

Shivam Sharma

sharma.shivam2705@gmail.com | [Scholar](#) | [LinkedIn](#) | [Website](#)

EDUCATION

New Jersey Institute of Technology

M.Sc in Data Science

New Jersey, USA

Aug 2019 – May 2021

The LNM Institute of Information Technology

B.Tech in Computer and Communication Engineering

Rajasthan, India

Aug 2015 – May 2019

PUBLICATIONS

RAG vs. Long Context: Examining Frontier Large Language Models for Environmental Review Document Comprehension [Paper](#)

H. Phan, A. Acharya, S. Chaturvedi, **Shivam Sharma**, M. Parker, D. Nally, A. Jannesari, K. Pazdernik, M. Halappanavar, S. Munikoti, S. Horawalavithana

Arxiv, 2024

Foundation models of scientific knowledge for chemistry: Opportunities, challenges and lessons learned [Paper](#)

S. Horawalavithana, E. Ayton, **Shivam Sharma**, S. Howland, M. Subramanian, S. Vasquez, R. Cosbey, M. Glenski, S. Volkova

BigScience Workshop, 2022

Expert: Public benchmarks for dynamic heterogeneous academic graphs [Paper](#)

S. Horawalavithana, E. Ayton, A. Usenko, **Shivam Sharma**, J. Eshun, R. Cosbey, M. Glenski, S. Volkova

Arxiv, 2022

Bang for your Buck: Performance Impact Across Choices in Learning Architectures for Crisis Informatics [Paper](#)

Shivam Sharma, C. Buntain

ISCRAM, 2022

Evaluating and explaining natural language generation with genx [Paper](#)

K. Duskin, **Shivam Sharma**, J.Y. Yun, E. Saldanha, D. Arendt

DaSH, 2022

An Evaluation of Twitter Datasets from Non-Pandemic Crises Applied to Regional COVID-19 Contexts [Paper](#)

Shivam Sharma, C. Buntain

ISCRAM, 2021

Method Comparison for Crisis Pipelines [Paper](#)

Shivam Sharma, C. Buntain

TREC-IS, 2021

Improving Classification of Crisis-Related Social Media Content via Text Augmentation and Image Analysis [Paper](#)

Shivam Sharma, C. Buntain

TREC-IS, 2020

Combining neural, statistical and external features for fake news stance identification [Paper](#)

G. Bhatt, A. Sharma, **Shivam Sharma**, A. Nagpal, B. Raman, A. Mittal

MSM, 2018

RESEARCH EXPERIENCE

Data Scientist

January 2023 – Present

Pacific Northwest National Laboratory

Washington, USA

- Contributed on the research and development of domain-adapted GPT-style LLMs for specialized applications in scientific and environmental datasets, optimizing model fine-tuning and domain-specific knowledge integration.

Research Associate

October 2021 – December 2022

Pacific Northwest National Laboratory

Washington, USA

- Assisted in the application of NLP techniques, including data normalization, feature engineering, and tokenization for effective data curation and post-processing workflows.

Research Intern

September 2020 – May 2021

Pacific Northwest National Laboratory

Washington, USA

- Contributed to the development and testing of application for model evaluation and refinement of text generated by GPT- and BERT-style LLMs to enhance the detection and classification of phishing emails

Research Assistant

October 2019 – September 2020

New Jersey Institute of Technology

New Jersey, USA

- Worked on training and evaluation of BERT-style LLMs to analyze and classify social media posts during crisis situations.

Research Intern

June 2017 – July 2017

Indian Institute of Technology

Uttarakhand, India

- Implemented NLP techniques such as semantic similarity, topic modeling for detecting stances in fake news content.

PROFESSIONAL ACTIVITIES

Conference and Journal Reviewing

- Program Committee, ER 2024
- Review Committee, PERTRA Socio-NLP 2024
- Review Committee, IEEE TAI 2023
- Review Committee, Humanities & Social Sciences Communications Journal 2023
- Review Committee, EMNLP 2023

Invited Talks

- Lowercarbon AI Hackathon for Clean Energy, 2024

PROGRAMMING SKILLS

Languages: Python, SQL, R

Libraries: HuggingFace Transformers, HuggingFace Accelerate, HuggingFace Datasets, PyTorch

Software: Git, Docker