

# Manan Suri

[manansuri.com](https://manansuri.com)

[manans@umd.edu](mailto:manans@umd.edu)

University of Maryland, College Park

+1(202)-999-8225

[Github](#)

[LinkedIn](#)

[Google Scholar](#)

## EDUCATION

**University of Maryland, College Park** 2024 - Present

- *Master of Science - Computer Science*

Advised by [Prof. Dinesh Manocha](#)

**Netaji Subhas University of Technology, New Delhi** 2020 - 2024

- *Bachelor of Technology - Computer Science and Engineering*

CGPA: 8.99

**K.R. Mangalam World School, New Delhi** 2016 - 2020

- *Higher Secondary Education*

Grade 12 - 97.2%, Grade 10 - 96.6%

## SELECT PUBLICATIONS

- **VisDoM: Multi-Document QA with Visually Rich Elements Using Multimodal Retrieval-Augmented Generation**  
**NAACL 2025**

Manan Suri, Puneet Mathur, Franck Dernoncourt, Kanika Goswami, Ryan A Rossi, Dinesh Manocha

- **DocEdit-v2: Document Structure Editing Via Multimodal LLM Grounding**

**EMNLP 2024**

Manan Suri, Puneet Mathur, Franck Dernoncourt, Rajiv Jain, Vlad I Morariu, Ramit Sawhney, Preslav Nakov, Dinesh Manocha

- **CoSyn: Detecting Implicit Hate Speech in Online Conversations Using a Context Synergized Hyperbolic Network**

**EMNLP 2023**

Sreyan Ghosh\*, Manan Suri\*, Purva Chiniya\*, Utkarsh Tyagi\*, Sonal Kumar\*, Dinesh Manocha

- **ACLM: Selective-Denoising based Generative Data Augmentation for Low-Resource Complex NER**

**ACL 2023**

Sreyan Ghosh, Utkarsh Tyagi, Manan Suri, Sonal Kumar, Ramaneswaran S, and Dinesh Manocha

- **Boosting Pre-trained Language Models with Task Specific Metadata and Cost Sensitive Learning**

**SemEval 2022 at NAACL 2022**

Manan Suri

## EXPERIENCE

**University of Maryland, College Park** Aug 2024 - Present

- *Graduate Research Assistant*

- \* Analyzing algorithmic behaviors across various social media platforms to assess their impact on information reach and engagement in high-stakes crisis situations, contributing to the creation of trustworthy AI frameworks.
- \* Collaborating with interdisciplinary teams, including Dr. Giovanni Luca Ciampaglia (UMD), Dr. David Broniatowski (GWU), and Dr. Erica Gralla (GWU), as a part of a TRAILS (Trustworthy AI in Law and Society) Seed project.

**University of Warwick, United Kingdom** June 2023 - Aug 2023

- *Data Science for Social Good Fellow*

- \* Constructed advanced, replicable pipelines for training, experimentation, and inference, leveraging Deep Learning to devise an innovative system that identifies and mitigates greenwashing in multi-modal social media content, spanning text, video, and image elements.
- \* Worked in close collaboration with the Algorithmic Transparency Institute and environmental specialists to understand project needs. Developed tailored strategies to promote the use of ethical and transparent algorithms crucial in combating greenwashing and supporting sustainability in the digital content space.

**Scalenut, New Delhi, India** Dec 2022 - Jan 2023

- *Data Science Intern*

- \* Lead the development of advanced analytical tools leveraging Large Language Models (LLMs), notably designing a high-performing fact attributor to enhance the capabilities of the Cruise Mode product.
- \* Initiated and successfully implemented optimization strategies for document ranking and retrieval functions, resulting in improved accuracy and efficiency across multiple products.

