

# Rachit Bansal

rachitbansal2500@gmail.com • <https://rachitbansal.github.io> • +91 9205677801

## RESEARCH INTERESTS

Interpretability and Explainability, Robustness, Computational Linguistics, Neural Machine Translation, Transfer Learning, Representation Learning

## EDUCATION

**Delhi Technological University**

Bachelor of Technology (**B.Tech**) in Electrical Engineering

New Delhi, India  
July 2022 (expected)

## EXPERIENCE

**Adobe, Media and Data Science Research (MDSR)**

*Research Intern* (Host: [Balaji Krishnamurthy](#))

Jan 2021 - Sept 2021

Towards making language models factual and commonsense reasoning aware.

- Developed a mechanism of mapping paths from structured knowledge graphs to sentences from a text corpus. Using this we built the first large collection of path-sentence pairs.
- Theorised and implemented a task-agnostic framework to contextualise a given input text through commonsense inferences. Achieved state-of-the-art results across a variety of tasks.

**University of Oxford, Cuneiform Digital Library Initiative (CDLI)**

*Research Intern* (Advisors: [Dr. Jacob Dahl](#) and [Dr. Niko Schenk](#))

Summer 2020

Investigating neural machine translation for low-resource cuneiform languages.

- Adapted NMT techniques from the three broad categories of data augmentation, knowledge transfer, and self-supervised pre-training for Sumerian-English translation.<sup>1</sup> Worked as a part of the [MTAAC](#) team to build an end-to-end information extraction pipeline for Sumerian.<sup>2</sup>
- Established a gradient and perturbation-based saliency framework to conduct thorough qualitative evaluations of various models along with human experts and assyriologists.

**IIIT Delhi, Laboratory for Computational Social Systems (LCS2)**

*Research Intern* (Advisor: [Dr. Tanmoy Chakraborty](#))

May 2020 - April 2021

Retrieving and detecting closed-domain misinformation across social networks.

- Created the first labeled dataset for COVID-19 misinformation by extracting more than 45M domain-specific Tweets and labeling a large portion through a semi-automated approach.
- With [Accenture Technology Labs](#), Bangalore, we introduced an end-to-end explainable neural model, trained using additional semi-supervised adversarial losses, to detect closed-domain misinformation early.

## PUBLICATIONS

- [1] *Evaluating Explanations: How much do explanations from the teacher aid students?*  
Danish Pruthi, [Rachit Bansal](#), Bhuvam Dhingra, Livio Baldini Soares, Michael Collins, Zachary C. Lipton, Graham Neubig, William W. Cohen.  
Transactions of the Association for Computational Linguistics (**TACL**) (conditionally accepted)
- [2] *CoSe-Co: Text Conditioned Generative CommonSense Contextualizer*  
[Rachit Bansal](#), Milan Aggarwal, Sumit Bhatia, Jivat Kaur, Balaji Krishnamurthy.  
– Workshop on Commonsense Reasoning and Knowledge Bases (**CSKB**) at **AKBC**, 2021 [[Print](#)]  
– International Conference on Learning Representations (**ICLR**), 2022 (under review)
- [3] *No Need to Know Everything! Efficiently Augmenting Language Models With External Knowledge*  
Jivat Kaur, Sumit Bhatia, Milan Aggarwal, [Rachit Bansal](#), Balaji Krishnamurthy.  
Workshop on Commonsense Reasoning and Knowledge Bases (**CSKB**) at **AKBC**, 2021 [[Print](#)]
- [4] *How Low is Too Low? A Computational Perspective on Extremely Low-Resource Languages*  
[Rachit Bansal](#), Himanshu Choudhary, Ravneet Punia, Niko Schenk, Jacob L Dahl, Émilie Pagé-Perron.  
**ACL-IJCNLP** Student Research Workshop (**SRW**), 2021 [[Print](#), [Slides](#), [Video](#)]
- [5] *Combining exogenous and endogenous signals with a co-attention network for early fake news detection*  
[Rachit Bansal](#), William Scott, Nidhi Sultan, Tanmoy Chakraborty.  
Pacific-Asia Conference on Knowledge Discovery and Data Mining (**PA-KDD**), 2021 [[arXiv](#), [Slides](#)]
- [6] *Cross-SEAN: A Cross-Stitch Semi-Supervised Attention Model for COVID-19 Fake News Detection*  
[Rachit Bansal](#), William Scott, Abhay Kaushik, Tanmoy Chakraborty, Shubhashis Sengupta.  
Journal of Applied Soft Computing [[arXiv](#)]

<sup>1</sup>Sumerian is the earliest documented language in Mesopotamia, and perhaps the world. Its earliest traces date back to the end of 4th millennium BC from current-day southern Iraq. [[Ref.](#)]

<sup>2</sup>[cdli-gh/Sumerian-Translation-Pipeline](#), [cdli-gh/Semi-Supervised-NMT-for-Sumerian-English](#)

## TEACHING

### Coding Blocks

New Delhi, India

– Student Instructor- Reinforcement Learning (Online)

March 2020 - May 2020

- Recorded 10-hours worth of lectures and held a number of live webinars. Collaborated with course mentors to build project ideas, assignments and quizzes.

– Teaching Assistant- Machine Learning with Deep Learning

June 2019 - August 2019

- Conducted classes and doubt sessions for a batch of 60 senior undergraduate students from all across the country. Built course quizzes and programming assignments in collaboration with other TAs.

## ACADEMIC PROJECTS

### Evaluating Explanations for NLP

Worked with [Danish Pruthi](#) to establish a student-teacher communication paradigm for automatic evaluation of gradient-based saliency attribution methods. I led the development of the paradigm for generative tasks, and enhanced the paper by inculcating empirical findings over multiple thorough review cycles.

### Chrome-SEAN: A Browser Extension to Detect Fake News<sup>3</sup>

Built an easy-to-use chrome extension to predict the possibility of a live tweet status being fake. Based on our work at LCS2, IIIT-Delhi, we deployed a Flask-based API of our misinformation detection model.

### Gaze localisation to Measure Sustained Attention

Worked in collaboration with Samsung R&D Lab, Noida, under the guidance of [Dr. Divyashikha Sethia](#) to curate an image processing module to analyze a person's attention. We engineered this using relative positioning of a person's gaze with respect to certain target regions on a mobile device.

## AWARDS & HONOURS

### Student- [LxMLS](#): 2021

One of the selected students to attend the 11th Lisbon Machine Learning Summer School.

### Fellow- Fatima Predoctoral [Fellowship](#), 2021

One of the selected few fellows for the 9-month fellowship, aimed at research collaboration and mentorship for aspiring PhD students.

### Literary Prodigy Award, 2015

Awarded by The Young Poets Network, UK, for my endeavours in the field of English Literature.<sup>4</sup>

## RELEVANT SERVICE & POSITIONS

- **Volunteer:** NAACL 2021, ICLR 2021, EMNLP 2020, NeurIPS 2020, ICML 2020 & ACL 2020
- **Co-Founder**, Code to School, An initiative to collaborate with schools across the country and teach high school students various programming languages and computer science skills.
- **Mentor**, Tensorflow, Google Code-In
- **ML Lead**, Google Developer Student Club, DTU Chapter
- **Joint Secretary**, Sahitya, the Literary and Debating Society of DTU

---

<sup>3</sup>[ChromeStore/Chrome-SEAN](#)

<sup>4</sup>Check out some of my [poetry](#)