

**HARSHIT JOSHI** New Delhi, India  
(+91) 767 864 0404 ◊ joshharshit@gmail.com ◊ harshitjoshi.in

## EDUCATION

---

**Cluster Innovation Centre, University of Delhi** *July 2017 - Present*  
Senior Year (4<sup>rd</sup> Year) Undergraduate Overall Percentage: 85.8% — Major: 87.39%  
B. Tech. (Information Technology and Mathematical Innovations) Department Rank: 2  
Minor in Management and Economics

## INTERESTS

---

NLP, Computational Social Science, Computer Vision, Deep Learning, Statistics

## RESEARCH EXPERIENCE

---

**Multimodal Digital Media Analysis Lab (MIDAS@IITD)** May 2020 - Present  
*Researcher* *New Delhi, India*

- Using NLP to tackle problems in Computational Social Science.
- Making interpretable models and highlighting ethics and bias in AI.

**Defence Research and Development Organisation (DRDO)** June 2019 - October 2019  
*Research Intern* *New Delhi, India*

- Used CityScape Dataset for Image Segmentation through implementation of DeepLabV3+
- Fine-tuned the model for two classes: void space and obstacle for cognitive mapper.

## PROFESSIONAL EXPERIENCE

---

**Supedio GmbH** January 2021 - Present  
*SDE Intern* *Dresden, Germany*

- Building the first data pipeline at Supedio
- Leveraging Graphs and Natural Language Processing for matching document layouts and content.

**Cronycle Ltd.** January 2019 - July 2019  
*Software Engineering Intern for Data Science* *U.K. - Remote*

- Ported batch jobs to live production using Kafka and Elastic Search, reducing latency by 5 minutes.
- Increased RSS collection dump by 10% by identifying new data sources and built a pipeline to send it to the MongoDB

**Google Summer of Code 2018** April 2018 - August 2018  
*Student Developer at Debian Project* *Remote*

- Organisation Link: [github.com/invoice-x](https://github.com/invoice-x) Project Report: GSoC 2018
- Extracting Data from PDF Invoices and Bills Details using a Regular Expression based Engine.
- Enhanced tesseract-OCR integration and increased code coverage by 16% by adding tests for functions.
- Developed a GUI application for three major OS (Linux, macOS, Windows) using Python and PyQt.

## TECHNICAL STRENGTHS

---

**Computer Languages** Python, C/C++, JAVA, R, SQL  
**Software & Tools** PyTorch, Tensorflow, OpenCV, MySQL, MongoDB, L<sup>A</sup>T<sub>E</sub>X, Git, Elastic Search, Kafka, MATLAB, Mathematica  
**Platform** Linux, Windows, Mac

## PUBLICATIONS

---

R. Sawhney\*, **H. Joshi\***, A. Nobles\*, and R. R. Shah, 2021. Tweets Classification to Assist Human Moderation for Suicide Prevention. In Proceedings of the International AAAI Conference on Web and Social Media 2021 (ICWSM).

R. Sawhney\*, **H. Joshi\***, R. R. Shah, and L.Flek, 2021. Suicide Ideation Detection via Social and Temporal User Representations using Hyperbolic Learning. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (NAACL)

R. Sawhney\*, **H. Joshi\***, L.Flek, and R. R. Shah, 2021. Phase: Learning Emotional Phase-Aware Representations for Suicide Ideation Detection on Social Media. In Proceedings of 16th Conference of the European Chapter of the Association for Computational Linguistics, EACL 21, Association for Computational Linguistics

R. Sawhney, **H. Joshi**, S. Gandhi, D. Jin, and R. R. Shah, 2021. Robust Suicide Risk Assessment on Social Media via Deep Adversarial Learning. Journal of the American Medical Informatics Association (JAMIA).

R. Sawhney, **H. Joshi**, S. Gandhi, and R. R. Shah, 2021. Towards Ordinal Suicide Ideation Detection on Social Media. WSDM' 21. Association for Computing Machinery.

R. Sawhney, **H. Joshi**, S. Gandhi, and R. R. Shah, 2020. A Time-aware Transformer based Model for Suicide Ideation Detection on Social Media. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).

## NON-ARCHIVAL ARTICLES

---

R. Sawhney, **H. Joshi**, S.Gandhi, and R. R. Shah, 2021. A Time-Aware Transformer Based Model for Suicide Ideation Detection on Social Media. Machine Learning for Health Workshop. Neural Information Processing Systems.

## NOTABLE PROJECTS

---

### **Cognitive Mapping and Planner for Navigation** *DRDO*

Scientist. S.P. Mishra  
*June 2019 - October 2019*

- Worked on the navigation of bot in the Gazebo environment.
- Used CityScape Dataset for Image Segmentation and used the features in the Gazebo environment
- Built an algorithm for depth perception using image segmentation results and creating disparity maps.

### **Strategizing Fantasy Football** *Probability and Statistics*

Associate Professor Dr. Sonam Singh  
*March 2019 - April 2019*

- Made new graphs for better analysis and visualization of players' performance by dividing the graphs into different grids into points vs cost axis.
- Used Gradient Boosting Trees, Stochastic Gradient Regressor, and Linear Regression to regress points scored against 32 features engineered through priori.

## POSITION OF RESPONSIBILITY

---

### **Govt. of India's Institutions Innovation Council** *Student Coordinator*

November 2018 - October 2019  
*University of Delhi*

- We achieved a 4-star rating, given by MHRD, Govt. of India
- Lead a team of top 10 innovators, who oversee the innovations across Delhi University.
- 10 teams were shortlisted for the Regional round from our council, which is the most from any University.

### **HashInclude - Computer Science Society** *External Affairs*

August 2018 - August 2019  
*Cluster Innovation Centre*

- Spearheaded a team of 40 people to conduct professional shows, exhibitions and talks successfully
- Conducted workshops on Open Source Development and Linux 101
- Organised two State-level student Hackathons with more than 500 participants.

## AWARDS AND ACHIEVEMENTS

---

Received Honorable Mention at 2020 **COMAP's** Mathematical Contest in Modeling (MCM). **Only team from India** to get Honorable Mention.

Invited to give a talk at **PyData Delhi Conference 19** on "Quantitative Finance with R"

Member, Football Team, Cluster Innovation Centre. Reached 2<sup>nd</sup> round of **Reliance Youth Sports 2019** in our inaugural year of participation.

Mathematical Finance Scholar under Focus Areas in Science and Technology **Summer Fellowship 2019**

**Google Summer of Code 2018** with Debian Project

Qualified for **ACM-ICPC 2018** Kolkata Kanpur Site contest held at UIET Kanpur