


Uma Gunturi







Master of Science in Computer Science | Virginia Tech

🌐 umagunturi.github.io/ @ uma.gunturi@ibm.com  github.com/UmaGunturi  Google Scholar
📍 IBM Silicon Valley Lab, 555 Bailey Ave, San Jose, CA 95141

Education

May 2023 Aug 2021	Virginia Tech Department of Computer Science Master of Science in Computer Science, GPA: 4.0/4.0 Thesis: <i>Linguistically Differentiating Acts and Recalls of Racial Microaggressions on Social Media</i> [PDF] Committee: <i>Dr. Eugenia H. Rho (Primary), Dr. Ismini Lourentzou, Dr. Edward Fox</i> ★ Received VT Paul E. Torgerson Research Excellence Award for best thesis	Blacksburg, VA
May 2021 Aug 2017	Bennett University Bachelor of Technology in Computer Science w/ Political Science (Minor), CGPA: 9.02/10.00	Greater Noida, India

Experience

Present July 2023	IBM Watsonx Orders Conversational AI powered drive-thrus  <i>Machine Learning Engineer</i> Projects: Building and deploying end-to-end spoken language understanding (ASR, NLU) systems optimized for noisy QSR (Quick Service Restaurant) drive-thru environments on resource-constrained edge devices, Extending real-time multilingual (Spanish) capabilities for our AI-driven dialog system "Archy" [Press]	San Jose, CA
Aug 2022 May 2022	IBM Research Foundations of AI – Knowledge Reasoning Team  <i>AI Intern Advisors: Dr. Tengfei Ma (Primary), Dr. Achille Fokoue, Dr. Alexander Gray</i> Project: Developed an interpretable neuro-symbolic model for time series classification, combining neural networks with signal temporal logics to uncover and describe data patterns in human-readable formulas	Yorktown Heights, NY
Present Aug 2021	VT Center for Human Computer Interaction (CHCI)  <i>Graduate Student Researcher at the Society + AI & Language (SAIL) Lab Advisor: Dr. Eugenia H. Rho</i> Projects: Examining semantic differences between acts and recalls of Racial Microaggressions on social media, Leveraging linguistics and NLP to help users counter gray areas of online toxicity, Using prompt-based LLMs to examine the relationship between social media language patterns and trends in national health outcomes	Blacksburg, VA
Dec 2021 Apr 2020	Indraprastha Institute of Information Technology (IIIT), Delhi MIDAS Lab  <i>Research Intern (Bachelor Thesis) Advisors: Dr. Debanjan Mahata (Primary), Dr. Rajiv Ratn Shah</i> Projects: Developed a framework for abstractive summarization of open-domain Hindi-English code-switched conversations, Understanding behaviors of users with suicidal ideation in COVID-19 mental health support communities	New Delhi, India
Aug 2019 May 2019	Georgia Tech Visual Intelligence Lab  <i>Visiting Research Student Advisors: Dr. Devi Parikh (Primary), Dr. Dhruv Batra</i> Project: Analyzed behavior of with-attention and without-attention Visual Question Answering (VQA) models, revealing insights to enhance model performance and guide future research directions	Atlanta, GA
Aug 2018 May 2018	Georgia Tech Chu Data Lab  <i>Visiting Research Student Advisors: Dr. Xu Chu</i> Project: Created a framework for composing and debugging machine learning model workflows using local interpretable model-agnostic explanations (LIME) for enhanced model transparency	Atlanta, GA

Publications

S=In Submission, W=Workshop, C=Conference

- [C.3] **Linguistically Differentiating Acts and Recalls of Racial Microaggressions on Social Media** [PDF], [Poster], [DEMO]
[Uma Gunturi](#)*, Anisha Kumar, Xiaohan Ding, Eugenia H. Rho
ACM Conference On Computer-Supported Cooperative Work And Social Computing [CSCW'24]
- [C.2] **Predicting Pandemic Health Decisions and Outcomes Through Social Media Language: A Fuzzy-Trace Theory Approach Leveraging Large Language Models** [PDF]
[Xiaohan Ding](#)*, [Buse Carik](#), [Uma Gunturi](#), Valerie Reyna, Eugenia H. Rho
ACM Conference on Human Factors in Computing Systems [CHI'24]

