

Rachit Bansal

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

RESEARCH INTERESTS

Computational Linguistics, Neural Machine Translation, Unsupervised and Semi-Supervised Learning, Representation Learning, Interpretability

TECHNICAL SKILLS

Languages: Python, C++, C, JavaScript, MATLAB, Unix Shell Scripting, HTML/CSS

Frameworks: PyTorch, Tensorflow, HuggingFace, FairSeq, Flask

Miscellaneous: Selenium, Seaborn, Material Design, Bootstrap, MongoDB

EDUCATION

Delhi Technological University

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

CPI: 8.72/10

New Delhi, India

July 2022 (expected)

RESEARCH EXPERIENCE

Laboratory for Computational Social Systems (LCS2)

Undergraduate Student Researcher

IIIT Delhi, New Delhi

May 2020 – Present

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

- **Project:** Retrieving and detecting closed-domain misinformation across social networks.

- Extracted more than 45M domain specific Tweets and labelled a part of them by modelling it as an NLI task between known facts and the query text, used BERT Sentence Embeddings and RoBERTa fine-tuned on the SNLI Corpus.

- 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.

Cuneiform Digital Library Initiative (CDLI)

Research Intern & Open-Source Contributor

University of Oxford, UK

June 2020 – September 2020

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

- **Project:** Investigating machine translation techniques for low-resource cuneiform languages.

- Curated, implemented and adapted techniques for Sumerian-English Translation under the three broad categories of Data Augmentation, Knowledge Transfer and Self-supervised Pre-training.¹

- Leveraged University of Oxford's Advanced Research Computing Cluster ([ARC](#)) to run the compute-expensive Transformer models, specially for pre-training. Used gradient and perturbation-based methods to interpret and evaluate the results across the various learning paradigms.

- Worked as a part of the [MTAAC](#) team to curate an end-to-end information extraction pipeline for Sumerian by integrating NER and POS Tagging models with the Semi-Supervised models.²

Samsung Research Lab

Undergraduate Student Researcher

DTU, New Delhi

October 2019 – May 2020

- **Advisor:** [Dr. Divyashikha Sethia](#)

- **Project:** Using Gaze Localisation to Study Sustained Attention on a Mobile Device.

- Experimented across various methodologies to curate an image processing module for analyzing a person's attention using relative positioning of the target and gaze points.

- Worked in collaboration with Samsung R&D Lab, Noida, for deployment of the model on an Android Application and tested it to measure sustained attention of the subjects.

PUBLICATIONS

- Danish Pruthi, **Rachit Bansal**, Bhuvam Dhingra, Livio Baldini Soares, Michael Collins, Zachary C. Lipton, Graham Neubig and William W. Cohen. “**Evaluating Explanations: How much do explanations from the teacher aid students?**” Transactions of the Association for Computational Linguistics (**TACL**) (under review)

- **Rachit Bansal**, Himanshu Choudhary, Ravneet Punia, Niko Schenk, Jacob L Dahl and Émilie Pagé-Perron “**How Low is Too Low? A Computational Perspective on Extremely Low-Resource Languages**” The 59th Annual Meeting of the Association for Computational Linguistics- Student Research Workshop (**ACL-IJCNLP SRW 2021**) (under review)

¹[cdli-gh/Semi-Supervised-NMT-for-Sumerian-English](#)

²[cdli-gh/Sumerian-Translation-Pipeline](#)

- **Rachit Bansal**, William Scott, Nidhi Sultan and Tanmoy Chakraborty “**Combining exogenous and endogenous signals with a co-attention network for early fake news detection**” The 25th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2021) [[Preprint](#)]
- William Scott, **Rachit Bansal**, Abhay Kaushik, Tanmoy Chakraborty and Shubhashis Sengupta “**Cross-SEAN: A Cross-Stitch Semi-Supervised Neural Attention Model for COVID-19 Fake News Detection**” Applied Soft Computing (under review) [[arXiv](#)]

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 Machine Translation

- Worked with [Danish Pruthi](#) to empirically investigate Saliency explanations for Machine Translation. Established a student-teacher communication paradigm for automatic evaluation of source-side attributions, without the need of human intervention and judgment.
- Evaluated techniques like Attention, Integrated Gradients, LRP, and DeepLift on the WMT En-Fr Dataset across Transformer-based Encoder-Decoder architectures.

Chrome-SEAN: A Browser Extension to Detect Fake News³

- Curated an easy to use chrome extension based on our research work of detecting misinformation, Cross-SEAN, to predict the possibility of a tweet status being fake with an accuracy of 95.4%.
- Integrated the functionality of Online Learning by taking real-time user feedback on the prediction and using it conditionally to improve model’s performance.

Power Forecasting using User Behaviour Learning⁴

- Experimented across models like Vanilla LSTMs, CNN-LSTMs, ARIMA, AR-Net and an Ensemble with XG-Boost, in order to forecast power and efficient energy utilisation in a household.
- Worked under the supervision of [Dr. Kapil Sharma](#) to integrate parallel functionalities including GHI Prediction, Appliance Scheduling, and Smart Plugs.

AWARDS & HONOURS

- **Fellow- Fatima Fellowship**, 2021
One of the selected few fellows for the 9-month fellowship, aimed at research collaboration and mentorship for aspiring PhD students.
- **Winner- HackData, IGDTUW**, 2019
Held the 3rd Position at the Nationwide Data Science Hackathon. Made an OCR system for medical prescriptions using CNN-Bi-LSTMs to auto-set health records and reminders in an allied application.
- **Academic Proficiency Award**, 2016
- **Literary Prodigy Award**, 2015
Awarded by The Young Poets Network, UK, for my endeavours in the field of English Literature.⁵

RELEVANT SERVICE & POSITIONS

- **Volunteer**: 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

³[ChromeStore/Chrome-SEAN](#)

⁴[RachitBansal/Power-Forecasting](#)

⁵Check out some of my [poetry](#)