Pranav Kamesh Sivakumar

ps4521@nyu.edu | (934) 210 9929 | Jersey City, NJ | LinkedIn | GitHub

EDUCATION

Master of Science in Computational Science – New York University – NY, USA Sep 2021 – May 2023
Coursework – Natural Language Processing, Computer Vision, Foundations of Finance, Fundamental Algorithms, Introduction to Deep Learning Systems, Numerical Methods 1, Programming Languages, High Performance Computing. **GPA: 3.52**

Bachelor of Technology in Chemical Engineering – IIT Madras, Chennai, India Jul 2017 – May 2021
Coursework – Computational Techniques, Multivariate Data Analysis, Probability, Statistics and Stochastic Processes, Foundations of Data Science, Pattern Recognition and Machine Learning, Differential Equations, Principles of Economics. **GPA: 8.38**

SKILLS

Programming Languages: Python, C++, R, MATLAB

Software: SQL, Tensorflow, Keras, PyTorch, Power BI, Pandas, Git, Google Cloud Platform, Numpy, Scikit-Learn, HuggingFace, NLTK, OpenCV, Flask, Slurm, Singularity, Bash, PostgreSQL, PySpark, Jupyter, AWS, OpenMP, MPI, CUDA

Data Science: Machine Learning, Deep Learning, Natural Language Processing, Pattern Recognition, Computer Vision, Algorithms and Data Structures

RELEVANT EXPERIENCE

Machine Learning Engineer – Deep Market Making Inc – Part Time Feb 2023 – May 2023
• Made a Proof of Concept for obtaining **Feature Importance** using **Shapley** values for the task of Corporate Bond price prediction
• Built **Normalization** architecture for Machine Learning models using **TensorFlow** and **Keras** frameworks.

Data Engineering and ML Intern – PartnerTap May 2022 – Dec 2022
• Created a **Normalization** script using **Flask** and **SpringBoot** to improve the performance and speed of existing Data Normalization methods.
• Solved the problem of **Duplication** in **Account Matching** by writing a code module using **Machine Learning** to handle deduplication for **20 million** records on AWS **EC2** instance within **30 minutes**, using **PostgreSQL** integration within **Python**.
• Built an **in-house** tool for **Account Matching** using **Machine Learning**, that performs **Record Linkage** between 2 data sources of size **1 million** records each by utilizing **Pandas**.

Automation Intern – AB-InBev May 2020 – Jun 2020
• Created code modules for Bot Automation using **Automation Anywhere** for collection of Employee Hiring data.
• Generated and verified **50** reports by creating code modules to automate Data Retrieval from **SAP** Database.
• Worked in an **Agile** Workspace environment, and generated Service Level Templates using **Power Query** and **Excel** for **Data Analysis**.
• Created live Data Visualizations using **Power BI** dashboards to track exports for use by the shipment team.

Data Analytics Intern – Cadla PTE ltd Apr 2019 – Jun 2019
• Cleaned the raw sales data from **1000** Supermarkets and departmental stores using **R** language and organized it on Google Cloud.
• Utilized **GCP** tool **BigQuery** for Data Warehousing of Sales data and verified correlations between sales volume and important annual events.
• Made Product and Brand based sales comparisons and visualized the findings using **R** language.

PROJECTS

Satellite Image Segmentation of Land Cover – Computer Vision – Multi Class Classification - NYU Sep 2022 – Dec 2022
• Trained **U-Net** models with **ResNet-18 and ResNet-50 encoder**s on the **LandCover.ai dataset** for **Satellite Image Segmentation**
• **Augmented** the given dataset with relevant **transformations** to balance the dataset, which improved the results.
• Compared the performance of Vanilla U-Net model with U-Nets with ResNet encoders, for both the original and Augmented datasets.
• Obtained a Classification accuracy of **0.88** for U-Net+ResNet-18 model, which is in line with previous benchmarks.

Unlearning Dataset Bias – Textual Entailment - Natural Language Processing – NYU Feb 2022 – May 2022
• Utilized concepts from **He et al's work** – (Unlearn Dataset Bias in NLI by Fitting the Residual) to train a BERT model for **textual entailment task**.
• Obtained an accuracy of **0.898** with **BERT** model on **the Stanford NLI** dataset, improving upon the benchmark set in He's work.

Emotion Masked Language Modelling – Natural Language Processing – NYU Nov 2021 – Dec 2021
• Reproduced the results of ***Sosea et al* 's work** – (eMLM: A New Pre-training Objective for Emotion Related Tasks), which introduced a new pre-training objective, **emotion Masked Language Modeling(eMLM)** to **BERT** models to improve their performance specifically for **Sentiment Analysis and Emotion Detection** tasks.
• Obtained similar results to the paper upon testing on 2 Sentiment Analysis and Emotion detection datasets.

Facial Emotion Detection – Image Processing – IIT Madras Sep 2020 – May 2021
• Implemented **a 2 stream Neural Network** to consider **Micro-Expressions** in **Facial Emotion Detection**.
• Showcased the difference in performance of 2 widely used Emotion Detection Models on Indian and mixed video datasets.
• Implemented procedures for generating spontaneous emotions and created an **Indigenous video emotion dataset**.
• Compared and verified the results with parallel Eye tracking and EEG wave analysis results for Emotion detection.

Data Contest – Pattern Recognition and ML – Classification – IIT Madras May 2020 – Aug 2020
• Developed a **model** to **predict Stay/Exit of an employee** from a company based on ratings and remarks given.
• Collected useful features after pre-processing data and appropriately **normalized** the features.
• Submitted an **XGBoost** model with **K-fold Cross Validation**, which scored a weighted accuracy of **0.8488** against the baseline of 0.7234.

ADDITIONAL EXPERIENCE

Graduate Teaching Assistant – Algebra and Calculus, Calculus I, Mathematics for Economics III – NYU Sep 2021 – May 2023
• Communicated Calculus concepts, cleared doubts, held office hours and graded Assignments for 200+ students in total.

AWARDS AND HONORS

Guinness World Record holder – Largest Multi-Sudoku Solved Puzzle. Represented India at the World Sudoku and Puzzle Championships from 2015-2019, 2022. Asian Sudoku Championship Gold Medallist. Times Sudoku Champion 2017.