found that bayesian estimators with NC prior although being same as MLE are an improvement since, unlike MLE, it encapsulates the past information whereas those with Jeffreys' prior were consistent and tend to be the same as MLE for large sample sizes
- [Github Link](#)

Zero Inflated Time Series Analysis of Terrorism in India
Course : Time Series Analysis | Supervisor : Prof. Sudipta Das — *Jul '21 - Dec '21*

- Bypassed *ARMA* models, which are generally restricted to continuous state-space by utilizing an observation driven model.
- Handled overdispersion using a Gamma distribution resulting in a *zero-inflated Negative Binomial regression* model. - Incorporated ARMA type structure to model the mean of the distribution.
- [Github Link](#)

Automation of Pacman game for single and multi-agent
Course : Artificial Intelligence | Supervisor : Prof. Br. Tamal — *Jul '21 - Dec '21*

- Implemented search algorithms: Breadth First Search, Depth First Search, A^* search for single agent optimal path finding.
- Implemented Q-Learning and Value Iteration to make the Pacman learn the optimal solution.
- [Project Overview](#)

COMPUTER SKILLS

Languages: Python3.x, R, Java(Prelim), SQL(Prelim), LaTeX
ML frameworks: Tensorflow2.x, scikit-learn, PyTorch, CasADi, Keras, OpenCV

AWARDS & ACHIEVEMENTS

CIO Special Award 2022, Dr. Reddy's Laboratories
First Rank Holder (Gold Medallist) 2019-21, RKMVERI
Awarded the **INSPIRE Scholarship** by DST, Govt. of India
Secured a **Rank of 2** in School in Higher Secondary Education

EXTRA CURRICULAR

Hobbies: Football, Wargames, Reading Novels