

SAKSHEE PATIL

(919) 453-8678 | saksheepatil05@gmail.com | [linkedin.com/in/saksheep](https://www.linkedin.com/in/saksheep) | sakshee5.github.io | github.com/Sakshee5

AI Engineer with 2+ years in Generative AI solutions at Deloitte, specializing in multi-agent LLM systems and RAG. Currently, pursuing a masters with a focus on Explainable AI for interpretable systems. Eager to apply AI expertise to drive innovation.

EDUCATION

M.Eng - Artificial Intelligence, Duke University, Durham, NC, USA **August 2024 - December 2025**

- GPA: 4.0, Relevant coursework: Explainable AI, Modeling Process & Algorithms, Deep Learning, LLMs
- Research Assistant for CREATE: Built Qubit - AI powered coding assistant to help high school students learn coding.
- Won the "First Grand Prize" and "Best Use of Agents" special prize in the DukeAI Hackathon amongst 500+ participants. Built Inception-Swarm, a multi-agent system capable of building multi-agent systems based on natural language prompts.
- Won 1st place in class Kaggle competition (data cleaning, feature engineering and classification) amongst 40 students.
- Published an article and code demo on Adversarial Attacks on LLMs (saksheepatil05.medium.com)

B.Tech - Mechanical Engineering, Indian Institute of Technology Indore, MP, India **August 2018 - August 2022**

- GPA: 3.47, Relevant coursework: Calculus, Linear Algebra and ODE's, Numerical Methods, Computer Programming
- Won silver medal in Inter-IIT Technical Meet, 2021: Led a 10-member team as part of the AI-ML Club.
- Published a [book chapter](#) titled "Deep learning-based methods for detecting surface defects in steel plates" in Elsevier.

SKILLS

Programming Languages and Frameworks: Python, HTML/CSS, SQL, NumPy, Pandas

Machine Learning: Transformers, Generative AI, PyTorch, HuggingFace, LLM, NLP, Deep Learning, Neural Networks, Recommendation Systems, Fine Tuning, RAG, LangChain, TensorFlow, Keras, Scikit-learn, XGBoost, OpenCV, LlamaIndex

PROFESSIONAL EXPERIENCE

AI Analyst - GenAI Tech, Deloitte, Mumbai, MH, India **August 2022 - August 2024**

- Prototyped and tested an end-to-end multi-agent LLM system utilizing OpenAI's Assistants API within a 10-person team, enabling autonomous web research, analysis, and deliverable generation (newsletters, reports, presentations). Generated 100+ first draft deliverables during the testing phase.
- Built and deployed a prototype (streamlit + GenAI) utilizing Retrieval Augmented Generation (RAG), Chain of Thought Prompting and Powerpoint XML for creating customizable and downloadable PowerPoint presentations, estimated to save 30% of consultant time in creating first drafts.
- Built a scalable topic modeling approach for advanced document analysis; adopted by 10+ teams internally.
- Fine-tuned a model to generate SVG code for Process Flow Diagrams by retrieving and adapting similar diagrams from a repository, allowing quick customization based on user requirements.

Summer Research Intern, National University of Singapore, Remote **April 2021 - August 2021**

- Developed custom OpenAI Gym environment with Pygame to test deep reinforcement learning algorithms, improving path planning and exploration strategies in partially interactive environments.
- Optimized the ImVoteNet repository usability, enabling streamlined object detection on point clouds and reducing setup time by eliminating additional data downloads. (saksheepatil05.medium.com)

SELECTED PROJECTS

Art Meets AI: Neural Style Transfer with Interactive Visualizations, XAI, Semester Project **August 2024 - Present**

- Built and deployed a web application that visualizes Neural Style Transfer (NST) in real-time, allowing users to interactively blend content and style images by adjusting parameters such as content/style weights and layer selections.
- Enhanced interpretability by integrating explainability techniques (XAI) like saliency maps and Grad-CAM, providing insights into how neural networks process and manipulate visual content during NST.

Alphanumeric Audio Dataset Collection, Sourcing Data for Analytics, Semester Project **August 2024 - Present**

- Completed end-to-end data collection process, from Institute Review Board (IRB) approval to open-source release, aimed at improving speech recognition for non-native English speakers.
- Designed and implemented a data collection protocol using Qualtrics, gathering over 500 audio samples from a diverse participant pool, along with comprehensive demographic metadata.
- Ensured high ethical standards in data collection, securing participant consent and ensuring anonymization throughout the process. Open-sourced the dataset, contributing to AI fairness by promoting inclusivity in voice recognition systems.

Unsupervised Segmentation of Peck-damaged Rice Grains **January 2022 - June 2022**

- Earned highest grade for conducting research in a field with limited existing literature, utilizing a novel dataset.
- Developed an unsupervised Convolutional Neural Network (CNN) for segmenting peck damage in rice grains, achieving a 92.54 Dice coefficient accuracy on the validation dataset.
- Further sub-classified the masks using K-means clustering, based on metrics such as damage scatter, area, and centroid positioning, potentially uncovering the underlying causes of peck damage across different classes.

Design and Vision-based Control of Miniature Rolling Capsule **April 2021 - January 2022**

- Designed & 3D printed a CAD model of a miniature rolling capsule fitted with a CCD camera for pipe inspection.
- Developed a modified U-net architecture for detecting the type, spread, and severity of defects on steel surfaces, achieving an accuracy 2% higher than the baseline results on the dataset.