Rohan Charudatt Salvi

✉ rohancsalvi@gmail.com   🏠 19-F, Jacqueline Rd,Waltham, MA

📞 217-974-5025   🔗 https://www.linkedin.com/in/rohancsalvi/

ᴳ https://scholar.google.com/citations?user=BennTKkAAAAJ&hl=en

About Me

I am Rohan Charudatt Salvi, a recent master's graduate from the University of Illinois Urbana-Champaign. My primary research interest lies in Natural Language Processing (NLP). I have prior experience and my ongoing research involves the application of NLP techniques to tackle diverse problems. Additionally, I also have a keen interest in working with language models, particularly in the areas of information extraction, conversational agents, biases,and multimodal data. I'm enthusiastic about pursuing a Ph.D. to further my learning and engage in research that advances the frontiers of NLP. Proficient in Python, R, C++, NLP, Deep Learning and Computer Vision.

Education

Aug 2021 – July 2023	**University of Illinois Urbana-Champaign** *MS in Information Science (GPA: 4.0/ 4.0)*
Aug 2016 – May 2020	**Visvesvaraya National Institute of Technology, Nagpur, India** *Bachelor of Technology in Computer Science & Engineering (CGPA: 8.56/10)*

Employment

Sept 2023 – Present — **University of Illinois Urbana Champaign-** *Graduate Research Assistant*
- Advisor: Catherine Blake
- Research focus: **Utilzing Large Language Models for Evidence-Based Medicine**

Sept 2023 – Present — **Kanu Ventures -** *AI Solutions Engineer*
- Provide strategic AI consultation to Kanu portfolio companies, driving project success and contributing to the successful implementation of innovative AI solutions.
- Conduct comprehensive testing and quality assurance for AI products, ensuring AI product reliability, performance, and functionality meet industry standards.
- Develop custom AI models to support data-driven decision-making and create computational tools to enhance data analysis capabilities, utilizing advanced AI techniques to address complex challenges in the ecology field.

May 2023 – July 2023 — **University of Illinois Urbana Champaign-** *Graduate Research Assistant*
- Compared performance of varied information extraction techniques to extract GDPR facets around the use of personal data in Privacy policies.
- Employing prompting methods to extract GDPR facets through large language models

May 2022 – Dec 2022 — **Bayer R&D Services-** *NLP Intern*
- Created a Natural Language Interface for Databases (NLIDB) for the company's Data Warehouse.
- Designed an NLP pipeline consisting of text preprocessing, parsers, and pattern recognition to extract columns, tables, and filtering conditions from the user's query.
- Developed a dialog system-based NLIDB solution capable of generating questions to resolve ambiguity or get more information from the user.

Aug 2021 – Dec 2021 — **WHO Collaboration Center, UIUC -** NLP Research Intern
- Collection of data related to Covid-19 from different social media sources.
- Cleaning the raw data and transforming it into features appropriate for Neural Network models.
- Train and evaluate deep text classification models such as CNN, BERT, and Bi-LSTM to classify information into six categories.

Jul 2020 – Jul 2021	**Tata Research Development and Design Center, India -** *Research Intern*
	● Developed a system dynamics model for covid-19 in India using stock and flow models
	● Simulated and analyzed the effects of interventions such as vaccines, testing, contact tracing, etc.
May 2019 - Jul 2019	**Credit Suisse, India -** *Intern Technology*
	● Built a Content Management System (CMS) for FAQs associated with Compliance.

Awards

iSchool Research Showcase 2022- Second runner up:
Poster: Stereotype Detection in LLM using Computational Grounded Theory.

Publications

- Zheng Qingxiao, Shengyang Xu, Lingqing Wang, Yiliu Tang, **Rohan Salvi**, Guo Freeman, and Yun Huang. 2023. Understanding Safety Risks and Safety Design in Social VR Environments. Proceedings of the ACM on Human-Computer Interaction, CSCW.
- Souvik Barat, Aditya Paranjape, Anwesha Basu, **Rohan Salvi**, Supratim Gosh, Vinay Kulkarni. Modeling and Simulation for the spread of Covid-19 in an Indian City. Proceedings of the 2022 Winter Simulation Conference.
- **Rohan Charudatt Salvi**, Catherine Blake, Masooda Bashir. PrivacyChat: Utilizing Conversational Language Model for Fine-Grained Information Extraction over Privacy Policies. Iconference, 2024.
- Day, K., Christl, D., **Salvi, R.**, & Sriram, P. (2023, March 24). *Video Pre-trained Transformer: A Multimodal Mixture of Pre-trained Experts*. ArXiv.org. https://arxiv.org/abs/2304.10505

Publications (In Review)

- Rohan Charudatt Salvi, Nigel Bosch. Investigating Perception of Stereotypes in Large Language Models: A Computational Grounded Theory Approach. Big Data & Society, 2023.
- Rohan Charudatt Salvi, Catherine Blake. Finding the Devil in the Details: Large language Models for High Precision Information Extraction. Journal of the American Medical Informatics Association, 2023.

Leadership and Teaching Experience

Jan 2022 - May 2022	**Mentored a Junior Student Neha Mathews, University of Illinois Urbana Champaign**
	Poster: Portrayal of COVID-19 by the American News Media, ischool student showcase 2022
Aug 2019 - Dec 2019	**CSL304: Neuro Fuzzy Techniques, Visvesvaraya National Institute of Technology, Nagpur**
	Undergraduate Teaching Assistant - Professor Poonam Sharma
	Created Jupyter notebooks as practice assignments and reference related to neural networks.
Aug 2018 - May 2020	**Student Mentor, Visvesvaraya National Institute of Technology, Nagpur**
	Planned mentorship programs, development seminars, and community bonding activities.

Service

Speaker at following events:

- **NLP and Computer Vision** at NLP Applications Workshop in National Center for Supercomputing Applications Student Research Conference, 2023
- **Introduction to Artifical Intelligence and AI for Ecology**, Ashoka Trust for Research in Ecology and Enviroment
- **Emerging Technologies for Enviroment Conservation Workshop**, Ashoka Trust for Research in Ecology and Enviroment

Student Volunteer for the following conferences:
Computer-Supported Cooperative Work and Social Computing 2022, Winter Simulation Conference 2022

Ongoing Research Projects

Sept 2023   **Professor Catherine Blake -** University of Illinois Urbana-Champaign (UIUC)
Project: Large Language Models for Evidence-Based Medicine
- Fine-Grained Outcome Extraction: Utilizing LLMs to extract granular information and relations from randomized clinical trials, with a particular focus on precise outcomes.
- Assessing medical reasoning capabilites of Pretrained language model chatbots in conversational setting.

Sept 2023   **Professor Nigel Bosch -** University of Illinois Urbana-Champaign (UIUC)
Project: Robust framework for measuring and mitigating stereotypical biases in LLM
- Prompting Methodologies: Developing innovative prompt based task to measure and identify stereotypical biases in LLMs across varied contextual settings. Comparing the model generation on hand crafted prompts against automated prompts
- Evaluation Metric: Creating a novel evaluation metric that analysis biases from multiple lens. We intend to divide the method into two spaces the observed space (probablities, sentiments, toxicity values) and decision space(definition, assumptions and multiple metrics). Our evaluation will also into account the noise in prompts, model parameters and their impact. This comprehensive approach will allows us to gauge biases more accurately.

Spoken Languages

English (Native Proficiency), Hindi (Native Proficiency), Spanish (Beginner Proficiency).