Harshit Joshi

harshitjoshi.in | LinkedIn | Github
josharshit@gmail.com | +91 991 026 8058

EDUCATION

UNIVERSITY OF DELHI
BTech in Information Technology
June 2021 | New Delhi
Cluster Innovation Center
Department Rank: 2
Major Percentage: 89.83%
Cum. Percentage: 87.88%
CGPA: 9.4 / 10

LINKS

Github:// **duskybomb**
LinkedIn:// **harshitjos**
Twitter:// **duskybomb**

COURSEWORK

UNDERGRADUATE
Software Engineering
Intro to AI
Object Oriented Programming
Computational Linguistics
Operating Systems
Computer Graphics
Graph Theory
Statistics and Probability
Numerical Analysis

SKILLS

PROGRAMMING
Over 5000 lines:
Python • C++ • R • LaTeX
• PyTorch
Over 1000 lines:
Java • ElasticSearch • C#
• MySQL • Kafka • MongoDB
Familiar:
Javascript • VueJs • F#

OPEN SOURCE

invoice2data: Core Contributor
invoicex-gui: Author

SERVICES

Volunteer
EMNLP 2020, EACL 2021, ICWSM
2021 Program Committee
ICML CA2MH 2021

EXPERIENCE

MICROSOFT RESEARCH | Research Fellow (PROSE)
November 2021 - Present | Remote | Redmond, US
- Designed and trained domain-specific LLMs. Covered by media outlets.
- Built SOTA production-ready neurosymbolic program repair framework (OOPSLA '22)
- Created multilingual repair framework, RING, built on top of LLMs (AAAI '23).

SUPEDIO GMBH | Research Software Engineer
January 2021 - October 2021 | Remote | HQ: Dresden, Germany
- Built automation tools, reducing person-hours employed per day by 400%
- Designed symbolic framework for the pdf data extraction for production.

CRONYCLE LTD. | Software Engineering Intern
Jan 2019 – Oct 2019 | New Delhi, India | HQ: UK
- Ported batch jobs to live production, latency by 5 minutes.
- Built a new data pipeline to increase RSS data dump by 10%.

RESEARCH

IIITD - MIDAS LAB | Researcher
May 2020 – Present | New Delhi, India
Working with **Ramit Sawhney** and **Prof Rajiv Ratn Shah** on Suicide Risk Assessment on social media using NLP and data mining techniques (EMNLP, WSDM, EACL, ICWSM, NAACL).

DEFENCE R&D ORGANISATION (DRDO) | Researcher
May 2019 – October 2019 | New Delhi, India
Worked on navigation of bot in Gazebo environment. Used CityScape Dataset for Image Segmentation and used the features in Gazebo environment, building algorithms for depth perception and creating disparity maps.

AWARDS

2020	Honorable Mention	COMAP Mathematical Contest in Modeling
2019	Invited Talk	PyData Delhi 19 Conference
2019	National	Mathematical Finance Scholar Summer Fellowship 2019
2018	International	Google Summer of Code 2018
2017	Honourable Mention	ACM-ICPC 2018 (Regionals)

PUBLICATIONS

1. **H. Joshi**, A. Ebenezer, J. Cambronero, S. Gulwani, A Kanade, V. Le, I. Radicek and G. Verbruggen. *FLAME: A small language model for spreadsheet formulas* Under Submission, 2023 `https://arxiv.org/abs/2301.13779`.

2. **H. Joshi**, J. Cambronero, S. Gulwani, V. Le, I. Radicek and G. Verbruggen. *Repair Is Nearly Generation: Multilingual Program Repair with LLMs.* AAAI 2023 `https://arxiv.org/abs/2208.11640`.

3. R. Bhavishi*, **H. Joshi***, J. Cambronero, A. Fariha, S. Gulwani, V. Le, I. Radicek and Ashish Tiwari. *Neurosymbolic Repair for Low-Code Formula Languages.* In In Proceedings of the ACM on Programming Languages (OOPSLA) 2022.