

# SNEHA SUNDAR

[ssundr.github.io](https://ssundr.github.io) | [linkedin.com/in/ssundr](https://www.linkedin.com/in/ssundr) | [408-508-9109](tel:408-508-9109) | [github.com/ssundr](https://github.com/ssundr) | [snehas9@illinois.edu](mailto:snehas9@illinois.edu)

## EDUCATION

**University of Illinois at Urbana-Champaign**  
*Bachelor of Science in Computer Science and Statistics*

May 2027  
GPA: 3.71/4.0

## SKILLS

**Languages:** Python, JavaScript, TypeScript, C++, C, Java, HTML, CSS, SQL, R

**Tools:** React.js, Git/GitHub, Flask, MongoDB, Neo4j, Pandas, OpenCV, PyTorch, Tensorflow, NumPy, Django, Linux

## EXPERIENCE

**University of Houston NSF REU Site | Software Engineering Intern**

May 2025 – Aug 2025

- Built a Python data pipeline using NumPy and OpenCV to preprocess and normalize 55K+ fingerprint images.
- Developed and optimized genetic algorithms to automate spoofing tests, raising success rate from 24% to 100%.
- Created automated benchmarking scripts to compare attack methods and assess system security.
- Designed a GPU-accelerated module to analyze fingerprint matching error rates, supporting debugging and system validation

**Language Interaction Lab | Back-End Developer**

Sep 2024 – Present

- Built scalable Python ETL pipeline (pandas, boto3) to ingest and process 1TB+ of gzip-compressed citations and full-text data from AWS S3, eliminating local storage bottlenecks and accelerating NLP experiments
- Improved citation accuracy for 150K+ academic references using a fine-tuned SciBERT model, automating similarity scoring and saving researchers 100+ hours of manual checks, enhancing reliability of citation tracking
- Implemented interactive dashboards that enhanced citation confidence tracking, improving model evaluation

**University of Illinois at Urbana-Champaign | Course Assistant (Data Science)**

Jan 2024 – Present

- Coordinated Git-based release of containerized Python Jupyter labs via CI/CD pipelines teaching pandas and Matplotlib for 1,000+ students, ensuring consistent version control and minimizing setup issues
- Mentored students in weekly office hours and collaborated with course staff as a team player to teach data loading, transformation, visualization, scikit-learn clustering and classification workflow, and debugging techniques

**Discover Financial Services | Sophomore Spark Participant**

May 2025

- Selected as 1 of 80 sophomores nationwide, gaining Agile (Scrum, sprint planning) and software engineering skills
- Explored MongoDB and AWS in fraud detection demos, applying data analytics to fraud, risk, and financial markets analysis

**Arrcus | Software Engineering Intern**

Jun 2024 – Aug 2024

- Engineered a multi-threaded Python log parser to sort device logs by protocol, reducing debugging time by 30%
- Deployed an automated Python/Postfix pipeline to generate and deliver error summaries to global engineering teams, cutting log distribution time from hours to minutes and improving incident response
- Created Python ETL pipelines (regex, pandas) to extract and structure log metrics for performance monitoring

## PROJECTS

**PickMe: Restaurant Recommender** | *React.js, Flask, Python, TypeScript, MongoDB, Git*

Jan 2024 – May 2024

- Developed full-stack React.js & Flask application to deliver personalized restaurant recommendations
- Leveraged Python KMeans clustering and Google Geolocation & Places APIs for location-aware suggestions
- Implemented Google OAuth and designed MongoDB schemas for user authentication, preferences, and history

**Long Texts Summarizer** | *Python, Flask, BeautifulSoup, NLTK, Jinja2, HTML/CSS, Git*

May 2024

- Engineered a full-stack Flask web app with requests and BeautifulSoup to extract and aggregate Wikipedia content
- Developed responsive front-end with HTML/CSS and Jinja2 templates; integrated NLTK for tokenization, stop-word removal, and extractive summarization with robust validation and error handling

## RELEVANT COURSEWORK

Computer Systems, Software Design Lab, Database Systems, Data Structures, Computational Social Science, Algorithms, Discrete Mathematics, Linear Algebra, Numerical Methods, Statistics and Probability 1&2, Statistical Modeling, Natural Language Processing, Applied Bayesian Analysis, Web Programming, Data Science Discovery