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)   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   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, Bachelor's Thesis   *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

 Runner-Ups on the leaderboard in Cogito's challenge for building low-latency speech emotion recognition apps.
 Winners of the EMNLP 2020' W-NUT Shared Task to design Art System for *Extracting COVID-19 Twitter Events*.
 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.

Open Source Involvements

JuliaText: Natural Language Processing in Julia Language, Lead Developer   *Nov 2018 - Sept 2020*
 Lead the development of packages such as TextAnalysis, WordTokenizers with 500+ GitHub stars and 5000+ downloads.
 Delivered a 30-minute talk+demo at JuliaCon 2020 on using the packages for modern Machine Learning based NLP.
 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 with bots, webapps, APIs, scrapers, and the wiki.
 Maintained the campus's wiki with 11k+ monthly users and added the auto-update feature using bots & scrapers.
 Organized, Curated the contents and led the workshops on Python and Git and GoLang with over 500 attendees.

Publications

 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**]

 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 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**]

Technical Skills

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