

Pujan Paudel

🏠 pujanpaudel.com ✉ ppaudel@bu.edu 📍 Boston, MA ☎ 725-724-0359 🌐 [codepujan](https://codepujan.com)

EDUCATION

Ph.D. in Computer Engineering, Cumulative GPA: 3.93 *September 2020 - August 2025 (Expected)*
Boston University (BU) *Boston, MA*
Research Areas: Natural Language Processing (NLP), Machine Learning (ML),
Information Retrieval (IR), Deep Learning (DL) and Computer Vision (CV)

Bachelor of Science (Honors) in Computer Science, Cumulative GPA: 3.75 *August 2016 - May 2020*
The University of Southern Mississippi (USM) *Hattiesburg, MS*
Honors Thesis: *The Blind Spot of Twitter Bot Moderation*

INDUSTRY EXPERIENCE

Machine Learning (ML) PhD Intern at **Truveta** *September 2024 - December 2024*

- Developed a Retrieval Augmented Generation (RAG) system for automatic semantic normalization of unstructured clinical data to standard medical ontologies.
- Achieved a 36% improvement in term mapping coverage by integrating the RAG pipeline with Truveta's large-scale electronic health record (EHR) datasets.

Natural Language Processing (NLP) Research Intern at **Merck Research Labs** *June 2024 - August 2024*

- Curated biomedical literature datasets, benchmarked state-of-the-art NLP models and developed a hybrid transformer-LLM pipeline for biomedical document classification.
- Integrated the document classifier pipeline on proprietary stakeholder data, improving identification of target biomedical documents by up to 8%.

ACADEMIC EXPERIENCE

Graduate Research Assistant at **Boston University Security Lab** *September 2020 - Present*

- Advanced discoverability of e-commerce scam websites for upstream online fraud mitigation.
 - ❑ Developed a feature distillation-based framework for Search Engine Ranking Pages (SERP) driven query mining system, ranking search queries by their propensity to discover scam websites.
 - ❑ Expanded discovery of scam websites by 21 times relative to baseline NLP models, accelerating discovery of emergent online scams in the wild.
- Enhanced contextual capabilities of moderation systems for precise flagging of false content [2].
 - ❑ Developed and implemented a new task for unsupervised stance detection called *Contrastive Textual Deviation*, leveraging transfer learning in Google's FLAN-T5 Large Language Model.
 - ❑ Achieved state-of-the-art stance detection performance on 3 different datasets reducing false positives of automated content flagging by 10 times.
- Improved identification of image-based misinformation at Twitter for visual content moderation [3].
 - ❑ Implemented an end-to-end reverse image search system for million-scale social media images search using perceptual hashing and Milvus.
 - ❑ Improved precision and relevance of visual similarity matching by 8%, enhancing visual content flagging by 13 times on Twitter.
- Built an automated keyword identification system to track tweets spreading false claims [4].
 - ❑ Developed and implemented an end-to-end Learning To Rank (LTR) based keyword extraction system outperforming traditional statistical and semantic similarity-based techniques.
 - ❑ Improved flagging of misleading textual content by 20 times on Twitter.

Research Assistant at **The University of Southern Mississippi** *January 2017 - May 2020*

- Conducted and published research in Twitter social bots using tools from topic modeling, network science, and information diffusion [5,6].
- Developed conversational agents using Amazon Alexa assisting USM Psychology department in feasibility study of voice assistant technology in retention of single-digit mathematical calculations for children with cognitive challenges.

TECHNICAL SKILLS

Programming: Python, C/C++, Javascript, Java, C#, Bash Scripting, R

Machine Learning: Scikit-learn, PyTorch, Pandas, WandB, OpenCV, SciPy, NetworkX

Technologies: Elasticsearch, Lucene, AWS, Azure, Docker, Flask, Spark, React.js, Django, Node.js, Github, vLLM, TGI, AWS SageMaker, Google Cloud Platform, Ray

Database: Milvus, FAISS, Cassandra, MongoDB, SQL, PostgreSQL

Concepts: Data mining, Exploratory Data Analysis, Information Retrieval, Knowledge Distillation, Knowledge Graphs, Generative AI, Transformers, Computer Vision, Large Language Models, Multi-Modal LLMs, Human-Computer Interaction, Vector database, Big Data, Cloud Computing

ADDITIONAL PROJECTS

American Rescue Plan Act (ARPA) Bill Earmark Analysis

March 2022 - May 2022

- Built a custom Named Entity Recognition (NER) model using Spacy to automatically infer policy buckets from ARPA amendment language, analyzing disproportionate distribution of earmarked funding

Scaling Remote Sensing Data Processing With Ray

September 2021 - December 2021

- Setup OpenTelemetry, Ray and Jaeger in Mass Open Cloud (MOC) for distributed profiling and finding bottlenecks on a NASA-JPL remote sensing application, proposing a new parallelization scheme with 3x speedup

Cyberwarfare: Longitudinal Trends and Effects on Foreign Policy

May 2021 - July 2021

- Crawled, compiled, and normalized a dataset of state-sponsored cyber attacks from three different data sources to analyze how cyber-severity of future attacks changes as an effect of policy actions between rival countries

AWARDS AND HONORS

Finalist, Best Applied Research Award, **CSAW**

2024

Usenix Security '24 Travel Grant, **Usenix Association**

2024

Pardee Center Graduate Summer Fellowship, **BU**

2021

Distinguished ECE PhD Fellowship, **BU**

2020

Runner Up, Undergraduate Research Symposium, **USM**

2019

Best Innovation Application, **CalHacks 2016**

2016

Best IBM Watson Hack, **HackRice 2016**

2016

SELECTED PUBLICATIONS

- [1] M.H. Saeed, S. Ali, **P. Paudel**, J. Blackburn, G. Stringhini, "Unraveling the Web of Disinformation: Exploring the Larger Context of State-Sponsored Influence Campaigns on Twitter," 27th International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2024), Padua, Italy, 2024.
- [2] **P. Paudel**, M.H. Saeed, R. Auger, C. Wells and G. Stringhini, "Enabling Contextual Soft Moderation on Social Media through Contrastive Textual Deviation," 33rd Usenix Security Symposium, Philadelphia, PA, USA, 2024.
- [3] **P. Paudel**, C. Ling, J. Blackburn and G. Stringhini, "PixelMod: Improving Soft Moderation of Visual Misleading Information on Twitter," 33rd Usenix Security Symposium, Philadelphia PA, USA, 2024.
- [4] **P. Paudel**, J. Blackburn, E. De Cristofaro, S. Zannettou and G. Stringhini, "Lambretta: Learning To Rank For Twitter Soft Moderation," 2023 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2023.
- [5] M. Singhal, C. Ling, **P. Paudel**, P. Thota, N. Kumarswamy, G. Stringhini, and S.Nilizadeh "SoK: Content Moderation in Social Media, from Guidelines to Enforcement, and Research to Practice," 2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P), 2023.
- [6] **P. Paudel**, TT. Nguyen, A. Hatua and AH. Sung, "How the tables have turned: Studying the new wave of social bots on Twitter using complex network analysis techniques," 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, Canada, 2019.
- [7] **P. Paudel**, TT. Nguyen, A. Hatua and AH. Sung, "User Level Multi-feed Weighted Topic Embeddings for Studying Network Interaction in Twitter." 2019: 8th International Congress, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, 2019.