**Northeastern University** Apr 2022 - Jan 2023

- *Research Intern*

- \* Conducted in-depth research on various significant themes such as Depressive Tendencies, Trigger Warnings, and Fine-grained Sexism Detection, demonstrating analytical capabilities in understanding and addressing critical societal issues, some of this work is published in workshops at top conferences including AAAI 2023, ACL 2023.
- \* Efficiently managed four distinct research teams, demonstrating leadership skills and proficiency in coordinating and executing complex research projects.

## PROJECTS

---

- **Nanovector — Lightweight Text-to-Vector Database:** Machine Learning, Embeddings, Database Systems, Docker
  - Developing an open-source, lightweight database designed for efficient text-to-vector storage and retrieval, optimized for seamless integration and deployment.
  - Implemented direct text-to-vector pipeline allowing streamlined conversion of text data to vector embeddings within the database, using Sentence Transformers for high-quality embeddings and PCA for dimensionality reduction.
  - Designed for versatility, Nanovector supports a wide range of applications including embeddings and feature storage, with deployment ease through Docker for containerized environments.
  - Tech Stack - *Python, NumPy, Docker, Sentence Transformers* (September '24)
- **Asatya — Fake News Browser Extension:** Machine Learning, NLP, Web Scraping, Web Development
  - Developed a robust suite of machine learning-based tools aimed at combating the spread of fake news.
  - Designed the system to generate key metrics for any given news item, including measures of reliability, biasness, and objectivity. Additionally, implemented functionality to provide users with a concise summary of the news report and recommendations for similar yet reliable news headlines.
  - Tech Stack - *Flask, BeautifulSoup4, pandas, tinydb, scikitlearn, joblib, react* (January '22)
- **Article — Comprehensive NLP Suite for Web Article Analysis:** Machine Learning, NLP, Web Scraping, LLM Deployment
  - Developed an innovative tool for extracting and structuring information from web articles, utilizing heuristic techniques and a wide array of NLP techniques such as Entity Recognition, N-gram analysis, keyword extraction, extractive summarization and text classification for comprehensive article analysis.
  - Incorporated advanced features like similar data mining, media bias detection, factuality checks and querying with Language representation models, providing a powerful content verification apparatus and enabling a broad range of media research applications.
  - Tech Stack - *Flask, BeautifulSoup4, pandas, RAKE, YAKE, scikitlearn, joblib, OpenAI* (February'23)

## TECHNICAL SKILLS

---

- **Languages:** Python, Java, C++, C, Javascript,HTML/CSS, SQL
- **Skills:** Machine Learning, Deep Learning, Natural Language Processing, Full Stack Web Development, Database Management
- **Tools & Technologies:** Scikitlearn, Pytorch, NumPy, Pandas, Flask, Node, React, Express, MongoDB, MySQL

## AWARDS AND ACHIEVEMENTS

---

- Research papers accepted at top conferences including **ACL**, **AAAI** and **NAACL**.
- **Adobe Career Academy** - Selected to be a part of the first Adobe Career Academy cohort by Adobe India.
- **TechForGood Hackathon** - Won **First Prize** amongst 3000 participants from over 36 countries.
- **KVPY** - Secured **AIR 159** amongst 150k students and was offered the fellowship by IISc Bangalore
- **JEE Advanced** - Secured **AIR 4290** among 150k students qualified from 1.5 million initial applicants
- **JEE Mains** - Secured **AIR 7172** among 1.5 million students with a percentile of 99.36.
- **Student of the Year, 2019** - Awarded by the **Times of India**, India's leading newspaper publication.
- **Chairman's Medal** for exemplary performance in academics throughout grades VI-XII.

## EXTRA-CURRICULAR

---

- Coordinating Project Agni which aims to build products by upcycling firehoses and empowers women in the process.
- Director, Technical at Enactus NSUT.
- Parliamentary Debating Mentor at The Debating Society of NSUT.
- Head Boy 2019-20, K.R.Mangalam World School.
- Founding President of TED-Ed Club at my highschool, K.R. Mangalam Worls School.