

Ayush Kaushal

Indian Institute of Technology Kharagpur

✉ ayushk4@utexas.edu

🌐 ayushk4.github.io

☎ +1 (737) 2885781

Education

Masters of Science in Computer Science

2021-2023 (Expected)

The University of Texas at Austin

B.Tech. in Computer Science

2017-2021

Indian Institute of Technology Kharagpur

9.54/10.0

Best Bachelor's Thesis Finalist [pdf]

Par Excellence Award

Publications and Talks

- Ayush Kaushal, Avirup Saha and Niloy Ganguly, "t-WTWT: Asserting the Role of Targets for Detecting Stance in Tweets", NAACL-HLT'21. [pdf | code | data | website]
- Ayush Kaushal*, Neel Karia* and Faraaz Mallick*, "Leveraging Multi-Staged Language Models for Extracting Measurements, their Attributes and Relations", ACL'21 SemEval. [pdf | code]
- Ayush Kaushal and Tejas Vaidhya, "Leveraging Event Specific and Chunk Span features to Extract COVID Events from Twitter", Published at EMNLP-2020 Workshops. (Oral Presentation, Shared Task Winners) [pdf | code]
- Tejas Vaidhya and Ayush Kaushal, "Domain Specific BERT for Named Entity Recognition on Lab Protocols", Published at EMNLP-2020 Workshops [pdf | code]
- Ayush Kaushal, Lyndon White, Mike Innes and Rohit Kumar, "Basic Tools for Tokenizing Natural Language in Julia", The Journal of Open Source Software. (JOSS) 2020. [pdf | code]
- Zhijing Jin, Julius von Kügelgen, Jingwei Ni, Tejas Vaidhya, Ayush Kaushal, Mrinmaya Sachan, Bernhard Schölkopf, "Causal Direction Matters: Causal and Anticausal Learning in NLP", Accepted at EMNLP'21 [pdf | code]
- Ayush Kaushal, "Natural Language Processing in Julia", JuliaCon 2020 Conference (Contributed Conference talk, video)

Work Experience

IBM Research Labs, India

Remote Research Intern

Bias Detection and Lossless De-biasing of News (GRM)

May 2020 - June 2020

🔗 Improved **fine-grained bias detection** by up to 4.1% F1 using auxiliary objectives of sentiment and VAD lexical score.

🔗 Developed novel task of **bias removal** from news while preserving the information and created its explainable systems.

Google, Summer of Code, 2019

Student Developer

Lead Contributor, DL-NLP Packages; 300+ GitHub stars

May 2019 - August 2019

🔗 Formulated novel regex-free approach for custom multilingual tokenization up to **4 times faster** than SpaCy, NLTK.

🔗 Proposed end-to-end CRF-LSTM model for general-purpose NER, POS tagging APIs with **93.1, 91.2 F1** on CoNLL'03.

Academic Research Experience

IIT Kharagpur, CSE Dept, Undergraduate Researcher

Jan 2020 - April 2021

🔗 Integrated semantics for stance detection using graph attention to improve up to **+10.7% accuracy** on WT-WT dataset.

🔗 Diagnosed spurious cues in stance detection datasets and built a **100k+ size** large target-aware benchmark.

University of Oregon, CS Dept, Remote Researcher

July 2020 - Dec 2020

🔗 Conceptualized **novel task** of *Spatio-temporal video dense captioning* and curated a **1.2M+ Captions** dataset.

🔗 Outperformed referring expression generation SoTA, with a show-attend-tell inspired model delivering **1.29 CIDEr**.

Competitions and Hackathons

🔗 State of the Art System for *Extracting COVID-19 Twitter Events* at W-NUT EMNLP 2020 Shared Task-3; **Oral Presentation**.

🔗 Domain-specific Bert for Named Entity Recognition over lab protocols; Accepted at **EMNLP 2020** noisy text workshop.

🔗 The **1st ranking system** for intelligently parsing doctor's handwritten prescription in the intra-university competition.

Technical Skills

Skills: Python, PyTorch, C++, L^AT_EX, Numpy, Tensorflow, Git, OpenCV, Stanford-Snap, NLTK, SpaCy, AllenNLP

Awards and Achievements

○ **Awards:** Par Excellence Academic Performance, Mitacs Research Scholarship, KVPY Scholarship by India Govt.

○ **Mentorship:** the project 'A Lite BERT' in Google Summer of Code 2020 and NLP mini-projects in Google CI 2019.

○ **Maintainer:** Metakgp, **10k+ monthly users** for university-centric tools, knowledge wiki, APIs, services, bots.

○ **Reviewer:** ACL-SemEval 2021, NAACL 2021, W-NUT-EMNLP 2020; **Volunteer:** EACL 2021, ICLR 2021, NAACL 2021