Pujan Paudel

pujanpaudel.com | ppaudel@bu.edu | Boston, MA | 725-724-0359 | codepujan

Education

Boston University (BU) — *Boston, MA*

Ph.D. in Computer Engineering — *January 2026*

Dissertation: *Data-Driven Approaches for Improving the Identification of Misleading content online*

Research Areas: Natural Language Processing (NLP), Machine Learning (ML),
Deep Learning (DL), Information Retrieval (IR), Computer Vision (CV), and Cybersecurity

M.S. in Computer Engineering, GPA: 3.93 — *December 2023*

The University of Southern Mississippi (USM) — *Hattiesburg, MS*

B.S. (Honors) in Computer Science, GPA: 3.75 — *May 2020*

Honors Thesis: *The Blind Spot of Twitter Bot Moderation*

Work Experience

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

- Developed an end-to-end Retrieval Augmented Generation (RAG) system for automatic semantic normalization of unstructured clinical data to standard medical ontologies.
- Improved term mapping coverage by 36% while deploying RAG system across Truveta's large-scale Electronic Health Record (EHR) datasets.

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

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

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

- Proactively discovering e-commerce scam websites for upstream online fraud mitigation [1].
  - ❏ Conceptualized and developed *Loki*, a feature distillation-based framework for Search Engine Ranking Pages (SERP) driven query mining system.
  - ❏ Expanded discovery of scam websites by 21 times relative to baseline NLP models, accelerating discovery of emergent online scams in the wild.
- Contextual content moderation system for precise flagging of harmful content [4].
  - ❏ Developed and implemented a new task for unsupervised stance detection leveraging transfer learning and fine-tuning Google's FLAN-T5 model.
  - ❏ Achieved state-of-the-art stance detection performance across 3 diverse social media datasets reducing false positives of automated content flagging 10 folds better.
- Identification of harmful media at Twitter for multimodal content moderation [5].
  - ❏ Implemented *PixelMod*, an end-to-end reverse image search system for million-scale social media image 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.
- Context adaptive keyword extraction system to track tweets spreading harmful content [6].
  - ❏ Developed *Lambretta*, a supervised Learning To Rank (LTR) and Open Information Extraction (OIE) based keyword identification system.
  - ❏ Improved the flagging of harmful content by 20 times on Twitter over a dataset of 4M tweets.

Technical Skills

Programming: Python, C, C++, JavaScript, Java, C#, Bash scripting, R

Machine Learning: Scikit-learn, PyTorch, TensorFlow, Pandas, Weights and Biases, OpenCV, SciPy, NumPy

Technologies: Elasticsearch, Lucene, AWS, Azure, GCP, Docker, Flask, FastAPI, Selenium, Spark, React.js, Django, Node.js, vLLM, Ollama, LangChain, Ray, Github Actions, Azure DevOps, Jenkins, MLflow

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, Recommender Systems, Search Systems, Human-Computer Interaction, 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 identifying bottlenecks on a NASA-JPL remote sensing application, optimizing a new parallelization scheme with 3x speedup.

Data-Driven Analysis of Targeted Cyberattacks and Effects on Foreign Policy. *May 2021 - July 2021*

- Crawled, aggregated, and normalized a dataset of state-sponsored cyber attacks from three heterogeneous data sources, analyzing the change in severity of future cyberattacks as an effect of policy actions between rival nations.

Geolocation-based Food Recommendation System. *March 2021 - May 2021*

- Benchmarked indexing and retrieval performance of KD-Tree and R-Tree based spatial indexes on 424K POIs, accompanied by an interactive companion web app that recommends nearby restaurants with multi-faceted search.

911 Overflow Management. *September 2019 - October 2019*

- Developed an IVR system to relieve 911 call surges during disasters, integrating Twilio for scalable VOIP routing and Dialogflow for NLP-based intent extraction, coupled with a web-app to aggregate related incident reports.

Voice Agent based Tutor for Mathematics Retention. *August 2017 - February 2018*

- Implemented a conversational agent using Amazon Alexa skills, assisting USM Psychology department in a study using conversational AI in retention of single-digit mathematical calculations for children with cognitive challenges.

Awards and Honors

Student Excellence Award, **BU Hariri Institute for Computing** *2025*

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*

Best IBM Watson Hack, **HackRice** *2016*

Selected Publications

[1] **P. Paudel**, G. Stringhini, "Loki: Proactively discovering online scams by mining toxic search queries," 35th Network and Distributed System Security Symposium (NDSS), 2026.

[2] A Galeazzi, **P. Paudel**, M. Conti, E De Cristofaro, G. Stringhini, "Revealing The Secret Power: How Algorithms Can Influence Content Visibility on Twitter/X." 35th Network and Distributed System Security Symposium (NDSS), 2026.

[3] 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.

[4] **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.

[5] **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.

[6] **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.

[7] 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.

[8] **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.