Sanika Chavan

+1 602 516 5693 • sanikac10@gmail.com • LinkedIn • GitHub

EDUCATION

M.S. Computer Science *May 2025*

Arizona State University, Tempe, AZ

SKILLS

Programming Languages: Python, Swift, HTML, CSS, JavaScript, TypeScript, Java, SQL

Tools and technologies: Kafka, Apache Spark, PostgreSQL, NoSQL, Azure, D3.js, React.js, Node.js, Tableau, Power BI, MATLAB, GitHub, Jenkins, Docker, SwiftUI, Bootstrap, Hadoop, Keras, MatPlotLib, Ski-kit Learn, Pandas, TensorFlow, Selenium, Spring boot, REST - APIs

PROFESSIONAL EXPERIENCE

Data Visualization Engineer Intern | *Conected VIT, Mumbai* *March 2022*

- Designed a user interface for Conected VIT's startup website, ensuring adherence to design standards.
- Developed over 5 different comprehensive **visualizations** to show the data and wireframes employing Adobe XD and Figma.
- The meticulously crafted design significantly enhanced user experience, leading to a 1.25x increase in user engagement.

Software Developer Intern | *Vidyalankar Institute of Technology, Mumbai* *August 2021*

- Led the design and implementation of a Student Repository System using Java and MySQL, efficiently managing and monitoring students over 10 different types of extracurricular activities.
- Deployed the Student Repository System, resulting in its widespread adoption by the institute in 3 months
- It led to a 2x reduction in overhead for schools in managing extracurricular activities and monitoring student participation.

TECHNICAL PROJECTS

Data Leakage in LLMs *December 2024*

- Designed and implemented a detection pipeline to identify key data leakage scenarios in large language models, including Seen Question and Answer, Seen Question, Similar Question, and Seen Relevant Information, using different datasets.
- Evaluated five diverse models (e.g., GPT-4o, Claude-3.5, Llama-3.2) to analyze susceptibility to leakage, revealing key insights about model generalization, reasoning, and reliance on training data patterns.
- Developed and validated the Relevant Information module, leveraging techniques like Part of Speech tagging and entailment checks to detect data leakage and highlight the need for robust training processes and dynamic evaluation benchmarks.

KAN You See the Bias? A visualization of Comparative Study of MLPs vs. KANs *December 2024*

- Faced the challenge of bias in machine learning models impacting human-centric decisions. Built a **storytelling visualization** using ADULT and WESAD datasets, integrating **heatmaps** and **bias distribution plots** to communicate these effects. Improved user engagement by 20%, effectively illustrating the real-world consequences of biased models.
- Analysed and compared Multi-Layer Perceptron (MLP) and Knowledge-Aware Network (KAN) architectures, demonstrating a 10% improvement in generalization and reduced bias with KANs. Visualized performance and fairness metrics across sensitive attributes, providing insights for building fairer, GDPR and HIPAA-compliant ML models.

WebDirect *March 2023*

- Developed a web portal using ReactJS and MongoDB, featuring a content-based recommendation system powered by Machine Learning and cosine similarity algorithms, providing an integrated platform for students to showcase their projects and receive relevant resource suggestions.
- Initiated with the Smart India Hackathon, where the project was selected as one of the top 10 out of 100 submissions.

PUBLICATIONS

WebDirect: A Web-based Project Repository | *IEEE Xplore Publication* *May 2023*

- The study presents a web-based project repository system using ReactJS, MongoDB, and cosine similarity for recommendations, addressing issues like complex UI and weak search filters.
- It advances web management and recommendation systems, with potential for adding features like author collaboration and adapting to platforms such as Android and iOS.

A Web-based Project Repository using ReactJS | *IEEE Xplore Publication* *October 2022*

- The paper tackles common issues with project portals, introducing enhancements like improved search filters, project recommendations, trending sections, and user collaboration features.
- It advances the study of project portals, providing insights into functionality and usability, with potential for future adaptation to new technologies and business needs.

ACADEMIC EXPERIENCE

Research Assistant, Arizona State University *December 2024*

- Researched and developed spatial query algorithms, including ST_Contains and ST_Within, using **Apache Spark** and **SparkSQL** to optimize large-scale data processing workflows.
- Conducted performance evaluation and benchmarking for spatial query solutions, implementing Hot Zone Analysis to identify high-density regions in large datasets.