Saurav Chennuri

Machine Learning Engineer / Software Engineer
Looking for full time roles starting **January 2024**

Boston, Massachusetts • +1-484-320-9175 • sauravchennuri@gmail.com • [Homepage](#) • [Google Scholar](#) • [Linkedin](#)

EDUCATION

Boston University, Boston, MA   05/2023
Master of Science — Artificial Intelligence
Indian Institute of Technology - Hyderabad, Hyderabad, India   05/2020
Bachelor of Technology —Engineering Sciences

SKILLS

Interests: Data Mining, Language and Vision, Pattern Recognition..
Cloud Services: Google Cloud, AWS cloud services, Snowflake.
Languages: Python, C++, Shell Scripting, Java, MySQL, FORTRAN.
Machine Learning: Pytorch, Pytorch Lightning, TensorFlow, TensorBoard, Wandb, Transformers, Pandas, scikit-learn.
Tech Stack: Docker, Tableau, PowerBI, DeepLabCut, Github, Kubernetes, Terraform, LSF, Linux(Ubuntu 16.04/20.04, MacOS), Github, GitLab, Jira, MATLAB,
Certifications: Snowflake Snowpro Core, CITI Human Subjects Protection Training: Social and Behavioral Focus.
Courses: Deep Learning, Computer Vision, Natural Language Processing, Operating Systems, Databases, Statistics, Probability.

SELECTED PROJECTS

Manipulating SGD for Data Ordering Attacks in Deep Learning approaches:
- 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.

Contrastive Learning to improve text embeddings for low resource language News Articles:
- 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 language 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 SVM, 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:
- Created embeddings in Poincare disk space of Hyperbolic dimension for Knowledge graphs of English, Swedish and Indonesian language obtained from DBPedia, queried through SPARQL
- Hyperbolic spaces are functions of negative curvature, of which spherical and Poincare disk are two forms of such space. The distance between entities increases exponentially (linearly for Euclidean) as we go radially outward making them attractive for embedding multi-relational knowledge graphs.
- 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).

Movie Recommender system - (*During undergraduate studies - 2018*)
- Built a movie recommender system using content based filtering, collaborative filtering and a hybrid of these two recommender systems, to recommend movies on Movielens(small) dataset
- For a content based recommender system, I build the item(movie) profile by vectorizing the movie description with TF-IDF vectorizer along with their corresponding keywords. Based on the cosine similarity of all the remaining movies with the chosen movie, I recommend the top - k most similar movies.
- For collaborative filtering based recommender systems, I used the *user rating - movie* utility matrix, and constructed the user profile matrix and item (movie) profile matrix through matrix decomposition approach using SVD decomposition with SGD based optimization and RMSE as a loss function.
  - Based on the constructed user and movie matrices, I filled in the sparse utility matrix through their corresponding dot products, and recommended the top-k best rated movies via interpolation to a user
- For the hybrid recommender system, given a user ID and an initial movie, I first shortlist top-k most similar movies through content based filtering and among those movies I use collaborative filtering to recommend the movies that the user might like.

WORK EXPERIENCE

Fractal Analytics, Bangalore, India   09/2020 – 08/2021
Artificial Intelligence startup with 2500+ employees and valued at $750 million.

Imagineer - Data Engineering division
- Created several intents and their corresponding responses in making a chatbot using Google DialogFlow Enterprise. Deployed webhook service to fetch the corresponding data from google BigQuery database at the backend to the DialogFlow interface. This application helped in reducing customer support response time.
- Contributed to end-to-end implementation of PowerBI dashboards for cross-functional teams, enabling real-time tracking of key performance indicators and facilitating data-driven decision making, resulting in an increase in overall efficiency.

Takenaka Corporation, Tokyo, Japan   05/2019 – 07/2019
A well established 400 year old Japanese architectural firm. Valued at more than 1 billion dollars.

Machine Learning Engineer Intern - Information and Communication Technology (ICT) group
- 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

Research Assistant, Center for Brain Recovery, Boston, MA   01/2023 – 05/2023
Academic Teaching Assistant, Boston University, Boston, MA   09/2021 – 06/2022
- Made course material and conducted lab sessions for different Computer Science courses for more than 140 students in each class. **Courses taught**: Computer Systems, Computational Tools for Data Science, Introduction to Computer Science.

PUBLICATIONS

Fusion Approaches to Predict Post-Stroke Aphasia Severity from Multimodal Neuroimaging Data
[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, Oct 2023.**

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

Towards Fast Crash-Consistent Cluster Checkpointing

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, Mar 2022.**

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