

Pranav Suresh Magadi

Applied AI / ML Systems

Chicago, IL | psureshmagadi@gmail.com | 312-721-2956 | [LinkedIn](#)

Lead Data Scientist with 6+ years in data science across healthcare and insurance. Builds production AI systems for clinical evidence extraction, retrieval, and document review, combining hands-on engineering with architecture, evaluation, and delivery. Chairs HCSC's Model Approval Committee, guiding enterprise technical review and evaluation practices for generative and agentic AI.

PROFESSIONAL EXPERIENCE

Health Care Service Corporation | Chicago, IL

Lead Data Scientist

Jun 2025 - Present

MedExtract | Clinical evidence extraction

- Originated MedExtract and led its architecture and development as the sole data scientist, partnering with an engineer on frontend and deployment; delivered a reusable, instruction-driven platform to **production for one team**.
- Built a LangGraph workflow with a ReAct extraction agent, document mapping, and task planning. Combined structure-guided retrieval, regex, and LLM extraction with source matching to produce summaries, evidence tables, and highlighted PDFs.
- Implemented offline LLM judges for plan quality, plan adherence, and extraction quality/completeness, alongside pilot feedback on evidence correctness, omissions, and miss severity.
- Implemented MLflow tracing, token/cost/latency diagnostics, and retry/error handling; partnered with engineering on caching intermediate and final results and the associated admin controls.
- Reduced processing time for **one observed 130-page document from over 2.5 minutes to about 75 seconds** through parallel OCR, request decomposition concurrent with OCR, and shorter LLM tool arguments.

AI Risk Adjustment | Chart screening and coder assistance

- Led architecture and production delivery of a shared retrieval-augmented extraction pipeline using Azure OpenAI and Azure AI Search, preserving diagnoses, visit dates, and source locations for screening and evidence-based coder assistance.
- Second-pass vendors confirmed additional risk-adjustable diagnoses in **25.1% of approximately 2,200 flagged records**; validated screening findings contributed **approximately \$1M in realized incremental reimbursement for benefit year 2025**.
- Designed and executed evaluation across vendor-data proxies, direct coder validation, retrieval relevance/ranking, and member-level fairness analysis; tuned prompts, retrieval, and custom chunking while prioritizing recall for human review.

Chair of the Model Approval Committee

Mar 2025 - Present

Concurrent enterprise leadership responsibility

- Redesigned technical model review with the co-chair, tailoring review depth to Responsible AI Committee risk assessments and introducing focused documentation and offline review pathways.
- Authored evaluation guidance for generative AI, RAG, and agentic systems; prepared data scientist reviewers through case walkthroughs and coached teams through office hours and direct evaluation support.

PROFESSIONAL EXPERIENCE CONTINUED

Health Care Service Corporation | Chicago, IL

Sr. Data Scientist

May 2022 - Jun 2025

- Led Medical Record Search architecture and development, scaling overnight batch and on-demand workflows to **thousands of documents daily across 8+ teams**; an A/B time study found **up to 40% savings in medical-record review time**.
- Developed diabetes-measure models incorporated into the HEDIS reviewer tool, using Clinical BERT and an LSTM with distributed training across **eight GPUs**. Resolved PyTorch Lightning training issues, trained annotators, and evaluated complete/partial matches with recall-weighted F2.
- Engineered a reusable RAG evaluation framework for relevance and faithfulness benchmarking; built the initial risk-adjustment LLM foundation later productionized as Lead Data Scientist.

Insurance Services Office, Inc. (Verisk Analytics) | Jersey City, NJ / Boston, MA

Data Scientist I

Feb 2020 - May 2022

- Built BERT classifiers for legal case documents, supplying liability categories to simulation models for insurance risk estimation and loss-distribution analysis.
- Developed GLMs and machine-learning models for property and casualty loss-cost forecasting and pricing; constructed multi-year projections for Lloyd's Realistic Disaster Scenarios to support loss forecasts and reserve capital planning.

Onco-care Analytics LLC | Chicago, IL

Data Scientist Intern

Jun 2019 - Dec 2019

- Built data pipelines, interactive analytics, and Monte Carlo cost simulations for oncology providers in value-based care, supporting cost analysis and risk-sharing strategies.

University of Chicago, Center for Translational Data Science | Chicago, IL

Data Science Intern

Jul 2019 - Dec 2019

- Developed a Random Forest model with class-imbalance mitigation to predict fire events using NASA GOES-R geospatial and satellite imagery data.

Accenture | Bengaluru, India

Application Development Associate

Nov 2016 - Jun 2018

- Developed service-oriented web integrations for orders, invoices, receipts, and banking data across ERP systems and databases; contributed to service architecture and cross-system data flows.

TECHNICAL SKILLS

AI and retrieval: Python, SQL, LangGraph, LangChain, Azure OpenAI, Azure AI Search, Azure Document Intelligence, RAG, hybrid search, structured extraction

Evaluation: LLM-as-a-Judge, judge calibration, human review, precision/recall/F2, nDCG, precision@k, MRR, relevance, faithfulness, fairness analysis

ML and engineering: PyTorch, PyTorch Lightning, Clinical BERT, MLflow, Databricks, Elasticsearch, PostgreSQL, FastAPI, Flask, GitHub Copilot

EDUCATION

University of Illinois Chicago | MS, Management Information Systems | Dec 2019 | GPA 3.9

BNM Institute of Technology | BE, Computer Science Engineering | Jun 2016