+1 (312) 998 - 9170
yosheej2@illinois.edu
github.com/YosheeJain
linkedin.com/in/yosheejain
yosheejain.github.io

Yoshee Jain

EDUCATION

University of Illinois Urbana-Champaign Champaign-Urbana, IL
- Bachelor of Science in Computer Science May 2026
- Bachelor of Science in Liberal Arts and Sciences Major in Statistics GPA: 4.0/4.0
- **Honors and Awards**
  - Chancellor's Scholar
  - Edmund J. James Scholar
  - Dean's List (FA22, SP23, FA23)
  - Summer Research Award (Campus Honors Program) (2024)
  - Illinois Engineering Achievement Scholarship (2024-2025, 2023-2024)
  - Illinois Engineering Outstanding Scholarship
  - Distinguished Undergraduate Researcher Certificate
  - Distinguished Active Member, Tau Beta Pi (The Engineering Honor Society)
- **Relevant Coursework:** Computational Social Science, Data Structures (C++), Systems Programming (C), Data Science Discovery, Statistics and Probability, Statistical Programming (R), Algorithms and Models of Computation, Data Visualization, Database Systems

RESEARCH EXPERIENCE

Human-Computer Interaction Institute. Carnegie Mellon University (& University of Memphis) Remote
Undergraduate Research Intern, Mentor: Dr. Jionghao Lin (& Liang Zhang) *May. 2024 - Present*
- Conducting a comprehensive literature review to identify the state-of-the-art techniques in **tutor-learner simulation** using AI and NLP methods: large language models, automatic multiple choice question generation, question validation.
- Implementing an innovative simulation using **Autogen involving a resource-augmented pipeline using LangChain and the OpenAI API** designed to automatically generate high-quality multiple choice questions.
- Designing a robust web application framework to support this simulation using HTML/CSS and JavaScript.

Department of Computer Science. University of Illinois Urbana, IL
Illinois CS Summer Research Program Intern, Mentor: Dr. Kathryn Cunningham *May. 2024 - Present*
- Develop a **web application framework using HTML/CSS, JavaScript, and Flask** to prototype an automatic plan identification system.
- Lead a design workshop and an evaluation study by interviewing instructors across various domains to validate the prototype.
- Administer validated instruments for evaluation, including the **NASA TLX cognitive load survey and customized usability surveys**, to analyze and interpret participant responses.

Department of Computer Science. University of Illinois Urbana, IL
Undergraduate Researcher, Social Computing Laboratory (SCUBA), Mentor: Dr. Eshwar Chandrasekharan Aug. 2023 - Present
- Explore prosocial reinforcement behavior on social media platforms like Reddit to ensure safe and healthy online communities.
- Identify treatment and control conditions on a **large dataset of ¿1,000,000 data points**. Perform **stratified propensity score matching and regression analysis** on the groups to understand the effect of positive reinforcement behavior on user interaction.
- Perform **LDA analysis and k-means clustering analysis on embeddings generated using BERT models** on Reddit data to yield insights into the type of content that receives prosocial engagement.

Department of Computer Science. University of Illinois Urbana, IL
Independent Study (CS 397), Mentor: Dr. Koustuv Saha *Aug. 2023 - Present*
- Understood the disparities between human comments and the outputs generated by LLMs for the same queries about Alzheimers.
- Collected the 50 top questions from the r/Alzheimers subreddit using Python Reddit API Wrapper (PRAW). **Trained classifiers and implemented LIWC analysis** to compare the results generated by large language models with the actual responses on the platforms.
- Performed qualitative analysis on Reddit generated and LLM generated results for the the top 30 posts using **inductive coding** to gain insights into this data.

Department of Computer Science. University of Illinois Urbana, IL
CS Student Ambassador/Research Scholar, Mentor: Dr. Kathryn Cunningham *Jan. 2023 - Present*
- Reviewed the literature to identify applicable techniques to identify code patterns in domains that interest non-majors. Assessed the capabilities of **code similarity algorithms, edit distances, and other systems for mining code idioms** to support pattern identification.
- Architect a prototype that can identify common code sequences from code repositories using the capabilities of **large language models and processing manipulation using clustering and similarity analysis**. Created a dataset by collecting over 20,000 code pieces from Stack Overflow across different Python topics including Pandas, TensorFlow, and MatPlotLib via web-scraping using Beautiful Soup in Python for testing.
- Conceptualize an interface to present these sequences as patterns in a way where they can be edited by instructors.

Department of Computer Science. University of Illinois Remote
Undergraduate Summer Research Intern, Mentor: Dr. Kathryn Cunningham *May. 2023 - Jul. 2023*
- Interviewed 10 educators who identified novel programming patterns to investigate the state of art of pattern identification for instructional purposes, its challenges, and how technology can help make this process more efficient.

- Performed **qualitative coding** on the interview transcripts to thematically analyze the methodology and draw conclusions about how educators find and create programming patterns. Discussed the design of a **theoretical socio-technical system** that can automate this process.

