Saurav Chennuri

Machine Learning Engineer with over 1.5 years of experience. Majored in Engineering and **Artificial Intelligence**. Expertise in NLP, Information retrieval, recommender systems and machine learning.

Proficient in **Python**, **C++**, **Pytorch**, **Tensorflow** and **SQL.**
Actively looking for MLE/SDE roles

Boston, MA | +1(484)320-9175 | saurav.varaprasad@gmail.com | Homepage

Selected Projects

Contrastive Learning to improve text embeddings for low resource language News Articles: ***(Natural Language Processing)*** **[Github]**

- Inspired by OpenAI's CLIP, I contrastively finetuned an ***mT5 transformer(small)*** to update Chichewa(Nyanja) language news article embeddings to be closer to its parallel english translated new article embeddings.
- English being a high resource language with the language models learning better representation space, a low resource language to have similar embedding structure could benefit from this. This was done for better classification accuracy of the Nyanja news article genre.
- Also implemented ***Mixup*** and ***NLPAug*** based text data augmentation approaches to address class imbalances.
- Final overall news article classification accuracy reached **61.2%** with enhanced embeddings and a ***DNN classifier***, while the accuracy varied between **50-60%** for CountVectorizer and TFIDF embeddings with Support Vector Machines, Random Forests, Naive Bayes and XGBoost classifiers. Observed that better representation of embeddings is more important than the classifier used.

Gender Bias Quantification from Knowledge Graphs in hyperbolic spaces: **[Github]**
(Knowledge Graphs | Deep Learning | Ranking)

- Created embeddings in Poincare disk space of Hyperbolic dimension for Knowledge graphs of English, Swedish and Indonesian language obtained from DBPedia, queried through **SPARQL**
- Leveraged **Pytorch** to implement Mobius addition, matrix multiplication, which are euclidean equivalent in hyperbolic spaces, and leveraged those operations for implementing **Riemann Optimization** for network update.
- Tilted the named person entities towards male and female entities (relation: gender entity) and checked the profession score for that named person entity before the tilt and after the tilt. This approach averaged across all named entities gave a score on the bias of a profession towards one of the sensitive entities (male, female).

Manipulating SGD for Data Ordering Attacks in Deep Learning approaches: (Deep learning | Machine Learning) **[Github]**

- Implemented a series of adversarial attacks on **ResNets** and **Vision transformers (Vit-b-16)** based on batch reordering and reshuffling methods in 4 attack policies to either reduce the performance of the model or slow down the training process.
- These **blackbox attack** methods exploit the stochasticity of the gradient descent update rule over the epoch, and make it very hard to observe and mitigate the attack.
- **ResNet-18,50** had a classification performance drop of 3-60% on CIFAR-10,100 benchmarks, while Vit-b-16 had a performance drop of 3-30% on SVHN benchmark. Surrogate model used is Lenet-5.

Education

Master of Science, Artificial Intelligence
Boston University, Boston, MA, (2021 - 2023); **GPA: 3.9**

Bachelor of Technology, Engineering Sciences
Indian Institute of Technology, Hyderabad, India, (2016 - 2020).

Publications

Fusion Approaches to Predict Post-Stroke Aphasia Severity from Multimodal Neuroimaging Data. [Github | Paper]
Saurav Chennuri, Sha Lai, Anne Billot, Maria Varkanitsa, Emily Braun, Archana Venkataraman, Janusz Konrad, Swathi Kiran, Prakash Ishwar, Margrit Betke;
Published at **ICCV** Workshop **Computer Vision for Automated Medical Diagnosis, 2023**.

Feature Analysis and Extraction for Post-Stroke Aphasia Recovery Prediction [Paper]
Saurav Chennuri, Anne Billot, Sha Lai, Prakash Ishwar, Margrit Betke, Swathi Kiran;
Published in **Medical Imaging Understanding Analysis Conference, 2022.**

Towards Fast Crash-Consistent Cluster Checkpointing [Github | Paper]
Andrew Wood, Moshek Hershcovitch, Ilias Ennmouri, Weiyu Zong, Saurav Chennuri, Sarei Cohen, Swaminathan Sundararaman, Daniel Waddington, Peter Chin.
Published in **IEEE conference on High Performance Extreme Computing, 2022.**

Sloshing Noise Classification in Fuel Tanks of Hybrid Vehicles using Convolution Neural Networks. [Paper]
Golla Siva Teja, Chennuri Saurav Vara Prasad, B. Venkatesham, K Sri Rama Murthy.
Published in the **Journal of Acoustic Society of America, 2021.**

Skills

Programing Languages
Proficient in: ***Python***, ***C++***, *Java*, ***SQL***
Working Knowledge: Shell Script, C#.
Technologies:
Pytorch, Pytorch Lightning, ***TensorFlow***, TensorBoard, Wandb, ***Transformers***, Pandas, Docker, PowerBI, ***Github***

Work Experience

Fractal Analytics

Data Analyst
Oct 2020 - Aug 2021

Data Analytics and DevOps

- Designed and implemented azure data pipelines from **Azure Datalakes**, leveraging azure data factory.
- Developed **powerBI** reports connected to the data pipelines, enabling data visualization and insights. Automated pipeline deployment and report publishing using Azure DevOps, ensuring efficient **CI/CD processes.**

Web Applications and Cloud Services

- Created a **web-app** as a prototype to fetch financial data, to have an interactive user experience powered by google Dialog Flow. This served as a virtual assistant to provide real time support and data.
- Set up **webhooks** in Dialog Flow to trigger HTTP callbacks to the backend server whenever the chatbot needed to fetch dynamic data or perform actions beyond static responses

Takenaka Corporation

Machine Learning Engineer Intern
May 2019 - Jul 2019

- Developed an ensemble model by leveraging **Xgboost**, **SVMs** and **Random Forests** to predict the required workforce per day through the construction duration for diverse construction projects, achieving a variance of 10 workers per day.
- Delivered comprehensive presentations to explain the observations to many cross-functional stake-holders to understand the impact of the work.

Academic Experience

Center for Brain Recovery, Boston University

Research Assistant
Jan 2023 - May 2023

- Worked on Quantifying the severity of Aphasia among the patients using **Machine Learning** approaches on the **multimodal data** of several kinds of fMRI, MRI scans and DTI images, through machine learning approaches.