Tuhin Subhra De

tuhinsubhrade@yahoo.in GitHub LinkedIn Google Scholar Website

Education

BTech & MTech, Indian Institute of Technology Kharagpur Jul 19 - Apr 24
Major: *Civil Engg* | ***Micro Specialization:*** *AI and Applications* | ***CGPA****: 8.75/10.0* *West Bengal, India*

Current Employment

Software Management Trainee July 2024 - Present
ICICI Bank India Pvt. Ltd (Technology Group) *Mumbai, India*

Research Interests

Deep Learning, Probabilistic Machine Learning and Generative Modelling, Neural Network Interpretation

Publications

1. **Tuhin Subhra De**, Madeti Karthikeya, and Sujoy Bhattacharya. A non-linear lasso and explainable lstm approach for estimating tail risk interconnectedness. Applied Economics, Pg :1–15, 2024. [doi]
    - Estimated the tail risk interconnectedness within the large banks of Japan and US using explainable ML
    - Used LassoNet (Lemhadri et. al, 2021), a L1 regularised neural network to obtain a parameter cum gradient based feature importance on the daily stock returns of banks within 5% confidence interval of quantile loss
    - Compared Interpretable-Multivariate LSTM (Guo et. al, 2019) which uses attention based variable importance, with LassoNet and other existing researches till date obtaining lowest objective loss by the LSTM model
2. **Tuhin Subhra De**, Pranjal Singh, and Alok Patel. A machine learning and empirical bayesian approach for predictive buying in b2b e-commerce. International Conference on Machine Learning and Soft Computing '24, page 17–24, New York, NY, USA, 2024. Association for Computing Machinery (ACM) [doi]
3. Anshul Balamwar, **Tuhin Subhra De**, Debabrata Das, Manoj Kumar Tiwari, Prediction of Turn Around Time using Neural Networks, IFAC Conference, 2022 [doi]
    - Predicted Turn Around Time of port ships by comparative analysis of various machine learning approaches within a certainty of 26.71 hours which initially had weeks of inconsistency upon 5 years of ships' log data

Research Experiences

Diffusion based Variational Autoencoder for Seq2Seq generation April 2024 – Present
Center for Excellence in Artificial Intelligence, IIT Kharagpur *Advisor: Prof. Adway Mitra*

- Implemented variational autoencoder to generate text with timestep wise latent spaces which regulated the posterior collapse problem (decoder ignores the latent space and copies the input) to a large extent
- Currently, to enhance latent space dependency and further mitigate the posterior collapse problem, we are testing a diffusion based denoiser whose parameters are dependent on timestep wise latent spaces and embeddings in-place of traditional RNN-LSTM based decoder

Hybrid ensemble method for potential customer recommendation May 2023 – July 2023
Udaan.com (Hiveloop Technologies Pvt. Ltd.) | led to publication no. 2 *Bangalore, India*

- Increased customer order conversion rate from 4% to 13% using predictive buying methods on Azure Databricks
- Collaborated with the business intelligence team to engineer features for customer orders using SQL from raw data
- Implemented an optimal grid-searched XGBoost classifier giving customer order probability and exhibiting 8% rate
- Stacked XGBoost with Poisson-Gamma model making Logistic Regression as meta learner reaching 13% rate

Knowledge distillation for unlabeled data scenario April 2022 – July 2022
Deakin University, Melbourne, Australia *Advisor: Prof. Jinho Choi*

- Merged response and feature based distillation under unlabelled data scenario between two different architectures
- Observed an 3% and 4% increase in accuracy for the student model (2 layered Perceptron) upon distillation from a teacher model (Multilayered CNN) when trained on Fashion MNIST and CIFAR-10 dataset respectively

- Simulated a data label free knowledge transfer scenario, where the student model is trained jointly by directing its prediction distribution close to the teacher's using KL Divergence and maximizing the cosine similarity between the high dimensional representation of the intermediate feature spaces

Multimodal Image generation from Attention based Conditional GAN April 2022 – June 2022
Center for Excellence in Artificial Intelligence, IIT Kharagpur *Advisor: Prof. Adway Mitra*

- Generated images with varying facial features using CGAN and simple convolutional upscaler for modal class mean
- Devised a TPU trained conditional GAN on Celecb-A dataset, equipped with channel and pixel based cross attention with the label vector at every layer of generator. Compared the results for CGAN and modal class means
- Developed an end-to-end pipeline to train and deploy models using Streamlit package in form of web application

Neural Approximation and Knapsack DP for booking optimization Sep' 2021 – Dec' 2021
Indian Institute of Management Mumbai *Advisor: Prof. Manoj Tiwari*

- Devised a method to maximize profit by selective cargo acceptance achieving 48% more profit than random process
- For an accepted booking data in bulk, we used Vehicle Routing Problem (VRP) to get the minimum delivery cost
- VRP being computationally expensive, jointly trained Neural Networks and its linear estimate as successive multiplied weight matrices on the above generated data to obtain an empirical cost on each cargo requests
- Used the empirical cost, revenue and capacity constraints for selective acceptance using dynamic programming

Achievements

- Qualified Regional Mathematics Olympiad to appear for Indian National Mathematics Olympiad (INMO) conducted by Tata Institute of Fundamental Research in 2017. I was one among 42 candidates in state and 500 candidates across nation qualified for INMO.
- Ranked in top 0.1% in Physics All India Senior Secondary Certification Exam (Class 12th) conducted by Central Board of Secondary Education among 2.5 million students wide nation in 2018.
- Secured 5542 All India Rank (General) among 22,000 candidates in country in Joint Entrance Examination Advanced in the year 2019.

Relevant Coursework

ML/AI: Programming and Data Structures, ML Foundations, Computer Vision and Sequential Modelling, Artificial Intelligence and Game Theory, Graphical and Generative Machine Learning, Big Data Processing

Maths and Statistics : Engineering Calculus, Probability, Linear Algebra, Regression Modelling

Technical Skills

Languages: C, C++ (STL), Python, SQL, Bash, JavaScript, HTML/CSS

Developer Tools: Git, Microsoft Azure Databricks, VS Code, Google Colab, Oracle Weblogic

ML Libraries: PyTorch, TensorFlow-Keras, Scikit-Learn, PySpark, OpenCV, Scipy, Seaborn, XGBoost

Mentorship Experience

ML Mentor (Regex Software Services): Mentored various students in theoretical and Hands-on ML projects

Campus Mentor (IIT KGP): Guided institute's juniors for various academic and non-academic activities

Extra-Curriculars

Social Work: Assisted in welfare of 100+ residents by volunteering at National Service Scheme IIT KGP
Conducted a cloth and book distribution drive for the orphan children during COVID-19

Sports : Represented school in inter-district cricket tournament held at Durgapur, WB (2016)
Actively participated and completed cycling marathon travelling 30 kms distance within 1.5 hours