Pradyumn Pathak

LinkedIn +1-8726642384 pradyumnp508@gmail.com Chicago, IL GitHub

bhadwa

SKILLS

- **AI/ML:** Agentic AI, Computer Vision, Deep Learning, Generative AI, LLMs, Machine Learning, Multi-GPU training, NLP, Transfer learning
- **Tools:** Amazon SageMaker, ArangoDB, Azure AI, BERT, C#, CUDA, Docker, FastAPI, GitHub, Google Cloud Platform, Hugging Face, LangChain, LLAMA, MediaPipe, Microsoft SQL Server, Python, PyTorch, OpenCV, OpenVINO, TensorFlow, Transformer
- **Soft Skills:** Agile development, Collaboration, Cross-functional teamwork, Decision making, Leadership, Problem solving, Project management,

EDUCATION

DePaul University – CDM, *Master of Science in Artificial Intelligence / CGPA 3.4* Sep 2023 – Jun 2025 | Chicago, IL

University of Mumbai, *Bachelor of Engineering in Electronics / CGPA 3.2* Aug 2016 – Nov 2020 | Mumbai, IN

PROJECTS

Agentic-AI based Resume Tailoring System, *Cognijin LLC* Jul 2025 – Present | Chicago, IL

- Architected and engineered "Pepper," a multi-agent, agentic AI system that automates resume tailoring using Google Gemini 2.5 Pro APIs, Fast-API, SQL Server, and LangChain, transforming a manual process into a recruiter-optimized, job-specific resume with 91%+ reduction in effort.
- Engineered persistent, scalable backend workflows integrating Gemini 2.5 Pro API, GCP, LangChain, and SQL Server, supporting real-time document analysis, cross-agent communication, and seamless cloud deployment via FastAPI and ngrok for instant user access.

RESEARCH EXPERIENCE

Gen-AI Research Assistant, *DePaul University* Jan 2024 – Jul 2025 | Chicago, IL

- Engineered domain-specific scientific LLMs for materials science by fine-tuning BERT-based LLMs, achieving a ~10% accuracy improvement and supporting active learning on temporal datasets from 1990 to 2024, enabling automated hypothesis generation research paper corpus.
- Designed and developed hierarchical context tree algorithms generating 300,000–500,000 semantically enriched definition sentences using LLMs like Microsoft PHY 3.5, LLAMA, producing a large-scale semantic knowledge graph that enhanced semantic drift resistance by up to 26% over traditional embedding methods, improving robustness in scientific knowledge representation automated discovery.
- Architected multi-GPU training and inference pipelines optimized for large-scale masked language modeling and hierarchical embedding extraction, realizing a 60% speedup in training and up to a 75% acceleration in embedding inference for efficient processing of scientific data.

PROFESSIONAL EXPERIENCE

Gen-AI Research Intern, *FPT - Americas* Jun 2024 – Aug 2024 | Austin, Tx

- Spearheaded end-to-end development of a real-time gesture and voice-controlled flowchart creation prototype using MediaPipe for hand gesture recognition and Azure Speech Services for speech-to-text commands, enabling intuitive, hands-free manipulation of flowchart elements.
- Designed and trained multiple YOLOv8 object detection models on augmented, multi-source datasets to identify 9 distinct flowchart components and arrow types with up to 0.81 mAP50 accuracy, improving shape detection robustness across hand-drawn, whiteboard, and digital flowcharts.
- Developed an automated pipeline that converts detected flowchart components into vector-based PowerPoint presentations using YOLOv8 and Azure AI Vision OCR for text recognition, reducing manual digitization time to ~1.1 seconds per flowchart and enabling seamless collaboration in meetings.
- Collaborated directly with clients to define project requirements and laid the ground word for a pioneering multi-stage AI neural style transfer pipeline for replicating culturally significant Vietnamese mosaic art, supporting digital preservation of heritage artworks in educational settings.

AI and Deep Learning Engineer, *Cynapto Tech. Pvt. Ltd.* Aug 2021 – Aug 2023 | Remote

- Architected and developed an end-to-end AI tooling platform integrating SQL Server, FastAPI, and object detection frameworks like YOLO, Detectron2 using TensorFlow and PyTorch, streamlining model training, evaluation, and deployment workflows, reducing annotation and development time by approximately 60%, resulting in estimated cost savings of around $10,000 per month and accelerated project delivery.
- Led generative AI research and development of a multilingual lip-syncing prototype for the film industry, delivered an MVP that achieved 85% lip regeneration quality in AI generated video content resulting in a client investment contract for further development.
- Accelerated AI model deployment cycles from typical industry timelines of 1 month down to 2 weeks by leveraging transfer learning and modular fine-tuning pipelines for rapid client-specific adaptation enabling client solutions delivery even with lean teams and agile, startup-paced timelines.
- Directed a medical image segmentation project with UNet models to detect leg defects in femur and tibia X-ray images, achieving a Dice coefficient of 0.85; improved diagnosis speed by 30% for bulk physiotherapy bottlenecks in government hospital collaborating with NGO.
- Implemented an end-to-end security AI solution for automated perimeter monitoring, by deploying object detection models on simple CCTV feeds, enabling cost savings estimated at $50,000 annually per client by decreasing reliance on manual security patrols and improved safety.

AI Intern, *Cynapto Tech. Pvt. Ltd.* Jan 2021 – Jul 2021 | Remote

- Contributed to fine-tuning and evaluating face recognition classifiers on a 10,000-image dataset, raising accuracy from 90% to 97% for a major construction industry client, ensuring higher precision in employee identification systems in real-time environment.
- Collaborated with teammates to automate benchmarking of computer vision models deployed on edge devices like Raspberry Pi 4, NVIDIA Jetson Nano, reducing inference latency to under 30ms per image and improving execution speed by at least 30% compared to baseline.
- Engineered end-to-end data preprocessing and augmentation pipelines using Python, TensorFlow, PyTorch, and OpenVINO, managing large datasets and reducing label noise to enhance model generalizability and robustness across diverse edge inference scenarios.

CERTIFICATIONS AND ACHIEVEMENTS

- Inducted as a lifetime member of Upsilon Pi Epsilon, the International Honor Society for Computing and Information (UPE). (2025)
- Represented DePaul by showcasing my generative AI research at ICSSI, highlighting innovations in AI-driven scientific discovery. (2024)
- Certified in Machine Learning and Neural Networks by Stanford University and DeepLearning.AI. (2019)