Ankita Gupta

NLP and Machine Learning Enthusiast | Ph.D. Student in NLP and Computational Social Science

+1 (612)-702-9134 | ankitagupta@umass.edu | /in/ankita-gupta-4b4571104 | /ankitaiisc | Google Scholar

- Researcher in the field of Machine Learning and Natural Language Processing.
- Presently working on
  - Event factuality models for political discourse analysis.
  - Crowd-sourcing multilingual and multi-domain annotations for Coreference Resolution.
- Current research interests in computational social science, language understanding & learning from limited data.

Education

2021-Present **Ph.D.** *University of Massachusetts Amherst*
Statistical Social Language Analysis Lab, College of Information & Computer Sciences
Advised by Prof. Brendan O'Connor and Prof. Mohit Iyyer
CGPA : 4.0/4

2017-2015 **Master of Engineering (Thesis)** *Indian Institute of Science, Bangalore*
CGPA : 7.0/8

2014-2010 **Bachelor of Technology** *Malaviya National Institute of Technology, Jaipur*
CGPA : 9.78/10

2010, 2008 **Schooling** *India International School, Jaipur*
High School : 93.4 %, Intermediate : 95.0 %

Publications

2021 **PoliBelief: A Multi-Source Epistemic Stance Dataset for Analyzing Political Ideology**
Transactions of the Association for Computational Linguistics, 2021. (under-review)

2020 **Ensemble Architecture for Fine-Tuned Propaganda Detection in News Articles**
SemEval, COLING, 2020. PDF

2019 **Knowledge Directed Multi-task Framework for Natural Language Inference in Clinical Domain**
BioNLP, ACL, 2019. PDF

2019 **Hyperpartisan News Detection using Lexical and Semantic Features**
SemEval, NAACL HLT, 2019. PDF

2019 **Question Factuality and Answer Veracity Prediction in Community Forums**
SemEval, NAACL HLT, 2019. PDF

2018 **An Online Power System Stability Monitoring System using Convolutional Neural Networks**
IEEE Transactions on Power Systems. PDF

2017 **Instability Prediction in Power Systems using Recurrent Neural Networks**
International Joint Conference on Artificial Intelligence (IJCAI). PDF, Slides

2015 **Optimal provision for enhanced consumer satisfaction and energy savings by an intelligent household energy management system**
IEEE International Conference on Power Systems (ICPS) PDF

Talks

2021 **Presented at Text as Data (TADA) 2021**
Presented my research on "PoliBelief: A Multi-Source Epistemic Stance Dataset for Analyzing Political Ideology."

2019 **Invited Speaker on AI for Social Good**
Applications in the field of social welfare such as flood levels prediction, early detection of skin cancer.

2018 **Invited Speaker on Optimization and its applications in Machine Learning**
Slides

2017 **Invited Speaker on Machine Learning with hands-on in Python**
Slides
Dayanand Sagar College of Engineering, Bangalore

Achievements and Honors

2021 **Graduate Scholarship**
Awarded *Anuradha and Hanuma Kodavalla* Graduate Scholarships in Computer Science

2021 **Graduate Scholarship**
Awarded *W. Bruce Croft* Graduate Scholarships in Computer Science

2019 **Samsung Citizen Award**
Awarded by CTO for extraordinary commitment and achievements beyond functional scope.

2014 **All India Rank 08**
Graduate Aptitude Test in Engineering

2014 **Gold Medallist**
B.Tech in Electrical Engineering

2012 **Scholarship**
Shortlisted for O.P. Jindal Engineering and Management Scholarship

2010 **Merit Award in Indian National Chemistry Olympiad**
Homi Bhabha Centre for Science Education and Indian Association of Chemistry Teachers

2010 **Certificate of Merit**
International Mathematics Olympiad and National Science Olympiad

2009 **KVPY Fellowship**
Young scientist fellowship awarded by department of science and education, Government of India

2008 **Certificate of Merit**
Top 0.1% academic performance in Social Science, Central Board of Secondary Education

Skills

- Programming : C, Matlab and Python, C++ (Certified Professional, Samsung)
- Neural Networks : TensorFlow, Keras, PyTorch
- Machine Learning and NLP : scikit-learn, NLTK, spaCy, Stanford CoreNLP
- Data Manipulation and Visualization : NumPy, SciPy, Pandas, SQL
- Data Streaming and Storing : Apache Kafka, Redis
- AWS Services : Amazon SageMaker, S3, Athena, DynamoDB, CloudWatch