POSTERS, PUBLICATIONS, AND PRESENTATIONS

ACM Conference on Computer-Supported Cooperative Work and Social Computing (Poster) *Jul. 2024*
C. Lambert, **Y. Jain**, K. Saha, E. Chandrasekharan. Investigating How Gilds Were Employed On Reddit. In Proceedings of the 2024 ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '24).

Under Review *Mar. 2024*
K. Saha, **Y. Jain**, C. Liu, S. Kaliappan, R. Karkar. AI vs. Humans for Online Support: Comparing the Language of Responses from LLMs and Online Communities of Alzheimer's Disease.

ACM Conference on International Computing Education Research (Poster) *Jul. 2023*
Y. Jain, K. Cunningham. Towards Methods for Identifying High-Quality Domain-Specific Programming patterns. In Proceedings of the 2023 ACM Conference on International Computing Education Research V.2 (ICER '23 V2). https://doi.org/10.1145/3568812.3603478

PROJECTS

AI@UIUC. University of Illinois
Technical Staff *Jan. 2023 - Present*
- Built a React Native-based minimal viable prototype of the front-end for the mobile application "Be+" for our client; the YouTuber DaxFlame.
- Engineered Be+ to utilize a short personality test and supplementary questionnaires to generate a unique chat-bot, referred to as a "+." Over time, these chat-bots become digital representations of users, learning to mimic their mannerisms and answer questions about them through ongoing conversations.
- Engaged in periodic check-ins with the client for client need analysis and presented project updates.

LMS Tracker
Independent Project *Jun. 2023 - Present*
- Design an Android application to consolidate course information (assignments, course material, due dates, announcements) of all students' courses in a single location. Added a feature to view the average expected grade for the semester based on the chosen courses and their average GPAs.
- Leverage Kotlin for the front-end and back-end development and integrated it with SQLite for database management.

Klimacc
Group Lead, Team Project *Jan. 2021 - Dec. 2021*
- Built a user-friendly weather application that analyzes past data to build a model that predicts future weather conditions.
- Designed the GUI using the Tkinter module and deployed the application using Python for the back-end.

Parivartan
Group Lead, Team Project, Quantahacks *Dec. 2019 - Jan. 2021*
- Engineered a prototype of a smart chip driving license system with a fingerprint recognition feature in the car to abate deaths due to underage driving. Created the minimal viable prototype using Raspberry Pi and Python.
- Awarded the third prize globally in Quantahacks, a hackathon organized by STEM for FEM. Featured among the top 500 of the 7200 innovations presented in the ATL Marathon organized by Niti Aayog, a Government of India think tank to promote research and innovation.

LEADERSHIP AND MENTORSHIP

Reflections Projections 2023. Association of Computing Machinery. University of Illinois Urbana, IL
Content Team Co-Director *Jan. 2023 - Present*
- Recruited keynote speakers, chief guests, educationists, and other sponsors and orchestrated communication with them.
- Organized 10+ tech talks by inviting speakers including Dr. Mark Guzdial from the University of Michigan and Dr. Josh Starmer from StatQuest.

Open Source @ Illinois Urbana, IL
OpenDev Co-Director *Jun. 2023 - Present*
- Supervise four projects in accessibility in CS, web-dev, AI/ML, and ECE, by providing feedback and ensuring goal achievement through periodic meetings.
- Optimize funding, coordinate recruitment, and organize workshops for skill enhancement.

EXPO Engineering Council Urbana, IL
Reverse Career Fair Lead *Jun. 2023 - Present*
- Facilitated communication between over 30 companies and RSOs to organize the *Reverse Career Fair* which is targeted at providing a low-stress setting for RSO leads to interact and form connections with industry representatives.
- Spearheaded logistics and communication with companies and registered students on the day of.

Unmasking Sustainability Delhi, India
Co-Founder *Jun. 2021 - Jan. 2022*
- Founded Unmasking Sustainability, a Non-Profit Organization; employed my school's female support staff who had lost their jobs, coordinated periodic meetings, gathered raw materials, and managed revenues to create a sustainable method of manufacturing masks.
- Sold over 800 masks; generated over 40,000 INR; donated over 200 masks over a period of 1 month.

CERTIFICATIONS

University of Illinois
Completion of Data Science Discovery *May. 2023*
https://d7.cs.illinois.edu/badges/stat107-sp23-chn6ItR904udc9rZktZmi9l1O8xMxP/

CITI Program
Completion of Course in the Protection of Human Subjects *Feb. 2023*
https://www.citiprogram.org/verify/?w7fde721f-bea7-4e56-92aa-713a5de11b1c-54560864

Microsoft
Microsoft Technology Associate: Introduction to Programming Using Python *Jan. 2022*
https://www.credly.com/badges/282e224d-b4a9-4e4a-b937-2e5302f75c5f/public_url

OpenEDG Python Institute
PCEP - Certified Entry-Level Python Programmer *Jul. 2021*
https://www.credly.com/badges/43894b36-ac44-4cf5-8ea6-c72287ff0f29/public_url

SKILLS

- **Programming Languages** - Python, Java, C/C++, SQL, Kotlin, R
- **Frameworks** - Pandas, Scikit-Learn, Numpy, Autogen, API Handling, Tkinter, Android Studio, SQLite, GitHub
- **Languages** - English, Hindi, German