Laiba Mehnaz

+1 929 584 4743| lm4428@nyu.edu | Website | linkedin | github

Education

New York University, Tandon School of Engineering May 2023
Master of Science in Computer Science
Relevant coursework: Design and Analysis of Algorithms, Information visualization, Computer Networking

Delhi Technological University, New Delhi, India May 2020
Bachelor of Technology in Software Engineering
Relevant coursework: Database Management Systems, Object Oriented Programming, Software Engineering

Work Experience

Research Assistant June 2020 – June 2021
MIDAS lab, Indraprastha Institute of Information Technology, Delhi *New Delhi, India*

- Developed a cross-lingual Hindi-English code-switched conversation summarization dataset and examined the results and limitations of pretrained language models on summarizing conversations. (published)
- Evaluated the domain robustness of pretrained language models and their distilled versions, layer-wise, using domain divergence and zero-shot probing. (published)
- Analysed the effect of domain-adaptive pretraining for medical data using a dataset of tweets mentioning medications. (published)

Research Intern June 2019 – Aug 2019
MIDAS lab, Indraprastha Institute of Information Technology, Delhi *New Delhi, India*

- Used transfer learning to combat class imbalance for classification and extraction of adverse drug reactions from tweets. (published)
- Secured 5th rank out of 103 teams in the Shared task on identification of offensive posts with targeted offense from Twitter, held at the SemEval workshop at NAACL-HLT 2019. (published)
- Secured 10th rank out of 34 teams in the Shared task on suggestion mining from online reviews held at the SemEval workshop at NAACL-HLT 2019. (published)

Research Intern June 2018 – Aug 2018
MIDAS lab, Indraprastha Institute of Information Technology, Delhi *New Delhi, India*

- Curated a dataset consisting of tweets that contain blood donation requests and showed the potential of the dataset to build systems for identification of emergency blood donation requests on Twitter, using feature extraction and machine learning. (published)
- Built a framework using Spacy for automatic extraction and ranking of keyphrases from articles using a supervised technique (Key2Vec), that leverages phrase embeddings for ranking keyphrases in ranking algorithms such as TextRank, SGRank, SingleRank, and TopicRank.

Technical Skills

Prrogramming Languages: Python, Java, C, C++, JavaScript, SQL
Frameworks: PyTorch, TensorFlow, Keras, SpaCy, SciKit-learn, NLTK, Gensim, Git

Projects

Unsupervised Domain Adaptation for Sentiment Analysis using BERT [**poster**] [**code**]

- Reduced the divergence between the representations of the target domain and source domain for sentiment analysis in BERT, to produce domain invariant representations with higher accuracy on the target domain, at each layer, using Domain Adversarial training.
- Awarded the first place in the The 17th SoC Term Project Showcase, at NUS. [**certificate**]

Leadership Positions

TedxDTU, Public Relations and Media Executive Jan 2018-Jan 2020

- Led the attendee management and hospitality team to provide a comfortable experience to the attendees.
- Planned and conducted several publicity campaigns in and around the campus.

Professional Services

Reviewer for SMM4H 2020 Workshop co-located with COLING 2020.

Publications

Laiba Mehnaz, Debanjan Mahata, Rakesh Gosangi, Uma Sushmitha Gunturi, Riya Jain, Gauri Gupta, Amardeep Kumar, Isabelle Lee, Anish Acharya, Rajiv Ratn Shah. (2021). GupShup: An Annotated Corpus for Abstractive Summarization of Open-Domain Code-Switched Conversations. *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).* [**paper**] [**code**]

Abhinav Ramesh Kashyap, **Laiba Mehnaz**, Bhavitvya Malik, Abdul Waheed, Devamanyu Hazarika, Min-Yen Kan, Rajiv Ratn Shah. (2021). Analyzing the Domain Robustness of Pretrained Language Models, Layer by Layer. *In Proceedings of the Second Workshop on Domain Adaptation for NLP. The 16th Conference of the European Chapter of the Association for Computational Linguistics.* [**paper**]

Laiba Mehnaz. (2020). Automatic classification of tweets mentioning a medication using pre-trained sentence encoders. *In Proceedings of the Workshop on Social Media Mining for Health Applications, Barcelona, Spain, December 2020. The 28th International Conference on Computational Linguistics.* [**paper**] [**poster**]

Laiba Mehnaz and Rajni Jindal. (2020). Using Transfer Learning for detecting drug mentions in tweets. *Proceedings of Fifth International Conference on ICT for Sustainable Development. Goa, India.* [**paper**]

Sarthak Anand, Debanjan Mahata, Haimin Zhang, Simra Shahid, **Laiba Mehnaz**, Yaman Kumar, Rajiv Ratn Shah. (2019). MIDAS@SMM4H-2019: Identifying Adverse Drug Reactions and Personal Health Experience Mentions from Twitter. *In Proceedings of the Workshop on Social Media Mining for Health Applications, Florence, Italy, July 2019. Association for Computational Linguistics.* [**paper**] [**poster**] [**code**]

Haimin Zhang, Debanjan Mahata, Simra Shahid, **Laiba Mehnaz**, Sarthak Anand, Yaman Kumar, Rajiv Ratn Shah, Karan Uppal. (2019). MIDAS at SemEval-2019 Task 6: Identifying Offensive Posts and Targeted Offense from Twitter. *In Proceedings of the International Workshop of Semantic Evaluation collocated with NAACL-HLT 2019, Minneapolis, USA.* [**paper**]

Sarthak Anand, Debanjan Mahata, Kartik Aggarwal, **Laiba Mehnaz**, Simra Shahid, Haimin Zhang, Yaman Kumar, Rajiv Ratn Shah, Karan Uppal. (2019). MIDAS at SemEval-2019 Task 9: Suggestion Mining from Online Reviews using ULMFiT. *In Proceedings of the International Workshop of Semantic Evaluation collocated with NAACL-HLT 2019, Minneapolis, USA.* [**paper**]

Puneet Mathur, Meghna Ayyar, Sahil Chopra, Simra Shahid, **Laiba Mehnaz**, Rajiv Ratn Shah. (2019). Identification of Emergency Blood Donation Request on Twitter. *In Proceedings of the 3rd Workshop on Social Media Mining for Health Applications held in conjunction with Empirical Methods in Natural Language Processing, EMNLP 2018.* [**paper**]