

# Ayush Kaushal

Indian Institute of Technology Kharagpur

✉ ayushk4@gmail.com

🌐 ayushk4.github.io

☎ +91 8938802245

## Education

### B.Tech. in Computer Science

Indian Institute of Technology Kharagpur, West Bengal, India

Graduating 2021

GPA 9.52/10 (after 7 semesters)

## Publications and Talks

- Ayush Kaushal, Avirup Saha and Niloy Ganguly, "Stance Detection is not Classification: Promoting the Role of Target Entities", **NAACL-2021** proceedings. (under review) [[pdf](#)]
- 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](#)]
- T.Y.S.S. Santosh, Ayush Kaushal, Avirup Saha and Niloy Ganguly, "Explainable Rumour Detection through Multiple Attention on Response Threads, Response Text and User Posting the Response", (In progress) [[pdf](#)]
- Ayush Kaushal, "Natural Language Processing in Julia", *JuliaCon 2020 Conference* (Contributed Conference talk, [video](#))

## Work Experience

### IBM Research Labs, India

*Bias Detection and Lossless De-biasing of Media Articles*

### Remote Student Researcher

May 2020 - June 2020

- Improved **fine-grained bias detection** by up to 4.1% F1 using auxiliary objectives of sentiment and VAD lexical score.
- Proposed the novel task of **bias removal** from news articles while preserving the information and created a system to de-bias the text in an explainable manner.

### Google, Summer of Code, 2019

*Deep Learning Sequence Labelling Models and Language Processing Packages*

### Student Developer

May 2019 - August 2019

- Leading Contributor for NLP Research packages WordTokenizers and TextAnalysis; **total 300+ GitHub Stars**.
- 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 tagger with **93.1, 91.2 F1** on CoNLL'03.

## Academic Research Experience

### Computer Science Department, IIT Kharagpur

*Research Areas: Natural Language Processing, Deep Learning*

### Undergraduate Researcher

Jan 2020 - Present

- Formulated a **3-level explainable** rumour detection model with up to **+6.3% Accuracy** gains on Twitter15,16 datasets.
- 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 augmented it to build a large bias-free dataset of **100k+ size**.

### Synergetic Holistic Intelligence Lab, University of Oregon

*Research Areas: Multimodal Machine Learning, Vision and Language, Video Captioning*

### Remote Researcher

July 2020 - Present

- Conceptualized **novel task** of *Spatio-temporal video dense captioning* and curated a **1.2M+ Captions** dataset.
- Devised a show-attend-tell inspired model for referring expression generation, achieving SoTA of **1.29 CIDEr**.

## Independent Research Projects

- 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 **1<sup>st</sup> ranking system** for intelligently parsing doctor's handwritten prescription in the intra-university competition.

## Technical Skills

**Skills:** Python, PyTorch, C++, L<sup>A</sup>T<sub>E</sub>X, Numpy, Tensorflow, Git, OpenCV, Stanford-Snap, NLTK, SpaCy, AllenNLP

## Awards and Achievements

- **Awards:** Par Excellence Academic, Mitacs Research Scholarship, KVPY Scholarship-India Govt.
- Among 50 students selected for the **Google India Summer School** AI Research 2020 - NLU Track.
- Mentored the project 'A Lite BERT' in Google, SoC 2020 and NLP mini-projects in Google CI 2019.
- *Maintainer, MetaKGP:* **10k+ monthly users** for university-centric tools, knowledge wiki, APIs, services, bots.
- *Executive Head, Kharagpur Open Source Society:* **2000+ participants** - Lead workshops and events.
- Reviewer: NAACL 2021 and EMNLP 2020 Workshops