PRANAV RAJESH CHARAKONDALA

Champaign, IL | +14479024780 | prc4@illinois.edu | LinkedIn: Pranav CR | GitHub: PranavCR01

EDUCATION

University of Illinois Urbana – Champaign — **Champaign, USA**
Master of Science Information Management, **GPA 3.93/4** — **May 2026**
Coursework: Applied Machine Learning, Human-Centered Data Science, Text Mining, Machine Learning for Cloud

PES University — **Bengaluru, India**
Bachelor of Technology Electronics and Communications Engineering, **GPA 7.72/10** — **May 2022**
Coursework: Artificial Neural Networks, Digital Image Processing, Programming with Python & C, Data Structures & Algorithms

TECHNICAL SKILLS

Languages: Python, SQL, C, Java, JavaScript, HTML, CSS, Apex, LWC, Node.js, R Programming
Tools: Power BI, Tableau, Jira, Git, GitHub, Microsoft Excel, Microsoft Word, Neo4j, Dataloader, Copado (CI/CD), Splunk
Libraries & Frameworks: PyTorch, Keras, TensorFlow, NumPy, Scikit-Learn, Pandas, Matplotlib, OpenCV
Technical Skills: Data Science, Machine Learning, Large Language Models, Computer Vision, Generative AI, Natural Language Processing, Agile Methodologies, Test-Driven Development, Full Stack Development, RESTful API, Cloud Technologies, DevOps, Salesforce Custom Development

PROFESSIONAL EXPERIENCE

CENTER FOR HEALTH INFORMATICS, UIUC – Research Lead, Applied AI — **May 2025 – Present**

- Led a 12-member cross-functional team to deliver a production-oriented multimodal GenAI system for health misinformation analysis in short-form video, partnering with the WHO to align system behavior with real-world operational and policy constraints across 120+ clips
- Built an end-to-end ASR + OCR + LLM pipeline (Whisper, EasyOCR, GPT-4, Mistral, LLaMA) that routes audio, visual text, and metadata through conditional processing paths, extracting 5K–20K OCR text snippets per clip with ~5–25s end-to-end inference latency
- Implemented a decision layer that generates structured rationales and 1–3 evidence snippets per clip, with optional grounding via a PubMed MCP server, enabling the system to select, justify, and validate outputs autonomously with 100% explanation coverage and >95% JSON parsing
- Designed an experimentation and evaluation framework to benchmark prompting strategies across multiple LLM providers, tracking latency, confidence calibration, and failure modes to support fast iteration and reliable deployment

DELOITTE CONSULTING, US OFFICES OF INDIA – Analyst — **June 2022 – July 2024**

- Implemented 50+ custom Salesforce FSI Cloud features for a major U.S. wealth management firm enhancing backend performance using APEX (Java) and frontend compliance with LWC (JavaScript), optimizing queries by 20%, and cutting page load time to 2 seconds
- Developed and integrated real-time workflows using Salesforce APIs, APEX, Visualforce, LWC, and Platform Events – including the "Interaction Notes" feature, commended by the client and deployed across 12,000+ client branches – defining custom objects and validation rules, streamlining cross-system communication, and reducing data lag by 40%
- Collaborated with stakeholders to refine 50+ user stories and engineered business-wide Salesforce workflows using APEX, async callouts, and OmniStudio, boosting validation accuracy, reducing manual edits by 35%, and cutting change requests by 25%
- Refactored legacy APEX logic across 20+ backend modules – revising system logic and modularizing business rules to reduce tech debt, while leading Salesforce data archival strategy using soft deletes with retained DB2 records to enhance reusability and ensure long-term retrieval
- Co-led resolution of 200+ Salesforce–DB2 data discrepancies as part of a cross-system reconciliation effort – built anomaly detection workflows using Excel, SOQL, and Snowflake to automate variance tracking and cut manual investigation time by 40%

DRONGO AI – Machine Learning Intern — **January 2022 – June 2022**

- Led R&D in medical image segmentation developing UNet (Keras) and PraNet (TensorFlow) models for osteoporosis and caries detection using X-rays, applying preprocessing, and noise reduction to improve dataset quality and increase Dice coefficient by 10%
- Optimized PyTorch-based model training pipelines using CUDA and NumPy, leveraging GPU acceleration to reduce training time and improve inference efficiency for large-scale image datasets
- Developed a real-time number plate detection system using YOLOv4 and PP-OCR, achieving 95% OCR accuracy at 30 FPS, with Pandas used for result aggregation and performance analysis
- End-to-end research and implementation of PP-OCR significantly enhanced the firm's RT-ANPR solution, with deployment via Docker on AWS cloud infrastructure, enabling scalable and efficient license plate recognition

PROJECTS

Safe RAG Mental Health Copilot – Human-Centered AI — **Champaign, USA – December 2025**

- Built a safety-first RAG-based conversational agent using FAISS vector-based semantic retrieval and session-aware memory for exam anxiety
- Integrated vetted clinical sources (WHO, CDC, APA) into a RAG pipeline with citation enforcement, safety guardrails, and abstention logic, achieving 100% citation-backed outputs, >95% structured validity, and zero unsafe responses across 40+ high-risk interactions

AIDE-OSS – AI Cybersecurity Assistant for Log Intelligence — **Champaign, USA – May 2025**

- Implemented a stateful, retrieval-augmented agent for Zeek network logs using FAISS-based semantic search and Mistral-driven reasoning
- Designed multi-step agent workflows that query, filter, correlate, summarize log events across TCP/IP, time windows, and source IP dimensions using Isolation Forest and Drain3, enabling automated investigation loops with session-aware context instead of single-turn log queries

AWARDS

- **"Defining Debutant" & 2x Spot Award** – Best performing fresher and best performer in the project (Deloitte)