Omar Sharif

PH.D. STUDENT | DEPARTMENT OF COMPUTER SCIENCE | DARTMOUTH COLLEGE

✉ omar.sharif.gr@dartmouth.edu | ⌂ omar-sharif03.github.io | ⌂ omar-sharif03 | in omar1303

🔍 Research Interest

- Natural Language Processing, Human-Computer Interaction, Deep Learning
    - Large Language Models Inference/Training and Representation Learning
    - Multi-granularity Information Extraction
    - Multimodal NLP (analyze text, image, speech and other meta signals to build intelligent systems)
    - NLP for Social Good (application focusing on Social Science and Healthcare)

🏛 Education

Sep 2022 - Dec 2026 (expected)	**Ph.D. in Computer Science** Dartmouth College
Aug 2019 - Jan 2022	**Masters in Computer Science & Engineering** Chittagong University of Engineering and Technology (CUET) [4.0/4.0]
Mar 2014- Nov 2018	**Bachelors in Computer Science & Engineering** Chittagong University of Engineering and Technology (CUET) [Summa Cum Laude]

💼 Appointments

Sep 2022 - Ongoing	**Graduate Research/Teaching Assistant**, Department of Computer Science, Dartmouth College
Jul 2022 - Aug 2022	**Assistant Professor**, Department of Computer Science & Engineering, CUET
Jul 2019 - Jun 2022	**Lecturer**, Department of Computer Science & Engineering, CUET, Bangladesh
Dec 2018 - Jun 2019	**Research Assistant**, CUET NLP Lab, Bangladesh

📑 Publications

To get the full list of publications please check: G [Google Scholar]　🎓 [Semantic Scholar]　® [ResearchGate]

Conferences

(*) indicates equal contribution.

C8. **Omar Sharif**, Joseph Gatto, Madhusudan Basak, Sarah Preum, "Explicit, Implicit, and Scattered: Revisiting Event Extraction to Capture Complex Arguments", (EMNLP-2024) [Paper]

C7. **Omar Sharif**, Madhusudan Basak, Tanzia Parvin et al., "Characterizing Information Seeking Events in Health-Related Social Discourse", (AAAI-2024) [Paper]

C6. Joseph Gatto, **Omar Sharif**, Sarah Preum, "Chain-of-Thought Embeddings for Stance Detection on Social Media", (EMNLP-2023 Findings) [Paper]

C5. Eftekhar Hossain, **Omar Sharif**, Mohammed Moshiul Hoque, Sarah Preum, "Deciphering Hate: Identifying Hateful Memes and Their Targets", (ACL-2024) [Paper]

C4. William Romano*, **Omar Sharif***, Madhusudan Basak, Joseph Gatto, Sarah Preum, "Theme-driven Keyphrase Extraction to Analyze Social Media Discourse", (ICWSM-2024) [Paper]

C3. Eftekhar Hossain*, **Omar Sharif***, Mohammed Moshiul Hoque, Sarah Preum, "Align before Attend: Aligning Visual and Textual Features for Multimodal Hateful Content Detection", (EACL-SRW 2024) [Paper]

C2. **Omar Sharif**, Eftekhar Hossain Mohammed Moshiul Hoque, "M-BAD: A Multilabel Dataset for Detecting Aggressive Texts and Their Targets", (CONSTRAINT@ACL-2022]) [Paper]

C1. **Omar Sharif**, Mohammed Moshiul Hoque, "Identification and Classification of Textual Aggression in Social Media: Resource Creation and Evaluation", (CONSTRAINT@AAAI-2021 [acceptance rate: 37.1%]) [Paper][Best paper award]

Journals

J2. **Omar Sharif**, Mohammed Moshiul Hoque, "Tackling Cyber-Aggression: Identification and Fine-Grained Categorization of Aggressive Texts on Social Media using Weighted Ensemble of Transformers", (Neurocomputing [IF: 6.00]) [Paper]

J1. **Omar Sharif**, Mohammed Moshiul Hoque, A. S. M. Kayes, Raza Nowrozy and Iqbal H. Sarker, "Detecting Suspicious Texts using Machine Learning Techniques", (Journal of Applied Sciences [IF: 2.83]) [Paper]

Workshops

W3. Joseph Gatto, **Omar Sharif**, Parker Seegmiller, Philip Bohlman, Sarah Preum, "Text Encoders Lack Knowledge: Leveraging Generative LLMs for Domain-Specific Semantic Textual Similarity", (GEM@EMNLP-2023) [Paper]

W2. **Omar Sharif**, Eftekhar Hossain, Mohammed Moshiul Hoque, "Offensive Language Detection from Multilingual Code-Mixed Text using Transformers", (DravidianLangTech@EACL-2021) [Paper]

W1. Eftekhar Hossain, **Omar Sharif**, Mohammed Moshiul Hoque, "Multilingual Code-Mixed Hope Speech Detection using Cross-lingual Representation Learner", (LTEDI@EACL-2021) [Paper][Top model in the challenge]

✎ Teaching

Role	Affiliation	Course	Class size	Syllabus
Teaching assistant	Dartmouth	Machine Learning and Statistical Data Analysis	150	📄
Teaching assistant	Dartmouth	Theory of Computation	20	📄
Teaching assistant	Dartmouth	Algorithms	55	📄
Co-instructor	CUET	Object-oriented programming (C++)	132	📄
Co-instructor	CUET	Structured programming (C)	130	📄

🏆 Honors and Awards

- **Presidential Graduate Fellow** (2022): Awarded to only two incoming graduate student in Dartmouth.
- **ICT division Bangladesh research grant** (2020-2021): 1000K BDT
- **CUET Ex-students merit award** (2015, 2016, 2017): For obtaining highest GPA in an academic year
- **Best paper award**: CONSTRAINT collocated with AAAI-2021
- **Volunteer or D&I Award** (ICWSM-24, EACL-[24, 21], EMNLP-[23, 21], ACL-22): To attend and present our papers

⬡ Technical/Programming Skills

Proficient C, C++, Python, PyTorch

Libraries Huggingface, Keras, TensorFlow, NLTK, Scikit-learn, Gensim, Pandas

Solved 1000+ competitive programming problems in different online platforms as part of my preparation for programming contests (ICPC, NCPC, etc). ⬀ Solved problems (GitHub)

Codeforces	Solved 500+ problems (omar_sharif)	**UVA**	Solved 200+ problems (cuet_fighter)
LightOJ	Solved 120+ problems (omar-sharif)	**CS Academy**	Solved 50 problems (omar_sharif)

☕ Misc

Reviewer
- > **Conferences:** EMNLP-[23, 24], ACL[23, 24], ICWSM-[23, 24], LREC-[22, 24], NAACL-24.
- > **Journals:** TALLIP, IEEE Access, Scientific Reports, TCSS, LREV