Mehul Jain

857-313-1399 jain.meh@northeastern.edu linkedin.com/in/mehuljain22/ github.com/mehulfollytobevice
Boston, MA | Available: May 2023 - December 2023

Education

Northeastern University, Khoury College of Computer Sciences Boston, MA
Master of Science in Data Science | **GPA: 4.00/4.00** Aug 2022 – May 2024 (Exp.)
Courses: *Supervised Machine Learning, Natural Language Processing Introduction to data management and processing, Algorithms*

Vellore Institute of Technology Vellore, India
Bachelor of Technology in Computer Science and Engineering | **GPA: 8.68/10.00** Jul 2018 – Jun 2022

Technical Skills

Languages: Python, Java, SQL, Go, Julia, C++, R

Databases: MySQL, Google BigQuery, MongoDB, Cloud Firestore, Hadoop

Frameworks: Beautiful Soup, PySpark, Scikit-Learn, PyTorch, Fast AI, Tensorflow Keras, Hugging Face, Matplotlib, Seaborn, Dash-plotly, Selenium

Publications

Deep Learning based intrusion detection system for IoT networks **July 2022**
- **Title used in journal:** A new Deep Learning approach enhanced with Ensemble Learning for accurate intrusion detection in IoT networks
- Accepted in Ad Hoc and Sensor Wireless Networks.
- **Keywords:** Deep learning, Ensemble learning, Machine learning, Intrusion Detection System, Cyber Security, Bot-IoT dataset.

Extreme Gradient Boosting for Toxic Comment Classification **November 2021**
- Published in Computational Methods and Data Engineering, Proceedings of ICCMDE 2021
- **Keywords:** XGBoost, Ensemble Learning, Natural Language Processing, Machine Learning.

A brief report on transfer learning for handling small data sets **May 2021**
- Presented at International Symposium on Computational Intelligence Issues in Block chain Technology, AI and ML, VIT Vellore.
- **Keywords:** Transfer learning, Deep learning, Convolutional Neural Networks, FastAI.

Data Extraction and Sentimental Analysis from "Twitter" using Web Scrapping **October 2019**
- Published in International Journal of Engineering and Advanced Technology (IJEAT).
- **Keywords:** Sentimental Analysis, Web Scrapping, Web Extraction, Classification Of Data using Machine Learning Algorithms.

Projects

Galaxy Morphology Classification using CNN **October 2022 - December 2022**
- Created a CNN based deep learning system to classify the morphology of distant galaxies in the universe. The dataset consists of 61,578 high quality images with 37 corresponding class probabilities. Project is based on the Galaxy Zoo Challenge on Kaggle.
- Deployed 5 different architectures of CNNs designed from scratch. Also utilized pre-trained architectures like Resnet-18 to improve the performance on the above task.
- Best model acheived a RMSE score of 0.1009 which corresponds to a rank in the top 50 on the Kaggle leaderboard.

Image Captioning System **September 2021 - October 2021**
- Collaborated with other interns to design an image captioning system based on the Flickr8K data set that contains over 8000 training images with associated captions.
- Integrated InceptionV3 pre-trained model from tensorflow keras and GloVe embeddings to construct CNN and LSTM architectures.
- Developed an interactive GUI application using Voila library to input images and easily generate captions.

Automating Legal Contract Reviews: A Transformer based Q & A System **January 2022 - June 2022**
- Developed a deep learning based system to expedite and facilitate contract reviewing process. The system automatically highlights important parts of a legal document based on queries from the user.
- Compared performance of 6 contemporary transformer architectures (BERT, DistilBERT, BART, BigBird, ELECTRA, Longformer) using different hyper parameters.
- Created a simple and straightforward application for end users without a technical background.

Machine Learning Micro-Project Repository **February 2020 - December 2021**
- Created an open source repository with over 30 micro projects based on numerous machine learning algorithms (e.g. Random Forest, XGBoost, SVN, K-NN, Logistic Regression, K-Means, Apriori, Thomson Sampling, etc).
- Contains easily understandable code to facilitate learning and reproducibility.

Leadership / Extracurricular

VIT Mathematical Association May 2020 – January 2021
Computer Science Project Chair VIT, Vellore

SOS - VIT Psychology Club (Smile Over Stress) January 2020 – January 2021
Co-founder/General Secretary VIT, Vellore