Course Work

Artificial Intelligence	Probabilistic Graphical Models, Advanced Natural Language Processing, Machine Learning for Signal Processing, Data Analytics, Game Theory, Pattern Recognition and Neural Networks, Data Mining, Dynamics of Linear Systems
Mathematics	Stochastic Models and Applications (Probability), Linear Algebra, Linear and Non-Linear Optimization, Mathematics-I (Differential, Integral and Vector Calculus), Mathematics-II (ODEs and PDEs), Mathematics-III (Laplace, Fourier, Z Transforms)
Computer Science	Quantum Information Systems, Data Structures and Algorithms, Computer Architecture and Organization, Computer Systems and Programming, Microprocessors, Switching Theory and Logic Design

Experience

Amazon

Applied Scientist (India Machine Learning) **2020-2021**

- Ranking deals & discounts on e-commerce platform.
  –Used Bayesian linear regression for estimating exact predictive posterior distribution of regression weights. This distribution is used to rank the deals similar to **Thompson sampling** (multi-arm bandits).
- Automatic curation of a store with products relevant to specific celebrations (e.g., thanksgiving). –Used metric-learning-based **meta-learning** approach with KL divergence as loss function to learn effective mapping of products.
- Inducing product taxonomy based on user search queries on an e-commerce platform. –Used **policy gradient** to train an agent which observes a user query and decides which internal node this query must be placed on such that the reward is maximized. Reward is governed by purchase behaviour and common-sense knowledge incorporated via ConceptNet.

Samsung Research Institute Bangalore

Lead Research Engineer **2019-2020**

- Worked on **fact verification** problem involving document retrieval using elastic search, sentence level semantic similarity using BERT and natural language inference using multi-task model with adversarial training.
- Designed machine learning models that evaluate **content quality** on parameters: hate speech, hyper-partisanship, exaggeration and sensationalism.
- Identification of **logically fallacious arguments** in a piece of text. Tackled specific types of fallacies such as ad-hominem, appeal to emotions and appeal to anonymous authority.
- Modelled the problem of **echo chambers** (which biases people to read only one side of a story) as a stance detection system, a multi-task framework to categorize opinions about a debatable issue in its favour/ against.
- Designed data collection and curation strategies for machine learning/deep learning based models.

Senior Software Engineer (Machine Learning) **2017-2019**

- Applied **machine reading comprehension** (BiDAF, QANet, RNet) to extract relevant parts of text with respect to a claim in fact-checking pipeline. Enhanced performance over existing benchmarks by incorporating constituency parsing and ELMo based trainable embeddings.
- Worked on **neural question generation** to convert a claim into question that can be used as a search query to enhance the coverage and relevance. Demonstrated performance of developed prototype on real life dataset.
- Worked on claim-extraction sub-module using **abstractive summarization and sentence ranking techniques**. Utilized **WordNet based text summarization** technique for fake-review summarization.

Master's Dissertation

Instability Prediction in Power Systems using Deep Networks **2017**

Prof. P.S. Sastry | Dr. Gurunath Gurrala

- Addressed the problem of **early prediction** of instability following a fault in an interconnected power system.
- Proposed heat-map visualization of **time series data** for instability manifestation. Used these identified patterns for image classification using **convolution neural networks**.
- Proposed an **outlier detection** based method to detect critical generators responsible for instability. Outlier is detected by fitting Gaussian density onto projected data in 2D.
- Also Proposed **multi-task framework** for instability detection and identification of critical generators. Common feature representation learnt by CNN makes predictions for both tasks.
- Assessed **robustness of the system to variability** such as noisy measurements, parameter changes, network topology changes.

Responsibilities and Positions

2021-2019 **Volunteer for Candidate Friday**
UMass Amherst CICS
Volunteered to share academic experiences as a graduate student at UMass with the new incoming graduate students.

2018-2019 **Reading Group Coordinator**
Advanced Technology Lab, Samsung Research and Development Institute, Bangalore
Conducted sessions on technical paper discussion every week to promote knowledge sharing.

2017-2018 **Campus Ambassador**
Samsung Research and Development Institute, Bangalore
Sharing work experience and opportunities available at Samsung with IISc student community.

2016-2017 **Student Placement Coordinator**
Indian Institute of Science, Bangalore
Connect with recruitment companies and share research being conducted by students at IISc. Responsible for execution of placement drive on campus.

2014-2015 **Young women professional representative**
Budget Meeting, Chief Minister Secretariat, Rajasthan, India
Invited to suggest ideas for urban development. Suggested underground electricity distribution system in Jaipur City for relieving urban congestion.

2013-2014 **Student representative of Departmental Under Graduate Committee**
Electrical Engineering Department, Malaviya National Institute of Technology, Jaipur
Assist faculty members in advising, counseling students in academic matters. Proposing new courses and programmes based on popular student demand.