

BHAVYAJEET SINGH

+91 97200 38855 ✉ bhavyajeet29@gmail.com [/in/bhavyajeet-singh](#) @bhavyajeet Google Scholar

Education

International Institute of Information Technology (IIIT), Hyderabad

Hyderabad, India

B. Tech. and M.S. in Computer Science and Engineering GPA: 9.59/10.0

2018 - 2023

- Graduated second in a batch of 65 students; Awarded the **Dean's List for 4 consecutive years** for being in the top 5% of the batch from 2018 to 2022.
- Awarded the **Dean's Research Award** for significant research contributions and multiple publications.
- Teaching associate for Information Retrieval and Extraction; and Music, Mind and Technology courses.
- Teaching assistant for multiple courses - Data and Applications, Introduction to Software Systems, Discrete Structures, Embedded Systems Workshop, Software Programming for Performance, Design and Analysis of Software Systems.

Research and Professional Experience

Microsoft Research

Research Fellow, Extreme Classification Team

August 2023 – Present

- Advisors: Yashoteja Prabhu, Niket Tandon, Manish Gupta, Manik Verma.
- Working with the **Bing Ads** team to augment language models with a **Memory Bank** to store auxiliary data. This allows late infusion of potentially private data from multiple sources into the model, lending improved retrieval capabilities to the model with a **sub-20ms latency**.
- Constructed efficient dual-encoder models for **large scale** tagging of over a billion Bing queries with ad-relevant named entities from a set of 3M entities, improving ad relevance.
- Benchmarked approaches of using LLMs for multi-label and multi-class classification with millions of candidate labels.

Kohli Center on Intelligent Systems, IIIT Hyderabad

Research Assistant

March 2020 – June 2023

- Worked with **Prof. Vasudeva Varma** and **Dr. Manish Gupta** on factually grounded cross-lingual data-to-text generation to enrich long-form articles like Wikipedia in low-resource languages. The work was funded by the **Wikimedia Foundation**.
- Formally defined **Cross-Lingual Fact-to-Text Generation (XF2T)** and **Cross-Lingual Fact-to-Long-Text Generation (XFLT)**, commonly encountered in business scenarios, as novel NLG tasks. Formulated novel pipelines for these tasks and introduced task-specific evaluation metrics; work published at **TheWebConf 2022** and **ECAI 2023**.
- Applied NLP to the field of Music Information Retrieval with **Prof. Ponnurangam Kumaraguru** and **Prof. Vinoo Alluri**. Analysed music sharing patterns and intents over various social media platforms like twitter and Reddit in the context of depression; Work published at **PLOS ONE 2023** and **ISMIR 2024**.

MotorQ, Inc.

Software Engineer Intern

May 2022 – August 2022

- Designed, developed and deployed an automated pipeline for compressing and re-hydrating developer logs using Kibana, reducing storage requirements by over 93%.
- Implemented real-time data-monitoring algorithms to detect anomalies and warn impending failures in critical pipelines.

Publications

Conference and journal papers

- **Bhavyajeet Singh**, Aditya Hari, Rahul Mehta, Manish Gupta, Vasudeva Varma. *XFLT: Exploring Techniques for Generating Cross Lingual Factually Grounded Long Text*. **ECAI, 2023**
- **Bhavyajeet Singh**, Kunal Vaswani, Sreeharsha Paruchuri, Suvi Saarikallio, Ponnurangam Kumaraguru, Vinoo Alluri. *"Help! I need some music!": Analysing Music Discourse & Depression on Reddit*. **PLOS ONE, 2023**
- Tushar Abhishek, Shivprasad Sagare, **Bhavyajeet Singh**, Anubhav Sharma, Manish Gupta, Vasudeva Varma. *XAlign: Cross-lingual Fact-to-Text Alignment and Generation for Low-Resource Languages*. **TheWebConf, 2022**
- Rahul Mehta, **Bhavyajeet Singh**, Vasudeva Varma, Manish Gupta. *CircuitVQA: A Visual Question Answering Dataset for Electrical Circuit Images*. **ECML PKDD, 2024**
- Pavani Chowdary, **Bhavyajeet Singh**, Rajat Agarwal, Vinoo Alluri *Lyrical Speaking: Exploring the Link Between Lyrical Emotions, Themes and Depression Risk*. **ISMIR 2024**
- Shivprasad Sagare, Tushar Abhishek, **Bhavyajeet Singh**, Anubhav Sharma, Manish Gupta, Vasudeva Varma. *XF2T: Cross-lingual Fact-to-Text Generation for Low-Resource Languages*. **INLG 2023**
- **Bhavyajeet Singh**, Pavan Kandru, Anubhav Sharma, Vasudeva Varma. *Massively Multilingual Language Models for Cross Lingual Fact Extraction from Low Resource Indian Languages*. **ICON 2022**.

