Usneek Singh

✉ usneeksingh1@gmail.com
🔖 Website
𝔾 Google Scholar

Education

Jul 2019- Jun 2023
Birla Institute of Technology and Science (BITS), Pilani, Pilani campus, India.
Bachelor of Engineering in Computer Science, CGPA: **9.68/10**
Relevant Coursework - Deep Learning, Probability and Statistics, Data Structures & Algorithms, Linear Algebra, Information Retrieval, Object Oriented Programming, Programming Languages, Compiler Construction
- Ranked in the **top 1%** among 1100 students in the academic batch of 2023.

Research Interests

AI4Code, Natural Language Processing, Neurosymbolic AI, Software Engineering, Deep Learning

Research Experience

Sep 2023- Present
LLM Fine-tuning for Excel Copilot Agent.
PROSE, **Microsoft Research**, Bengaluru – Advisor: Dr José Cambronero, Dr Aditya Kanade, Dr Sumit Gulwani
- Developed **natural language validation** strategies to obtain high-quality synthetic data for fine-tuning.
- Managed end-to-end pipeline for fine-tuning LLMs on **NL to spreadsheet formula** task in Excel Copilot.
- Conducted research on **Structured data** representation to optimize code generation using language models.
- Designing a low-latency system for formula auto-completion in Excel based on a **neuro-symbolic approach** to extract relevant context from spreadsheets.

Jan 2023 - Jul 2023
Comparative Analysis of Transformers for Modeling Tabular Data.
AI Labs, **American Express**, India – Advisor: Dr Piyush Arora
- Trained 4 **Transformer** architectures on credit fraud prediction problem, achieving a best F1 score of **0.89**.
- Adapted these techniques to **industry** data, enhancing performance by **3%** over the base LightGBM approach.
- Conducted detailed experiments to study the impact of Upsampling and dimensionality on model performance.

Aug 2022- Dec 2022
Text and Multimodal Relation Extraction.
ADAPT Lab, BITS Pilani – Advisor: Prof Poonam Goyal
- Designed a **sequence-to-sequence** model to perform relation extraction in out-of-domain textual content.
- Adapted a **hybrid transformer** architecture for extracting relations from correlated image and text sources.
- Devised an approach to efficiently extract relations when multiple images are interconnected with textual data.

Jun 2022- Aug 2022
Retrieval Augmented Long Document Question Answering.
UKP Lab, **TU Darmstadt**, Germany – Advisor: Prof Iryna Gurevych
- Devised a **retriever-reader** approach for long document question-answering task on **QASPER dataset**.
- Improved the SoTA baseline by **5 F1** points, integrating a Cross Encoder as the Retriever and T5 as the Reader.
- Conducted a comprehensive analysis of language model behaviour, focusing on the impact of relevant information presence in input data.

Jan 2022- Jun 2022
Efficient Image Super-Resolution for Mobile Devices.
Visual Computing Lab, BITS Pilani – Advisor: Prof Pratik Narang
- Devised four-stage super-resolution network incorporating **multiception** convolution layers in distillation block.
- Implemented a **pixel attention** layer post-upsampling, and employed **smooth L1** loss for training.
- Refined the RFDN model, achieving a **28%** reduction in parameters while preserving accuracy on **DIV-2K**.

Aug 2021- Dec 2021
Multilingual Chatbot for Indian Languages.
Web Intelligence and Social Computing Lab, BITS Pilani – Advisor: Prof Yashvardhan Sharma
- Developed a multilingual chatbot leveraging **MuRIL BERT** for fixed response Q/A in 17 Indian languages.
- Enhanced the chatbot to include context-based question-answering by fine-tuning it with the **SQuAD** dataset.
- Conducted benchmarking of multilingual language models for fixed response accuracy, utilizing top-k evaluation.

Jun 2021- Dec 2021
Digital restoration of ancient Indian murals through image Inpainting.
Council of Industrial and Scientific Research **(CSIR-CEERI)**, India – Advisor: Dr Dhiraj Sangwan
- Developed an automatic damage detection method for images, using a **U-net** with optimized training time.
- Innovated a **novel data augmentation** strategy, effectively addressing dataset scarcity in mural restoration.
- Implemented an **inpainting** technique with **partial convolution** for selective restoration of damaged regions.

