Srijan Bansal

Graduate Research Assistant
Language Technologies Institute, Carnegie Mellon University

srijanb@cs.cmu.edu
GitHub | LinkedIn | Website | Google Scholar

EDUCATION

Carnegie Mellon University, Pittsburgh, USA *Aug'21 - May'23*
Master of Language Technologies; Research Track (Advisor : Prof. Eric Nyberg) GPA: **4.1/4.0**

Indian Institute of Technology Kharagpur, India *Jul'16 - May'20*
Bachelor of Technology in Electronics and Electrical Communication Engineering, GPA: **9.42/10**
Computer Science and Engineering [Minor] (Advisor : Prof. Animesh Mukherjee)

SKILLS AND INTERESTS

Skills: Natural Language Processing, Research, Prompt Engineering, Deep Learning, Machine Learning, Artificial Intelligence, Data Science, Transformers, Neural Networks, Hugging Face, PyTorch, Python
Interests: Large Language Models (LLMs), Question Answering, Model Compression, Conversational AI, Parameter Efficient Learning, Dialogue Systems, Information Retrieval, Multimodal ML, Multilingual NLP, Code-Switching

SELECTED PUBLICATIONS

1. **Srijan Bansal**, Semih Yavuz, Bo Pang, Meghana Bhat, Yingbo Zhou. **Few-shot Unified Question Answering: Tuning Models or Prompts?**, *arxiv preprint* [**pdf**]

2. Srinivas Gowriraj, Soham Dinesh Tiwari, Mitali Potnis, **Srijan Bansal**, Teruko Mitamura, Eric Nyberg. **Language-Agnostic Transformers and Assessing ChatGPT-Based Query Rewriting for Multilingual Document-Grounded QA**, ACL 2023 DialDoc Workshop (accepted)

3. **Srijan Bansal**, Suraj Tripathi, Sumit Agarwal, Teruko Mitamura, Eric Nyberg. **PRO-CS : An Instance-Based Prompt Composition Technique for Code-Switched Tasks**, EMNLP 2022 [**pdf**] [**code**]

4. **S. Bansal**, S. Tripathi, S. Agarwal, S. Gururaja, A. Veerubhotla, R. Dutt, T. Mitamura, E. Nyberg. **R3: Refined Retriever-Reader pipeline for Multidoc2dial**, ACL 2022 DialDoc Workshop [**pdf**] [**code**]

5. **Srijan Bansal**, Vishal Garimella, Ayush Suhane, Animesh Mukherjee. **Debiasing Multilingual Word Embeddings: A Case Study of Three Indian Languages**, ACM Hypertext 2021 [**pdf**] [**code**]

6. **Srijan Bansal**, Vishal Garimella, Ayush Suhane, Jasabanta Patro, Animesh Mukherjee. **Code-switching patterns can be an effective route to improve performance of downstream NLP applications: A case study of humor, sarcasm and hate speech detection**, ACL 2020 [**pdf**] [**code**]

7. Jasabanta Patro, **Srijan Bansal**, Animesh Mukherjee. **A Deep-Learning framework to detect Sarcasm Targets**, EMNLP 2019 [**pdf**] [**code**]

8. T.Y.S.S Santosh, **Srijan Bansal**, Avirup Saha. **Can Siamese Networks help in Stance detection?**, ACM CoDS-COMAD 2019 [**pdf**]

WORK EXPERIENCE

Amazon AWS, Lex Santa Clara, CA
Applied Scientist Intern | Supervisor: Arshit Gupta *May'23 - Aug'23*
- Adaptation-aware Large Language Model (LLM) compression using low-rank attention-head factorization

Salesforce Research Palo Alto, CA
Research Intern | Supervisor: Yingbo Zhou *May'22 - Aug'22*
- Thorough examination of soft prompt tuning towards unified question answering under a low-resource setting

Carnegie Mellon University Pittsburgh, PA
Graduate Research Assistant | Supervisor: Eric Nyberg *Aug'21 - Aug'23*
- Knowledge-grounded conversational question answering in open-domain setting; parameter-efficient training methods

Hitachi Vantara Bangalore, India
Data Scientist | Supervisor: Anirban Basu *Apr'20 - July'21*
- Designed platform to assist DevOps teams to optimize and localize the root cause of issues in containerized applications