Workshop and shared tasks

- Himanshu Maheshwari **Bhavyajeet Singh**, Vasudeva Varma. *Scibert sentence representation for citation context classification*. Proceedings of the Second Workshop on Scholarly Document Processing, **ACL-2021 (Awarded the best paper award for the task)**.
- **Bhavyajeet Singh**, Aditya Hari, Rahul Mehta, Tushar Abhishek, Manish Gupta, Vasudeva Varma. *Cross-lingual Multi-Sentence Fact-to-Text Generation: Generating factually grounded Wikipedia Articles using Wikidata*. In Proceedings of **The Wiki Workshop 2023**
- Three accepted system description papers at The 17th International Workshop on Semantic Evaluation, 2023.

Projects

Taiga : Conversational Assistant

NLP, Flutter

- Built a **conversational digital social companion** by finetuning generative models over the empathetic dialog dataset designed to cater to the elderly.
- Packaged the conversational agent in the form of a mobile application built using flutter with **elderly friendly UI** and other helpful features.

Computer Vision Toolkit

OpenCV, python, pytorch

- Built a Computer Vision toolkit for tasks like image mosaicing, background removal (using GrabCut), and Scene recognition (using bag-of-visual-words approach).
- Created a pipeline for multi class image classification using imagenet and CNN.

Wikipedia Search Engine

IRE, Python

- An efficient indexing and search engine for English Wikipedia with blocked-sort indexing and TF-IDF based ranking to serve general and field queries on 64 GB of XML in less than 1 seconds.

Project PreQL

Database Systems, C++

- Implemented a CLI based DBMS from scratch for Insertion, Deletion, Selection and Aggregation commands, similar to that of SQL, with support for B+ tree and linear hashing type indexes.

Linux Shell

Operating Systems, C

- A shell for linux implemented from scratch in C with support for piping, redirection and multiple system commands.

Achievements and Activities

Microsoft Garage Hackathon

Microsoft, Sept 2023

- Secured **second rank** in the executive challenge at the annual hackathon by developing a finetuning multimodal language models for the task of radiology report writing.

Google Research Week

Bengaluru, Jan 2023

- One of the **top 70 students** selected from across the country to attend the NLP track at Google Research Week 2023.

India-EU ICT Smart City Hackathon

BITS Goa, Jan 2020

- Secured **first rank** at the annual hackathon aimed at providing solutions for smart cities by building an innovative low cost air quality monitoring and filtering tool using IoT and one M2M protocols.
- Was awarded a **fully funded Europe trip** to the Bits & Pretzels Entrepreneurs' Conference 2022 at Munich, Germany.

Hack For Positivity 2021

Jan 2021

- Won an international hackathon, sponsored by **Google and MongoDB Atlas**, by creating a Browser extension which can filter out the negative and hateful news articles using sentiment analysis and recommend positive or wholesome news.

IBM Hackathon

IISc Bangalore, Jan 2019

- Won the prize for **best results** at pravega 2019 IBM hackathon at IISc Bangalore for building a fake news detector using IBM watson.

ACM ICPC

2020

- Qualified for the ACM ICPC Asia-West Amritapuri regionals, securing the 71st place amongst 270 teams.

Leadership Positions

IIIT Hyderabad, 2019-2022

- Elected as a **member of Students' Parliament** for 2 consecutive years, working closely with the faculty and the student community of IIIT.
- Part of the Apex body at IIIT, a 16 membered team responsible for planning the admissions and induction for the incoming batches and mentoring over 500 students.
- Served as the **Coordinator of Literary Club at IIIT-H** for a duration of 2 years and organised multiple events like treasure hunts, book reading clubs, quizzes etc.