Publications

2024   **Asking language models how to represent data for fine-tuning**.
U. Singh, A. Singha, A. Awasthi, S. Gulwani, A. Kanade, V. Le, M. Singh, G. Verbruggen
Under review (NAACL) | PDF

2024   **An Empirical Study of Validating Synthetic Data for Formula Generation**.
U. Singh, J. Cambronero, S. Gulwani, A. Kanade, A. Khatry, V. Le, M. Singh, G. Verbruggen
Under review (NAACL) | PDF

2024   **Comparative Analysis of Transformers for Modeling Tabular Data: A Casestudy using Industry Scale Dataset**.
U. Singh, P. Arora, S. Ganesan, M. Kumar, S. Kulkarni, and S. R. Joshi
CODS-COMAD | PDF

2023   **Multilingual Chatbot for Indian Languages**.
U. Singh, N. Vora, P. Lohia, Y. Sharma, A. Bhatia, and K. Tiwari
ICCCNT | PDF

2023   **Ancient Indian Murals Digital Restoration through Image InPainting**.
U. Singh, S. Maiti, A. Saini, and Dhiraj
IEEE SPIN | PDF

2022   **MFDN: Multiception Feature Distillation Network**.
S. Sameen*, **U. Singh***, and P. Narang
TENCON | PDF

Awards

2023   **ACM Student Grant**, *Received a travel grant worth Rs. 30,000 from ACM to attend the CODS-COMAD conference in Bengaluru, India*.

2022   **DAAD-WISE Scholarship**, *Received a scholarship worth Rs. 3,00,000 from DAAD, Bonn for a research internship in Germany. This prestigious award is granted to the top 100 students across India.*.

2017   **NTSE Scholarship**, *Selected among the top 750 students out of 2 lakh applicants in India in the national science examination. Awarded a scholarship worth Rs. 2,00,000 by the Govt. of India.*.

Teaching Experience

Aug 2022- Dec 2022   **Computer Programming**, *Teaching and Lab Assistant*.
Instructed 100 students in weekly 2-hour labs, focusing on fundamental concepts in C programming.

Jan 2022- May 2022   **Data Structures & Algorithms**, *Teaching and Lab Assistant*.
Facilitated weekly 3-hour programming labs for 80 students. Developed lab exercises with contemporary topics and industry-relevant questions to aid in placement preparation.

Leadership and Volunteer Work

Aug 2019- Jun 2022   **Nirmaan Organisation**, *Project Leader*.
Led a team of 20 teaching volunteers in Baas village, India. Devised a comprehensive plan to enhance spoken and written English skills, benefiting over 100 Nirmaan beneficiaries.

Aug 2020- Dec 2020   **Student Faculty Council, BITS Pilani**, *Student Representative*.
Served as a student representative in the Department of Computer Science faculty meetings, playing a key role in incorporating student feedback into course design and curriculum development.

Academic Projects

Feb 2022- May 2022   **Creation of custom compiler [Code]**.
Course Project: **Compiler Construction**, Advisor: Prof Vandana Aggarwal
- Collaboratively designed and developed a custom language compiler in C within a team of five members.
- Implemented all phases of the compiler, from parsing to assembly code generation from scratch.

Mar 2022- May 2022   **Sentiment Analysis of Blogger's Opinion [Code]**.
Course Project: **Information Retrieval**, Advisor: Prof Poonam Goyal
- Developed a sentiment analysis model using a BiLSTM, augmented with attention layers for improved accuracy.
- Crafted an embedding strategy that integrates Glove embedding, TF-IDF scores, & WordNet sentiment scores.

Sep 2021- Nov 2021   **Person Identification using Forehead images [Code]**.
Course Project: **Deep Learning**, Advisor: Prof Kamlesh Tiwari
- Curated a dedicated forehead image dataset by extracting cropped images from 100 individuals.
- Compared performance on baseline models- SIFT, SURF, FaceNet and ResNet50. **Best accuracy** - **89%**.