- **EPFL (Machine Learning and Optimization Lab)** Lausanne, Switzerland
  Research Intern | Supervisor: Martin Jaggi *May'19 - July'19*
  - Designed momentum-based learning rate scheduler for black-box (zero-order) stochastic convex optimization
- **Hitachi Ltd., R&D Team (Centre for Technology Innovation)** Tokyo, Japan
  Research Intern | Supervisor: Martin Klinkigt *Dec'18 - Jan'19*
  - Explored the applicability of multi-layered capsule neural networks for object recognition and semantic disentanglement
- **IIIT-Allahabad (Speech Image and Language Processing Lab)** Allahabad, India
  Research Intern | Supervisor : U.S. Tiwary *June'18 - July'18*
  - Flexible framework to implement stacked bi-directional LSTM models for image captioning with ResNet architecture

TEACHING

- **CMU 11797 Question Answering** Spring 2023
  Teaching Assistant - Designed & graded course assignments; mentored course projects for several teams

PREVIOUS RESEARCH PROJECTS

- **Re-evaluating the role of human rewritten queries for conversational retrieval**
  Research Study | Supervisor : Prof. Eric Nyberg *Jan'23 - May'23*
  - Evaluated the role of human rewritten queries in conversational retrieval across multiple open-domain conversational question-answering datasets and compared it against LLM (Alpaca-7B) generated rewritten queries.
- **PEFTDB: Parameter Efficient Debiasing of Language models across multiple bias axes**
  Research Study | Supervisor: Prof. Emma Strubell and Prof. Maarten Sap *Jan'23 - May'23*
  - Novel method for debiasing language models with minimal parameter usage, and emphasized the task-agnostic nature of these parameters; evaluated on multiple datasets and bias axes. [**pdf**]
- **Prompt Pruning for Large Language Models (LLMs)**
  Independent Study | Supervisor : Prof. Graham Neubig *Sep'22 - Dec'22*
  - Explored strategies for prompt pruning to incorporate more examples and dynamically adjust prompts for online learning. Assessing the applicability of pruning strategies across various LLMs. [**pdf**]
- **Benchmark for open-domain Conversation Question Answering**
  Research Study | Supervisor : Prof. Eric Nyberg *Sep'22 - May'23*
  - Evaluation benchmark for conversation QA model in full-data and few-shot setting across multiple domains. Propose a flexible framework to distill knowledge through rewritten queries for dense and sparse models.
- **Multihop Multimodal Question Answering**
  Course Project | Supervisor : Prof. Yonatan Bisk *Jan'22 - Apr'22*
  - Propose a two-stage pipeline that employs a corpus-level text retriever and a novel multimodal and multi-hop retriever to achieve fine-grained retrieval of information sources for the WebQA multimodal dataset. [**pdf**]

RELEVANT COURSES

Computational Ethics, Art & Machine Learning, Convex Optimization, Question Answering, Multimodal Machine Learning, Advance Natural Language Processing, Machine Learning for Text and Graph-based Mining, Artificial Intelligence

ACADEMIC ACHIEVEMENTS

- **First** place in the Multidoc2dial shared task (unseen-setting) at the Dialdoc Workshop in ACL 2022.
- Research fellowship provided by CMU includes full tuition coverage and a stipend for the Master's program. (2021-2023)
- Awarded for best student of B.Tech (ECE) having a mix of academic and research potential at the end of 4 years.
- Young Researchers' Symposium **Special Mention Award** for short paper at CoDS-COMAD 2019 Conference
- Scholarship recipient for **Kishore Vaigyanik Protsahan Yojana-2016** and **National Talent Search Exam-2014**

VOLUNTARY INVOLVEMENTS

- Program Committee Member (Reviewer) : Dialdoc Workshop 2023, ICON 2020
- Department Executive Committee Member, Language Technologies Institute, CMU (June'22 - May'23)
- Program Committee Member (Volunteer): ACL 2020

EXTRA CURRICULAR ACTIVITIES

- Member of the **Gold-medal** winning Water Polo team, National Level **Inter-IIT Championship 2019**
- Operational & Technical Assistant (Student Volunteer) at **Rotary India's** health, blood donation & adult literacy camps
- **National** and **Regional** level prizes in Philately, Participation in Hall Activities : Literary, Dramatics & Choreography