

KARAN PRAHARAJ

Ph:+1 (720) 779-5552

[karanpraharaj.github.io](https://github.com/karanpraharaj)

karan.praharaj@colorado.edu

EDUCATION

University of Colorado Boulder
Masters in Computational Linguistics.

August 2020 -

Shiv Nadar University
B.Tech. - Electronics and Communication Engineering

July 2015 - May 2019

TECHNICAL SKILLS

C, Python, SQL, Bash, Git, HTML, PHP, Tensorflow, Pytorch, JavaScript, Microsoft AzureML, AWS, GCP, Docker, Kubernetes, ReactJS, Django, Photoshop, PHP, AJAX, Angular

EXPERIENCE

AI Intern, [Reveal Data](#)
Advisor : Dr. Irina Matveeva

Feb 2021 - Present
Chicago, IL

- Working on extending multi-lingual capabilities to Reveal Data's deep learning based e-discovery and investigation models.
- My work is to focused on investigating different fine-tuning regimes for multi-lingual BERT in the context of e-discovery and Technology Assisted Review (TAR).
- Studied the effect of shuffling, random seed, layer-wise learning rate decay, number of epochs, early-stopping, category-combination training, sequential fine-tuning and custom pre-training.

Research Assistant, [CLEAR Lab](#) @ CU Boulder
Advisor : [Professor Martha Palmer](#)

Feb 2021 - Present
Boulder, CO

- Working on DARPA's Active Interpretation of Disparate Alternatives (AIDA) Program to develop a multi-hypothesis semantic engine that generates explicit alternative interpretations of events, situations, and trends from a variety of unstructured sources, for use in noisy, conflicting, and potentially deceptive information environments.
- My work is to build the interface which provides visualization of schemas induced from information extraction. It is to be designed as a tool for making human curation easier by supporting event-event and event-entity relations, temporal ordering, and schema editing.

Research Intern, [GREYC Lab](#) @ University of Caen-Normandy
Advisor : [Professor Gael Dias](#)

Nov 2019 - Present
Normandy, France

- Worked on Patch-based identification of lexico-semantic relations at the University of Caen Normandie, France. (HULTECH group - HUMAN Language TECHNOLOGY)
- The aim of the project is to devise multi-task strategies for the identification of lexico-semantic relations like hypernymy, synonymy, meronymy and co-hyponymy between a pair of words.
- We generate word patches populated using the k closest neighbours to the words, to identify the most likely semantic relation between any two given words.
- Received a full travel grant and subsistence allowance.

Research Intern, CVIT Lab @ IIIT Hyderabad

Advisor : Professor CV Jawahar

Jan 2019 - June 2019

Hyderabad, India

- Developed an automated academic conference review matching system using Machine Learning, Natural Language Processing and Information Retrieval, at the [Centre for Visual Information Technology](#) at IIIT Hyderabad.
- The system incorporates statistical machine learning to automate the process of assigning submitted research papers to the most suitable reviewer (based on domain expertise) and thereby, consequentially raise the percentage of expert reviewers by a significant margin.
- *Algorithms used* : Latent Dirichlet Allocation, Latent Semantic Indexing

PROJECTS

An Attention and Copy Approach For Synthetic Corpora Generation of Code Switched English-Hindi Sequences

Sep 2020 - Dec 2020

- This project was aimed at solving the problem of a scarcity of data for Hindi-English code-switching by proposing a language model that generates artificial Hindi-English code-switched data.
- Created a language model that leverages a standard transformer augmented with a pointer-generator network to enable the ability copying of words from source texts.

A Comparative Analysis of Beam Search Variants

Sep 2020 - Dec 2020

- Conducted a performance analysis of Beam Search and its weaknesses and compared them against two of its performing variants which alleviate those weaknesses – BULB search and Beam Stack Search.

Ensemble Learning for Regression Applications

July 2018 - Jan 2019

- Analysed the performance (in terms of accuracy and robustness) of ensemble algorithms like Bagging, Boosting and Random Forests against the performance of other machine learning algorithms like Support Vector Regression, Lasso and Ridge Regression, Neural Network Regressor, k-NN and k-means clustering.

The Semantics and Pragmatics of Racial Epithets

Jan 2021 - May 2021

- Studied the relations of slurring and racially charged language with their descriptiveness and if negative expressiveness can be explained by entailment (semantics) or conversational implicature (pragmatics).
- This work was submitted as my final paper submission for the Semantics and Pragmatics class in the Spring term of 2021.

RFID based Navigation System for the Visually Impaired

Jan 2018 - May 2018

- Developed an RFID based system to assist the blind find and identify an object in their proximity. The model uses two-way localization, one using acoustic stimuli and the other using RFID.
- The proposed model provides an affordable solution to the existing problem of lack of sophisticated technological assistance to the VIPs.
- **Published in:** [32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems \(VLSID\)](#)

TEACHING EXPERIENCE

Teaching Assistant, Dept. of Computer Science, CU Boulder
University of Colorado Boulder

Spring 2021

- Conducted programming tutorials for MATLAB and Python.
- Held office-hours to help students with their doubts and questions.
- Designed and graded assignments.
- Moderated a Piazza page (a discussion board) for the class.

Teaching Assistant, TalentSprint
Foundations of AI and ML (Industry Course)

Spring 2019
IIIT Hyderabad

- Taught basic concepts of Machine Learning and Deep Learning techniques for Computer Vision and Natural Language Processing.
- Designed & graded assignments, evaluated exams, conducted tutorials and mentored lab sessions.

Undergraduate Teaching Assistant, Shiv Nadar University
Engineering Mathematics

Monsoon 2017

- Conducted tutorial sessions for 1st Year B Tech students (average class strength of 70).
- Responsible for grading assignments, quizzes, mid-terms and final-term examinations.

COURSEWORK (SELECTED)

Natural Language Processing
Applied Machine Learning
Data Structures
Object Oriented Analysis and Design
Calculus & Linear Algebra
Semantics and Pragmatics

Design and Analysis of Algorithms
Digital Signal Processing
Morphology and Syntax
Calculus - I and II
Probability and Statistics
Linear algebra

POSITIONS OF RESPONSIBILITY

Secretary of Food Affairs
SNU Food Committee

May 2017 - April 2019
Shiv Nadar University

- Oversaw all things related to food and drink on campus and acted as a conduit between university administration, food vendors and the university community of close to 2300 people.
- Spearheaded a team of 13 people to draft a constitution and have it ratified.
- Successfully organized a Food Festival which had a footfall of 800+.
- Introduced benchmarks for food quality and hygiene, an increased variety of food options.

President
FACTION - SNU's Quizzing Society

February 2017 - February 2018
Shiv Nadar University

- Presided over 8 quizzes in my tenure. Average attendance of each quiz was over 80.
- Organised F! - The Inter-college Media and Entertainment quiz. Signed up three sponsors, who helped us give out an overall prize money of Rs. 30000.
- Served as club designer in the previous term.