Piyush Chawla

Personal Website: https://pijusch.github.io/
Contact: +1 380-895-4401, Columbus Ohio
chawla.81@osu.edu
https://www.linkedin.com/in/chawla-piyush/
google-scholar

Deep Learning, Machine Learning

Education (The Ohio State University, Columbus OH)

Research: Applied Machine Learning, Model Interpretability, Natural Language Processing

The Ohio State University **GPA: 3.98/4.0**
*Master of Science, Computer Science Engineering (**Transferred from Ph.D.**)* **2018 - Dec 2022**

Birla Institute of Technology & Science **GPA: 9.7/10.0**
Bachelor of Engineering (Hons), Computer Science Engineering **2014 - 2018**

Technical Skills

Programming Python, Pytorch, Hugging-Face, Torch-Geometric, Git, Version Control, Scikit Learn, SciPy, Gensim
Key Courses Machine Learning, NLProc, Data Mining, Optimization, High Performance (Distributed) Deep Learning

Research/Work Experience

PhD Student, The Ohio State University, Columbus, OH **Aug18 - Now**

- GRAVITY Lab (Dr. Prof. Han-Wei Shen)
- Graduate Teaching Assistant

Research Scientist Intern, Bosch Research, Sunnyvale, California **May - Aug 22**
Project: Explainable Graph Similarity for Workflow Graphs

- Developed Explainable AI solution for the problem of Graph Similarity on PascalVOC keypoints.
- Leveraged Self Supervised Learning (SSL) to achieve SOTA on node alignment based explainability.
- Result: Proposed novel structure centric SSL pretext tasks and ML framework to push limits of Graph XAI.
- Result: Filed research patent for the idea.

Research Intern, IBM Almaden, San Jose, California **May - Aug 21**
Project: Log Anomaly Detection

- Generated a large-scale dataset from publicly available GitHub repositories.
- Designed downstream tasks to gauge the quality of the log dataset.
- Trained language models Doc2Vec and RoBERTa for log anomaly detection tasks achieving SOTA.
- Result: Plugged-in these language models into IBM's proprietary log anomaly detection pipeline.

Research Assistant, Smart Data Analytics, Bonn-Germany **Dec 17 – July 18**
Bachelor's Thesis: Link prediction in Multi-Lingual Knowledge Graphs (Python, TensorFlow)

- Leveraged interlanguage links (e.g., English-German) in DBpedia knowledge graph to train KG embeddings.
- Implemented the mTransE model to improve link prediction accuracy.
- Result: Novel method dubbed NLTransE to add semantic knowledge from word embeddings to KG embeddings.

Intern (Research Scholar), University of Bonn DAAD WISE (1 of 100 all-over India) **May – Jul 17**
Project: Knowledge graph completion using latent vector models (Python, TensorFlow, C++)

- Explored different latent-vector models for KG link prediction. TransE, TransR, TransH, DistMult etc.
- Result: Conducted large-scale experiments on DBpedia Knowledge Base to train KG completion models.

Key Projects

Question Answering On Table Data *(Python, Pytorch)* **Fall 2022**

- Working to develop novel Question Answering method for hybrid data (structured and unstructured).

Pattern Recognition in Academic Publications *(Python, Pytorch, D3, JavaScript)* **Spring 2022**

- Clustering of text document is commonly used in pattern recognition but explaining clusters is a challenge.
- Proposed a novel clustering approach in topic model latent space using Mixture of Gaussians.

- Used clusters to find interesting research themes emerging in the academic publications.
- Result: A novel visualization system to discover temporal trends in documents. (Publication Under Review)

Probing Static Word Embeddings for Relational Information *(Python, Pytorch)* **Spring 2021**

- Relation induction has been used to uncover relational information in static word embeddings.
- The existing works focus on linear models and only consider vector offset as the feature.
- We developed MLP based non-linear probes and found that features like vector product and sum contain significant amounts of information. Achieved State of the art (SOTA) on relation induction problem.
- Findings: Publicly available GloVe embeddings contain richer information than Skipgram.
- As an extension, we train models in meta learning setting (MAML, Reptile), improving the few-shot prediction.

Understanding Convolutional Neural Networks for text *(D3, JavaScript, Python, Flask)* **Spring 2020**

- ConvNet visualization (explanation) has been explored for computer vision. However, understanding this class of architecture remains an open problem in the context of natural language (text) applications.
- Result: Developed a novel approach (Token-wise Sentiment Detection) to visualize the contribution (+ve/-ve) of each word in a sentence towards the predicted label (sentiment).
- Result: Discussed adversarial and error analysis strategies for ConvNet-based sentiment-analysis classifiers.

Finished Publications

- Piyush Chawla, S. Hazarika, HW Shen (2020) Token-wise sentiment decomposition for ConvNet: Visualizing a sentiment classifier. PacificVis 2020
- Piyush Chawla, D. Esteves, K. Pujar, J. Lehmann (2019) SimpleLSTM: A Deep-Learning Approach to Simple-Claims Classification. EPIA-2019
- D. Esteves, J. Marcelin, Piyush Chawla, A. Fischer, J. Lehmann (2021) HORUS-NER: A Multimodal Named Entity Recognition Framework for Noisy Data. IDA 2021
- D. Esteves, A. J. Reddy, Piyush Chawla and J. Lehmann (2018) Belittling the Source: Trustworthiness Indicators to Obfuscate Fake News on the Web. EMNLP 2018

Awards and Achievements

- University Fellowship, The Ohio State University 2018-2019
- Bachelor's Thesis Scholarship, BITS Pilani 2017-2018
- Thesis Research Scholarship, University of Bonn 2017-2018
- WISE scholarship, Deutscher Akademischer Austauschdienst (DAAD WISE) Summer 2017
- Merit scholarship, BITS Pilani 2014-2018

Professional Service

Secondary Reviewer

IEEE TKDE 2019, KDD 2020, IEEE VAST, ACL 2021, ISVC 2021