Ankita Gupta

ankitagupta@umass.edu
ankitaiisc.github.io

RESEARCH INTERESTS

Current research interests in text comprehension and information extraction (who did what to whom), language understanding, analyzing social harms from current NLP models, computational social science and misinformation, learning from limited data.

I am presently working on

1 Event factuality models for political discourse analysis.

2 Crowd-sourcing multilingual and multi-domain annotations for coreference resolution.

EDUCATION

University of Massachusetts Amherst
Ph.D., Computer Science   2021 – Present

Indian Institute of Science Bangalore
M.E. with Thesis, Electrical Engineering and Computer Science   2017

Malaviya National Institute of Technology Jaipur
B.Tech., Electrical Engineering   2014

RESEARCH EXPERIENCE

UMass Amherst   2021-Present
Position: Graduate Research Assistant
Advisors: Brendan O'Connor, Mohit Iyyer

- Multi-source **epistemic stance (event factuality)** modelling.
  Fine-tuned BERT modelling with intermediate fine-tuning using a dataset drawn from books with diverse U.S. political ideologies. Corpus analysis to identify significant belief holders mentioned in the text. Comparison with traditional named entity recognition.

- Crowdsourcing for **coreference resolution**.
  Researching a human-friendly paradigm to collect crowdsourced coreference annotations for multiple domains and languages with reduced annotation costs.

Amazon Bangalore   2020-2021
Position: Applied Scientist

- Ranking deals & discounts on e-commerce platform.
  Thompson style sampling from a predictive posterior distribution estimated via Bayesian linear regression.

- Automatic curation of theme-specific products.
  Metric-learning-based **meta-learning** approach with Kullback–Leibler divergence loss.

Samsung Bangalore

Position: Lead Engineer

- **Fact Verification** 2018-2020
  Document retrieval using elastic search and search tree (trie).
  Sentence similarity using BERT.
  Inference using multi-task modelling with adversarial training.

- **Content Quality**
  Machine learning models for identification of hate speech, hyper-partisanship, logical-fallacy (appeal to anonymous authority).

- **Echo Chambers** (biases people to read only one side of a story)
  Stance detection using a multi-task model to distinguish favourable and opposing opinions about a debatable issue.

- **Miscellaneous** 2017-2018
  Question answering system (BiDAF, QANet and RNet) to extract snippets of text which are relevant to a claim in fact-checking pipeline.
  Neural question generation to convert a claim into question that can be used as a search query to enhance the coverage and relevance.
  Claim-extraction using **abstractive summarization** and sentence ranking techniques.

PUBLICATIONS

Under Review

PoliBelief: A Multi-source Epistemic Stance Dataset for Analyzing Political Ideology.
Ankita Gupta, Su Lin Blodgett, Justin H. Gross, and Brendan O'Connor.

Journal Publications

An Online Power System Stability Monitoring System using Convolutional Neural Networks.
Ankita Gupta, Gurunath Gurrala, and P. S. Sastry.
IEEE Transactions on Power Systems. 2018.

Conference and Workshop Publications

Instability Prediction in Power Systems using Recurrent Neural Networks.
Ankita Gupta, Gurunath Gurrala, Pidaparthy S Sastry.
In *Proceedings of the Twenty-Sixth IJCAI.* 2017.

Question Factuality and Answer Veracity Prediction in Community Forums.
Ankita Gupta, S Sahoo, D Prakash, R.R Rohit, V Srivastava, and Y H Kim.
In *Proceedings of the 13th International Workshop on Semantic Evaluation.* 2019.

Hyperpartisan News Detection using Lexical and Semantic Features.
V Srivastava, **Ankita Gupta**, D. Prakash, S Sahoo, R.R Rohit, and Y H Kim.
In *Proceedings of the 13th International Workshop on Semantic Evaluation.* 2019.

Knowledge Directed Multi-task Framework for Natural Language Inference in Clinical Domain
S Chopra, **Ankita Gupta**, and A Kaushik.
In *Proceedings of the 18th BioNLP Workshop and Shared Task.* 2019.

HONORS AND AWARDS

UMass W. Bruce Croft Scholarship (one student in NLP)	2021
UMass Anuradha and Hanuma Kodavalla Scholarship	2021
Samsung Citizen Award	2019
Gold Medallist (B.Tech)	2014
KVPY Fellowship (Young scientist fellowship, Government of India)	2010

SKILLS

Python, C++ (Certified Professional, Samsung), C and Matlab, TensorFlow, PyTorch, scikit-learn, NLTK, spaCy, Stanford CoreNLP, NumPy, SciPy, Pandas, Amazon SageMaker, AWS Services and Amazon Mechanical Turk.

RELEVANT GRADUATE COURSEWORK

Advanced Natural Language Processing, Probabilistic Graphical Models, Machine Learning for Signal Processing, Data Analytics, Game Theory, Pattern Recognition and Neural Networks, Data Mining, Linear and Non-Linear Optimization.

TALKS

A Multi-Source Epistemic Stance Dataset for Analyzing Political Ideology. 11th Annual Conference on Analyzing Text as Data (TADA).	2021
Optimization, Machine Learning. Dayanand Sagar College of Engineering.	2017, 2018

SERVICE AND OUTREACH

Ph.D. Applicant Support Program mentor, UMass CICS.	2021
Volunteer, Candidate Friday, UMass CICS.	2021
Reading Group Coordinator, Samsung.	2018-2019
Young Women Professional Representative, Budget Meeting, Chief Minister Secretariat.	2014
Student Representative, Departmental Under Graduate Committee, MNIT, Jaipur.	2013-2014