- [C.1] **GupShup: Summarizing Open-Domain Code-Switched Conversations** [PDF], [🔗]
 Laiba Mehnaz*, Debanjan Mahata*, Uma Gunturi*, Amardeep Kumar*, Rakesh Gosangi, Riya Jain, Gauri Gupta, Isabelle Lee, Anish Acharya, Rajiv Ratn Shah (* = Equal Contribution)
The 2021 Conference on Empirical Methods in Natural Language Processing [EMNLP'21]
- [W.1] **ToxVis: Enabling Interpretability of Implicit vs. Explicit Toxicity Detection Models with Interactive Visualization** [PDF], [DEMO]
Uma Gunturi*, Xiaohan Ding*, Eugenia H. Rho
The CHI'23 Workshop on Combating Toxicity, Harassment, and Abuse in Online Social Spaces [CHI'23 Workshop]
- [S.3] **CounterQuill: Investigating the Potential of Human-AI Collaboration in Online Counterspeech Writing** [Coming Soon]
 Xiaohan Ding*, Kaiké Ping, Uma Gunturi, Buse Carik, Sophia Stil, Lance T Wilhelm, Taufiq Daryanto, James Hawdon, Sang Won Lee, Eugenia H. Rho
Under Review at ACM International Conference on Supporting Group Work [GROUP'25]
- [S.2] **Understanding LLM Use Among Neurodivergent Users: Experiences, Challenges, and Workarounds.** [Coming Soon]
 Buse Carik, Xiaozheng Wang, Xiaohan Ding, Uma Gunturi, Eugenia H. Rho
Under Review at ACM Conference On Computer-Supported Cooperative Work And Social Computing [CSCW'24]
- [S.1] **Transformers at the Edge: Challenges and Recipes for Productizing Neural Machine Translation for Regular Languages** [Coming Soon]
Uma Gunturi*, Pooja Voladoddi*, Mostafa Varzaneh*, Tanmay Bakshi*
Under Review at ACL Rolling Review [ARR'24]

Select Research Projects

Intertextuality in linguistic interactions

Aug'21 - Present

Advisors: *Dr. Eugenia Rho*

- > Led foundational mixed-methods HCI research to characterize the language underlying acts and recalls of racial microaggressions in the context of racism in the U.S. providing broader implications to the current challenges in content moderation systems on social media. [Accepted@CSCW 24]
- > Introduced Role-Based Incremental Coaching (RBIC), a prompt-based LLM approach grounded in fuzzy-trace theory, to extract critical "gists" from COVID-19 related discussions and assessed their impact on online engagement and national health trends such as vaccine uptake and hospitalizations. [Accepted@CHI 24]
- > Visualizing semantic opportunities and limitations of transformer-generated counter responses against implicit online aggressions through an interactive UI playground for users to learn and craft responses through machine-generated texts. [🔗] [Accepted@CHI'23 Workshop]

Abstractive Summarization of Open-Domain Code-Switched Conversations

Apr'20 - Dec'21

Advisors: *Dr. Debanjan Mahata, Dr. Rajiv Ratn Shah*

- > Introduced the task of abstractive summarization for Hindi-English (Hi-En) code-switched conversations leading to developing the first code-switched conversation summarization dataset *GupShup*. [🔗] [Accepted@EMNLP'21]
- > Experimented with state-of-the-art abstractive summarization models (T5, Pegasus, BART etc) and conducted an extensive qualitative analysis of the results to provide insight into some of the model shortcomings.

Generating Low Resource Language Data from Code Switching

Sept'22 - Present

Advisor: *Dr. Teruko Mitamura, Dr. Zachary Jagers*

- > Proposed a novel, low-cost data generation framework to generate Low Resource Languages from code-switched (CS) corpora. [Paper]
- > Experimented with English-Telugu and English-Hindi to generate datasets for machine translation, sentiment analysis, and commonsense reasoning tasks.
- > This work is done in collaboration with the Language Technologies Institute at Carnegie Mellon University and is partially funded by the Air Force Research Laboratory under agreement number FA8750-19-2-0200.

Neuro-symbolic Models for Interpretable Time Series Classification

May'22 - Aug'22

Advisors: *Dr. Tengfei Ma, Dr. Alexander Gray, Dr. Achille Fokoue*

- > Developed a novel Neuro-Symbolic LSTM (NS-LSTM) that integrates Signal Temporal Logic (STL) descriptions and neural networks (NNs) to accomplish downstream tasks such as TSC using multi-view data representation leaving models human-readable and interpretable
- > We test NS-LSTM on real-world datasets obtained from the BETR DARPA project and other benchmark datasets from the UCR time-series repository, demonstrating that NSTSC achieves comparable performance compared to state-of-the-art models.

Honours and Awards

Paul E. Torgersen Research Excellence Award, 2023 | Virginia Tech One of the four recipients in the Master's program

Teaching and Leadership Roles

SAIL NLP Reading Group, VT *Organizer* Aug'21 - Present

- > Organize a weekly lab-wide Reading Group focused on research in the areas of Human Computer Interaction, Natural Language Processing. We read recent and classical papers as well as arrange for invited talks in related areas.

Software Design and Data Structures (CS 2114) *Graduate Teaching Assistant* Aug'21 - May'23

- > Responsibilities included evaluating labs, and helping students with the coursework and home/lab assignments.

Skills

- > **Languages:** Java, Python, C++, HTML
- > **Frameworks:** Pandas, NumPy, Gensim, Matplotlib, OpenCV, Keras, Tensorflow, PyTorch, TextBlob, NLTK, HuggingFace
- > **Tools:** Git, GCP, Eclipse, Android Studio, Flask, FireBase, Visual Studio, Stanford Parser, Elasticsearch
- > **Relevant Coursework:** Machine Learning, Information Retrieval, Data Analytics, Human AI Interaction, Natural Language Processing, Deep Learning, Discrete Mathematics, Probability and Statistics

Academic Service

Reviewer CSCW '22, CHI'22

Volunteer ARR'24, ACL'22