

Shivam Mittal

Madison, WI | smittal39@wisc.edu | (+1) 559-213-5557 | [Linkedin](#) | [Github](#)

Education

University of Wisconsin- Madison Madison, WI Master of Science (M.S.)	Aug 2025 - May 2027
Major in Computer Science, specializing in AI, Natural language processing and Generative AI	
<u>Coursework</u> - Advanced Natural language processing, Foundation models, Advanced Machine learning, Operating Systems	
M.S Ramaiah Institute of Technology Bengaluru, India Bachelor of Engineering (B.E.)	Aug 2016 - Aug 2020
Major in Electrical and Electronics Engineering	CGPA - 8.79/10

Work Experience

Graduate Teaching Assistant - Computer Science, Part-time Madison, WI, USA	Aug 2025 - Present
<ul style="list-style-type: none">Assist in teaching CS540 - Introduction to Artificial Intelligence, focusing on Deep learning, Natural language processing (NLP), and ProbabilityCreate homework assignments, hold office hours, support students with concepts related to Machine learning and AI.	
ZS Associates - Senior AI Software Engineer Bengaluru, India	June 2024 - Aug 2025
<ul style="list-style-type: none">Designed and deployed production-ready Generative AI multimodal agents with advanced text and vision capabilities leveraging Retrieval-Augmented Generation (RAG) improving response accuracy by 30%.Built GraphRAG, a knowledge graph algorithm for document QA, achieving 20% higher accuracy than RAG agents.Led multimodal, scalable Multi-Agent AI SaaS applications deployed across 3+ industries with thousands of users.Specialized in deploying local LLMs with sub-second latency, 100% data privacy compliance for client-facing projects.	
Iron Mountain - Senior AI Software Engineer Bengaluru, India	Oct 2022 - May-2024
<ul style="list-style-type: none">Developed production-ready Intelligent Document Processing AI solutions achieving 99% data extraction accuracy across diverse industries using Natural language processing and Computer vision.Engineered synthetic signature generation and developed a generalized signature detection model with 98% accuracy.Fine-tuned large language and multimodal models on domain-specific documents, improving performance by 15%.Developed production ready AI SaaS applications for document question answering, deployed to 10+ clients.	
Tensorlot - Machine Learning Software Engineer Bengaluru, India	Oct 2021 - Oct-2022
<ul style="list-style-type: none">Developed time series forecasting models achieving 97% accuracy to predict employee attrition rates.Created a customized Alexa skill via the Alexa Developer Console to enhance customer purchase journeys.Developed a deep neural network, achieving 98% accuracy in detecting 15 distinct wall paint defects.Implemented predictive maintenance models for anomaly detection in machinery and sensors.Engineered an end-to-end A/B testing MLOps CI/CD pipeline on AWS, adaptable for machine learning use cases.Developed a customer journey tracking solution leveraging computer vision algorithms.	
ABB - Management Trainee (Computer vision researcher) Bengaluru, India	Jan 2021 - Sept-2021
<ul style="list-style-type: none">Developed AI vision inspection for robotics manufacturing lines, boosting defect detection and reducing manual labor.	
Vijna Labs - Artificial Intelligence Intern Bengaluru, India	July 2018 - Aug-2018
<ul style="list-style-type: none">Worked on real time object detection, image processing, computer vision and training neural networks.	

Projects & Achievements

U.S Patent - Associated with Iron Mountain

Topic: Generative AI Agent for Intelligent Document Processing | (Filed, Pending with USPTO) | Inventor

- Developed a Generative AI Agent for intelligent document processing with a use-case driven routing mechanism.

Undergraduate Project - Autonomous vehicle | M.S Ramaiah Institute of Technology | Bengaluru, India

- Built a self-driving electric vehicle using Deep Learning and NVIDIA Jetson Nano with 90% accuracy on lane detection.

Student Organization - Team Co-Lead | Stier Racing | M.S Ramaiah Institute of Technology | Bengaluru, India

- Led a team to design and develop a Formula Electric Vehicle. Achieved first place in Formula racing competitions.

Awards - Top Performer of the Year, Project Champion, InGenius Awards, and Employee of the Month

Hackathon - Winner of the Iron Mountain AI hackathon, led a team of 3 members, awarded best engineering design

Technical Skills

Data structures, Algorithms, System design, Distributed Systems, Microservices, Object oriented programming, SOLID principle

Programming Languages & Frameworks: Python, C++, MySQL, Cypher, API design, Git, FastAPI, ApacheAGE, GPU's

Core AI/ML/Deep learning/Vision/NLP: Neural networks, PyTorch, JAX, SKLearn, XGBoost, YOLO, OpenCV, Stable Diffusion

Generative & Agentic AI: Foundation models, Vision-language models, LangChain, LangGraph, HuggingFace, Transformers, LLM finetuning, Agentic Orchestration, Information Retrieval, LLM tool utilization, Multi-turn Conversationality, LLMOps

Deployment & MLOps: AWS services, AWS SageMaker, GCP, Azure, Docker, Kubernetes, Gitlab, vLLM, Langfuse, LiteLLM