

Ayush Kaushal

The University of Texas at Austin

✉ ayushk4@utexas.edu

🌐 ayushk4.github.io

☎ +1 (737) 2885781

Education

Masters 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*

Work Experience

IBM Research Labs, India **Remote Research Intern**
Bias Detection and Lossless De-biasing of News (GRM), Deep Learning, NLP
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**
Practical Neural APIs for Sequence Modelling in Text, Deep Learning, Software
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.
🔗 Delivered a 30-minute talk+demo at JuliaCon 2020 on these ML/NLP packages. Also published at JOSS Journal.

Competitions and Hackathons

- 🔗 Best System Design, AllenNLP's 2021 Hackathon, Improved GPT-3 for Text Simplification using BERT.
- 🔗 Winner, Cogito's 2021 challenge to build low-latency multimodal speech emotion recognition systems.
- 🔗 Winners, EMNLP 2020's W-NUT Shared Task, State-of-the-art System for *Extracting COVID-19 Twitter Events*.
- 🔗 Winners, OpenSoft 2018 Kharagpur on softwares for intelligently parsing doctor's handwritten prescription.

Publications

- Asserting the Role of Targets for Detecting Stance in Tweets, **NAACL-HLT'21**. [[pdf](#) | [code](#) | [data](#) | [website](#)]
- Causal Direction Matters: Causal and Anticausal Learning in NLP, Accepted at **EMNLP'21** [[pdf](#) | [code](#)]
- Multi-Staged Language Models for Extracting Measurements, Attributes & Relations, **ACL'21 SemEval**. [[pdf](#) | [code](#)]
- Leveraging Event Specific & Chunk Spans to Extract Twitter COVID Events, Oral **EMNLP'20 Workshops**. [[pdf](#) | [code](#)]
- Basic Tools for Tokenizing Natural Language in Julia, *The Journal of Open Source Software*. (JOSS) 2020. [[pdf](#) | [code](#)]
- Domain Specific BERT for Named Entity Recognition on Lab Protocols, **EMNLP-2020 Workshops** [[pdf](#) | [code](#)]

Academic Research Experience

IIT Kharagpur, CSE Dept, Bachelor's Thesis, *Deep Learning* 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, *Computer Vision* 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**.

Open Source Involvements

JuliaText: Natural Language Processing in Julia Language, Lead Developer Nov 2018 - Sept 2020
🔗 Lead developer of JuliaText packages – TextAnalysis, WordTokenizers with 500+ GitHub stars and 5000+ downloads.
🔗 Mentored student developers working on these packages in Google Summer of Code 2020 and Google Code-In 2019.

MetaKGP: University Centric Apps at IIT Kharagpur, Developer and Maintainer May 2018 - April 2020
🔗 Developed and maintained the University centric apps, wiki, bots and APIs with 11k+ monthly userbase.
🔗 Organized, Curated the contents and led the workshops on Python and Git and GoLang with over 500 attendees.

Technical Skills

Skills: Python, Git, Bash, C++, Numpy, Scikit-learn, PyTorch, Tensorflow, Flask, HTML, OpenCV, SpaCy, SQL