Tuhin Subhra De

tuhinsubhrade@yahoo.in Gituhin tuhin-subhra-de Tuhin Subhra De +91-9475338369

Education

Indian Institute of Technology Kharagpur (IIT KGP) West Bengal, India
Majors: *B.Tech. and M.Tech. (Dual Degree) in Civil and Environmental Engineering* *Jul 19 - Apr 24*
Minors: *Mathematics and Computing, Artificial Intelligence Applications* **CGPA:** 8.7/10.0

Experience

Wharton School, University of Pennsylvania July 2023 – Present
Deep Learning Research Intern *Philadelphia, US*

- Working on explainable **Attention models** focusing on the interpretable features primarily affecting the predictions
- Analysed important features using first order **Saliency** method merged with **Layer-wise relevance propagation**
- Performed a sentiment analysis, observing the visual patterns between the embeddings and outputs whilst training

Udaan.com (Hiveloop Technologies Pvt. Ltd.) May 2023 – July 2023
Data Science Intern (led to publication no. 3) *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

Deakin University April 2022 – July 2022
Machine Learning Research Intern *Melbourne, Australia*

- Applied **Knowledge distillation** to semantically transfer information between two heterogeneous ML models
- Constrained a Multi-class Logistic Regression model (student) to **50%** and trained a CNN (teacher) to **90%** accuracy
- Passed unlabeled batches of data from the Fashion MNIST dataset obtaining the predictions from both the models
- Increased student's accuracy from **50%** to **67%** by using **KL-Divergence** on student - teacher model prediction gap

Venwiz Technologies Pvt. Ltd. Dec' 2022 – Jan' 2023
Data Science Intern *Bangalore, India*

- Enhanced website's search retrieval accuracy from **65%** to **78%** by building methods to tag client description texts
- Used ensemble of different regex and Levenshtein distance-based models to obtain most probable tags associated
- Formed a **K-means clustering** model on the vector **cosine similarity** grouping new tags and existing incorrect tag

Projects

Profit maximization in Cargo booking scenario | *IIM Mumbai* Sep' 2021 – Dec' 2021

- Devised a method to maximize profit by selective cargo acceptance, achieving **48%** more profit than random process
- Simulated stochastic booking data with Capacitated **Vehicle Routing Problem** solving for minimum delivery cost
- Formulated the above step by a Neural Network and derived a linear estimate of it for empirical cost of each cargo
- Used the empirical cost, revenue data and capacity constraints to maximize profit using **dynamic programming**

AI Face Generator Application | *Center for Excellence in AI, IIT KGP* May 2022 – June 2022

- Made a web app having **200+** viewers for generating random and customised face images from selected facial feature
- Built a **convolutional variational autoencoder** generating random human faces with data compression ratio of 96
- Devised a convolutional decoder generating customised faces from 40 facial features with data expansion ratio of 1228
- Trained both models using **PyTorch XLA** TPU(**3.5** hrs), hosted on **Git LFS** and deployed using Streamlit Cloud

Dynac Web Browser Extension | *MetaKGP Development group, IIT KGP* Oct' 2022 – June 2023

- Created a web browser extension providing dynamic and real-time updates in CV making portal of ERP, IIT KGP
- Leveraged **JavaScript** event listeners to required HTML elements in CV building form post element inspection
- Kept client-side rendering and hosted on Chrome Web Store, it reached **1000+** users within **2 months** of launch

Publications

1. ***De, T.,*** *Kartikeya, M., Bhattacharya, S.* ***A Non-Linear Lasso approach for estimating tail risk interconnectedness*** *(Under review, manuscript submitted to Applied Economics journal, submission ID: 235453376)*
   - Analysed tail risk interconnectedness during COVID-19 and Russia-Ukraine war among the public banks of Japan
   - Built a novel approach on **10yrs** data by **LassoNet** (Neural Network with LASSO) clubbed with **quantile loss**
   - Devised feature selection metric on **L1-Norm** parameter merged with **gradient sensitivity** of neural networks
   - Used the selected features to quantify an interconnectedness graph and found high market cap banks are risk-prone

2. *Balamwar. A.,* ***De. T.,*** *et.al (2022).* ***Prediction of Turn Around Time using Neural Networks - A Case Study of Shipping Port,*** *IFAC PapersOnLine, Vol 55, I-10, Pg 389-394, ISSN 2405-8963, doi: https://doi.org/kss9*
   - Predicted Turn Around Time of port ships within a certainty of **26.71** hrs initially having weeks of inconsistency
   - Compared performance of various machine learning models over **5 yrs** data of Mumbai port enhancing schedules

3. ***De. T.,*** *Singh. P, Patel. A, (2024).* ***A Machine learning and Empirical Bayesian Approach for Predictive Buying in B2B E-commerce,*** *8th International Conference on Machine Learning and Soft Computing, Singapore*
   - Built an ensemble stack approach combining boosted trees and Bayesian model giving probability for a customer buying a product. This is the outcome of the internship done at Udaan.com, more details are presented under the Experience section

Achievements

- Qualified Regional Mathematics Olympiad (**RMO**), appeared INMO — HBSE, TIFR India (2017)
- Ranked in top **0.1%** of Physics AISSCE (12th) among 2.5 million students — CBSE Board (2018)
- Secured **5542** All India Rank (General) among 22,000 candidates in country — JEE Advanced (2019)

Technical Skills

Languages: C, C++ (STL), Python, SQL, JavaScript, HTML/CSS —— **Frameworks**: Streamlit, Flask
Developer Tools: Git, Google Cloud Platform, Microsoft Azure Databricks, VS Code, Google Colab (TPU)
Libraries: Numpy, Pandas, TensorFlow, Keras, PyTorch, Scikit-Learn, PySpark, OpenCV, Scipy, Seaborn
Softwares: MySQL, MS-Office, Adobe Express, AutoCAD, ABAQUS

Relevant Coursework

Artificial Intelligence: Programming and Data Structures, (Machine Learning, Deep Learning, Artificial Intelligence)-*FA, Big Data Processing, Dependable and Secure AI/ML; **FA: Foundations and Applications*
Statistics : Probability and Statistics, Linear Algebra, Regression and Time Series Modelling

Mentorship Experience

Mentor (IIT KGP): Guided institute's junior students for various academic and non-academic activities
ML Mentor (Regex Software Services): Mentored various students in theoretical and Hands-on ML projects

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)
Participated and secured 3rd position in Mini-Marathon covering a distance of 10 kms
Actively participated and completed cycling marathon travelling 30 kms distance within 1.5 hours