Mudit Tiwari

Data Scientist
MSc in Mathematics & Computing
Bangalore, India

Email : connect[at]mudittiwari[dot]com
https:/www.linkedin.com/in/mudit-tiwari/
https://github.com/mudittiwari255
Mobile : +91-912-751-2129

Education

- **Indian Institute of Technology, Guwahati** — Guwahati, India
  Master of Science in Mathematics & Computing; CGPA: 6.9 — *July 2016 – May 2018*
- **Sri Venkateswara College, University of Delhi** — Delhi, India
  Bachelor of Science in Mathematics; CGPA: 8.7 — *July 2013 – June 2016*

Experience

- **CyborgIntell** — Bangalore, KA
  Data Scientist — *March 2019 - Present*
  - **Explainable AI**: Explainable AI is a explainability module of fully automated machine learning product. I worked on implementation and research of local as well as global Importance criteria. Implemented algorithms like null-importance, LIME, SHAP, PDP etc for PySpark, SKlearn & H2O based models.
  - **Decisioning Engine**: Developed an advance decisioning & optimization engine to take and recommend important decisions on Customer level. It is most use full for BFSI, Health and Retail sectors. This algorithm used a blended idea of Baysian optimization and partial dependence plot.
  - **Pyspark Automated Pipeline Development**: Implemented end to end pipeline for fully automated Pyspark modeling. This pipeline included(not limited to) data treatment, feature selection, model selction, hyper parameter tunning etc.
  - **Plot API**: Created set of APIs to visualize the model performance for regression/ forecasting AutoML. Some examples of the APIs are: Original Vs Predicted, distribution of residuals etc. These APIs populated the points to be plotted and then were sent to frontend for plotting excercise.
  - **Auto Documentation**: Developed an Auto Documentation module for AutoML product, this document is automatically made using Python and LaTeXafter creation of every project. This document contains results and details of each step taken by AutoML.
  - **Customer Analytics**: Worked on real industry use cases from BFSI, HR analytics, Web analytics, retail & hospitality etc. I mainly lead the path on data understanding, problem formalization, data preperation, model development, feature engineering & post model analysis.
- **Teknuance** — Chennai, TN
  Research Analyst - Machine Learning and Mathematics — *July 2018 - March 2019*
  - **Algorithm Design**: Designed the set of core algorithms for a file based database. Main elements of this set algorithms were: Data dumping into file based architecture, SQL Queries backend, caching layer development. These algorithms were later implemented in C++ by other team.
  - **Graph Based Caching Layer**: Designed the algorithm and architecture of a caching layer in a file based DB. This caching layer was based on graphs and the whole idea was to use the semanticity of queries to save on computational time.
  - **Knowledge Graph**: Researched and designed the architecture of core knowledge graph's algorithm. It is based on entity-relation-entity theory, and was used onto top of a chatbot. Later, implemented in python to go in production.
  - **Chatbot**: Trained a sequence to sequence model for training bunch of dataset for english & tamil language chatbot. The deeplearning framework used was keras & pytorch.
  - **Character Recognition**: Developed a Image classification algorithm for tamil character recognition, the algorithm was trained using pytorch and architecture used was RESNET50.
  - **Text to SQL**: Researched and used NLP fundamentals, and sequence to sequence modeling for training a model which turned the normal english to SQL query.
- **Teknuance** — Chennai, TN
  Research Intern - Mathematics — *Summer 2018*
  - **Distribution Differentiator**: Developed the distribution differentiator, which used several statistical tools & regular expressions to hypothesise if two sets are from same distributions or not.
  - **Meeting Schedular**: Created a python based application to schedule internal meetings. This was used by maintaining a global meeting portal, and every employees calender. It handled the slot collision and provided the recommendation to user of favourable time for employees.

Projects

- **Flocking Algorithm: A Simulation Study**: GUI based simulation of Flocking Algorithm using Python. Also came up with a new flocking algorithm using KD Tree to reduce the overall time complexity of the old algorithm. This project was done under guidance of Prof. Partha Sarthi Mandal.
- **Restricted Boltzman Machine**: Implaementation of geoffrey hinton's famous paper on Dimensionality Reduction using RBM. I used python and tensorflow to build and train model from scratch.

Technical Skills

- **Languages**:Python, C, MATLAB
- **ML Technologies**: Numpy, Pandas, SKLearn, XGboost, CatBoost, LightGBM
- **Frameworks**: Pytorch, H2O, PySpark
- **Miscellaneous**: Linux, Vim, GCP, Excel

Achievements

- **Academics**
  - **Joint Admission for Masters - 2016**: Secured AIR 249 among 10000+ candidates in JAM 2016 taken in Mathematics.
  - **GATE - 2018**: Qualified GATE MA with 902 rank.
  - **Inspire Scholarship**: Received Inspire Scholarship (2013 batch) from MHRD, Govt of India.
  - **Other**: JEE Main-40234/1.2M, Qualified SAU entrance with rank 61, Qualified NSTSE.
- **Competitive Data Science [Only Showing Few]**
  - **Analytics Vdihya - Bain & Company Forecasting Challange**: Secured rank 2 among 2500+ participants.
  - **Analytics Vidhya - AmExpert by American Express**: Secured rank 5 among 3000+ participants, won the prize money, and top5 medal.
  - **Kaggle - Ashrae Great Energy Predictor**: Got rank 53 among 3600+ submissions, won a silver medal.
  - **Kaggle - Catch Me if You can [Intruder Prediction]**: Got 84 rank among 3800+ submissions.
  - **HackerEarth - HDFC ML Hackathon**: Got rank 12 among 3000+ registered candidates.
  - **HackerEarth - LMG ML Hackathon**: Got rank 20 among 2000+ registered candidates.