

Laiba Mehnaz

+91 9971737074 | laibamehnaz@gmail.com | [Website](#) | [linkedin](#) | [github](#)

EDUCATION

Delhi Technological University, New Delhi, India

Aug. 2016 – May 2020

B.Tech in Software Engineering

CGPA: 7.88/10.0

Thesis on “Using Transfer Learning for Drug Detection from Tweets” [[pdf](#)]

Indian School Muscat, Oman

2016

All India Senior School Certificate Examination (Class XII), CBSE

96% (3rd rank in the school)

Indian School Muscat, Oman

2014

All India Senior School Certificate Examination (Class X), CBSE

GPA: 9.8/10

WORK EXPERIENCE

Research Assistant

June 2020 – Present

MIDAS-IIITD, Indraprastha Institute of Information Technology, Delhi

New Delhi, India

- Worked on cross-lingual conversation summarization of code-switched Hindi-English conversations to English summaries. Wrote a Python framework to calculate metrics for quantifying the complexity of code-switching in a code-switched Hindi-English dataset. Also, analysed the challenges due to code-switching in the current state-of-the-art language models. (preprint)
- Worked on analysing domain robustness of pretrained language models and their distilled versions, layer by layer, using domain divergence and zero-shot probing. (published)
- Worked on automatic classification of tweets mentioning a drug for the Shared Task of SMM4H at COLING 2020. Our system ranked 7th out of 26 teams. (published)

Research Intern

June 2019 – Aug 2019

MIDAS-IIITD, Indraprastha Institute of Information Technology, Delhi

New Delhi, India

- Worked on identification of adverse drug reactions from tweets, as well as identifying the spans mentioning those adverse drug reactions. Published in the Shared task of SMM4H at ACL 2019. [[code](#)]
- Worked on identification of offensive posts with targeted offense from Twitter. Published in the Shared Task of SemEval at NAACL-HLT 2019. Our system ranked 5th out of 103 teams.
- Worked on suggestion mining from online reviews. Published in the Shared Task of SemEval at NAACL-HLT 2019. Our system ranked 10th out of 34 teams.

Research Intern

June 2018 – Aug 2018

MIDAS-IIITD, Indraprastha Institute of Information Technology, Delhi

New Delhi, India

- Collected and curated a dataset consisting of tweets that contained blood donation requests. Using machine learning, worked on showing the potential of the dataset to build systems for identification of emergency blood donation requests on Twitter. (published)
- Using Spacy built a framework for automatic extraction and ranking of keyphrases from articles using a supervised technique (Key2Vec) that leverages phrase embeddings for ranking keyphrases. Implemented this in ranking algorithms such as TextRank, SGRank, SingleRank, and TopicRank.

PROJECTS

Unsupervised Domain Adaptation for Sentiment Analysis using BERT

[[poster](#)] [[code](#)]

- Using Domain Adversarial training, reduced the divergence between representations of both the target domain and source domain for the task of sentiment analysis in BERT, producing domain invariant representations with higher accuracy on the target domain, at each layer.
- Our project won the first place in the The 17th SoC Term Project Showcase, at NUS. [[certificate](#)]

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. *ArXiv*, *abs/2104.08578*. [[paper](#)]

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](#)]

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](#)]

PROFESSIONAL SERVICES

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

TECHNICAL SKILLS

Languages: Python, C/C++, Java

Frameworks: TensorFlow, Keras, PyTorch, SpaCy, Sci-Kit, NLTK, Gensim, Git

Natural Languages: Fluent in English and Hindi, beginner in Arabic

LEADERSHIP POSITIONS

TedxDTU, Public Relations and Media Executive

Jan'18-Jan 2020

- Planned and conducted several publicity campaigns in and around the campus
- Led a team for attendee management as well as the hospitality for the speakers arriving for the talk