AMANUL HAQUE

https://ahaque2.github.io/ | ahaque2@ncsu.edu |+1(919)-946-6094

EDUCATION

Doctor of Philosophy (PhD) in Computer Science — *GPA - **4.0/4*** — *Expected Dec 2023*
North Carolina State University, Raleigh, NC
With a focus on: Artificial Intelligence, Machine Learning, and Natural Language Processing

Master of Science in Computer Science — *CGPA - **4.0/4*** — *May 2019*
North Carolina State University, Raleigh, NC

Bachelor of Engineering in Information Science — *CGPA - **8.76/10*** — *May 2015*
PES University, Bangalore, India

RESEARCH AND PROJECTS

- ***Identifying Preferred Entity in User Generated Text*** (Github) ***Python, PyTorch***
  Created Pixie, a gold standard dataset for Preference Classification to identify the preferred entity in user-generated comparative sentences from app reviews. Our dataset includes implicit and indirect comparisons that have been overlooked by previous works. We achieved an F1-score of 84% by fine-tuning BERT with segment embeddings to demarcate compared entities, while SOTA achieved 74% F1 score on Pixie *(Published at ACL 2022)*
- ***Emotional response to COVID-19 News Coverage*** (Github) ***Python, PyTorch, Gensim***
  Created a large-scale dataset of Covid-19 related news tweets and user responses containing ~20M tweets from 12 countries across five continents. Fine-tuned the RoBERTa model to recognize emotions in tweets and used LDA to identify discussion topics. We found that anger was the predominant emotion in user responses, while sadness in news tweets. User emotions differ across population groups and discussion topics *(Published at IEEE Access 2022).*
- ***Understanding Dynamics of Polarization via Multiagent Social Simulation*** (Github) ***Python, mesa***
  Designed a social simulation to study polarization dynamics in social networks. Adopted Social Judgment Theory and modeled user behavior based on empirical evidence from past studies. Analyzed the influence of user tolerance and selective exposure on user satisfaction and polarization via content sharing *(under review in AI & Society)*
- ***Political Bias in Large Pertained Language Models*** (Github) ***Python, PyTorch***
  Identified political bias in pertained BERT model. We demonstrate how training data impacts political bias and its impact on downstream tasks. We use a prompt-based validation and find that language models are sensitive to data used for training and fine-tuning. We propose a new political bias detection dataset and metric.

PROFESSIONAL EXPERIENCE

Coupang, Mountain View, California, Machine Learning Summer Intern ***May 2022 - Aug 2022***
- Conducted an ablation study of Coupang's Deep Cross Network (DCN) and compared it with Google's.
- Benchmarked the DCN model and created scripts for easy-to-run experiments on public datasets for comparison.
- Improved DCN model performance via feature engineering and parameter tuning and reduced features being used.

Seagate, Longmont, Colorado, Machine Learning Summer Intern ***May 2020 - Aug 2020***
- Designed a graph-based unsupervised abstractive multi-document text summarizer for a social listening tool to identify trending online topics and summarize related documents.
- Implemented an unsupervised aspect-based sentiment analyzer for online user reviews.

Lenovo, Morrisville, NC, Computer Science Summer Intern ***May 2018 - Aug 2018***
- Designed modules to automate test plan generation using Requirement Management Kanban (RMK) and historical test results to reduce test suite execution time.
- Designed an information extraction model to identify executable commands from unstructured text in RMK.

Oracle, Bangalore, India, Software Developer ***July 2015 - June 2017***
- Collaborated with a global team to design and develop modules for Service Deployment Infrastructure (SDI) which governs the provisioning flow for all Oracle Public Cloud (OPC) subscription life-cycle.
- Implemented modules for a data center-level load balancer and a loosely coupled execution mode to improve the runtime and increase parallelism in execution.

SKILLS

- *Programming Languages*: Python, Java, Shell, SQL, PHP, HTML, JavaScript
- *API and Libraries:* PyTorch, Tensorflow, Keras, Scikit Learn, Gensim, Numpy, Pandas, REST