ROHAN SAHA

PhD Student
Department of Computing Science
University of Alberta

Email: rohansaha60@gmail.com   Webpage: simpleparadox.github.io   LinkedIn: rohansaha60

Research Interests

Deep Learning, Language Models, Human Inspired Learning, Sample Efficient Training

Education

University of Alberta   2022-Present
PhD - Computing Science. GPA: 4.0/4.0   Edmonton, Canada
    Efficient training for small language models.
    Supervisor: Alona Fyshe

University of Alberta   2019-2021
MSc - Computing Science. GPA: 3.5/4.0   Edmonton, Canada
    Using language models to understand how the brain processes language.
    Supervisor: Alona Fyshe

Kalinga Institute of Industrial Technology   2015-2019
Bachelor of Technology, Computer Science and Engineering. GPA: 9.48/10.0   Bhubaneswar, India

Employment

- **Natural Resources Canada, Government of Canada**   2023

  NLP Research Intern - Led a successful project to upgrade the language model based recommendation engine for suggesting similar records on geo.ca Supervisor: Xinli Cai Bo Lu. Location: Ottawa, Canada (Remote)

- **Alberta Machine Intelligence Institute (AMII)**   2023, 2024

  - **AI Career Accelerator Program Participant (Consulting)** - Course content contributor for multiple modules for the ML training course offered by AMII. Location: Edmonton, Canada.
  - **Work Integrated Learning (Consulting)** - Provided valuable consultation to Canadian Startups for integrating machine learning products. Location: Edmonton, Canada.

- **Robert Bosch Engineering and Business Solutions**   2019

  Software Engineering Intern - Teamed with 4 interns to design and develop a mixed reality proof-of-concept application using Microsoft HoloLens to visualize levels of ambient radiation. Supervisor: Ajay Devulapalli. Location: Bengaluru, India.

- **Endurance International Group (Newfold Digital)** 2018

  Software Engineering Intern - Worked with a team of 2 people to upgrade the primary payments collection system to run on the latest version of Ruby on Rails. Supervisor: Gopal Rohra, Adesh Kala. Location: Mumbai, India.

Research

* Equal Contribution

- Rohan Saha, Abrar Fahim, Alona Fyshe, Alex Murphy. Exploring Curriculum Learning for Vision-Language Tasks: A Study on Small-Scale Multimodal Training, CoNLL BabyLM Challenge 2024. Paper

- Rohan Saha*, Jennifer Campbell*, Janet F. Werker, and Alona Fyshe. Insights into early word comprehension: Tracking the neural representations of word semantics in infants, September 2023 Paper.

- Arif Shahriar*, Rohan Saha*, and Denilson Barbosa. 2023. Relational Extraction on Wikipedia Tables Using Convolutional and Memory Networks. July 2023 Paper.

- Rohan Saha, Cassidy Pirlot. Analysis of Evolutionary Program Synthesis for Card Games. January 2021 Paper.

- Rohan Saha. Homonym Identification using BERT - Using a Clustering Approach. April 2020 Paper.

- Rohan Saha. Comparing Classification Models on Kepler Data. Jan 2021 Paper.

Teaching

F = Fall semester, W = Winter semester

- **Teaching Assistant - University of Alberta** F2024

  CMPUT-461/501: Introduction to Natural Language Processing - Assisted course facilitation for teaching with teaching fundamental concepts of NLP.

- **Teaching Assistant - University of Alberta** W2023

  CMPUT-302: Introduction to Human Computer Interaction - Assisted the facilitation of the course to teach and design principles surrounding educational technology applications.

- **Teaching Assistant - University of Alberta** F2020

  PSYCO-576: Machine Learning for Psychology - Designed 10+ hours of jupyter notebooks assignments to teach applications of machine learning to neuroscience data.

- **Teaching Assistant - University of Alberta** F2019,W2020,W2021

  CMPUT-301: Software Engineering - Involved in teaching hands-on labs mentoring teams of undergraduate students for building an android mobile application. Technologies covered: Android, Java, Firestore NoSQL, Unit / Intent testing, Git.

Talks

- **Upper Bound AI Conference:** Large Language Models - A Fundamental Overview. I have a talk at Upper Bound 2023 on the components of language models, covering topics such as attention, pre-training strategies, and scaling laws. Slides.

- **Google ML Study Jam Series**: Hands on walkthrough of a Kaggle Learn course. I gave a talk on how one can train their first machine learning model in python and talked about other foundational ML topics (May 2023). Video.

- **Tech Talk - Linear Models in Machine Learning**: Lead speaker for a session on Linear Models at the Machine Learning Bootcamp organized by Google Developers Group Edmonton and Calgary Data Science Academy (October 2020) Video.

- **Parallels between computer and human representations of meaning (with Dr. Alona Fyshe)** at Broad Institute: How machine learning can be used with brain-imaging data to understand the processing of information in the human brain. (September 2020) Video.

Awards

Name	Year
KIIT Merit Scholarship (INR 30000)	2019

Service

Reviewing

Association of Computational Linguistics (ACL), Neural Information Processing Systems (NeurIPS), North American Chapter for the Association of Computational Linguistics (NAACL), Conference on Computational Natural Language Learning (CoNLL).

Department Service

- Computing Science Graduate Student's Association - Vice President — 2021-2022
- Computing Science Graduate Student's Association - Academic Director — 2020-2021

Other

- Google Developers Groups Cloud Edmonton - Executive — 2021-Present
- Neuromatch Academy (Deep Learning) - Content creator Link. — 2021