

AVINASH PANDEY

aopandey24@gmail.com | Boston, MA | <https://www.linkedin.com/in/avinashopandey>

EDUCATION

Northeastern University

Expected May 2028

Master of Science in Artificial Intelligence, Incoming Fall 2026

Planned Coursework: Foundations of Artificial Intelligence, Linear Algebra and Probability for Data Science

Purdue University

December 2024

Bachelor of Science in Computer Science, GPA: 3.3/4.0

Coursework: Database Systems, Data Science, Introduction to Data Analysis and Mining

TECHNICAL SKILLS

Languages: Python, SQL, Java, R

AI/ML: PyTorch, scikit-learn, XGBoost, Hugging Face Transformers, NumPy, Pandas

LLM & RAG Systems: LangChain, OpenAI API, Gemini API, Claude API, FAISS, ChromaDB

Data & MLOps: PySpark, Docker, Kubernetes, AWS EC2, AWS S3, AWS Lambda, MLflow, Git/GitHub

AI Development Tools: Cursor, Claude Code, GitHub Copilot

PROFESSIONAL EXPERIENCE

Purdue University

Feb 2025 – Feb 2026

Research Assistant

Indianapolis, IN

- Designed a multi-LLM summarization pipeline using Google Gemini 2.0 Flash and GPT-4o-mini to generate personalized biomedical literature summaries, achieving BERTScore-F1 of 0.86 across 20 biomedical queries.
- Built an asynchronous retrieval and metadata-enrichment workflow using PubMed Entrez and NIH iCite to collect article metadata, citation signals, and relevance features across biomedical literature.
- Developed a citation-aware ranking approach combining ontology-guided keyword extraction, cosine similarity, and Relative Citation Ratio (RCR) to surface high-value papers for downstream summarization.
- Implemented multi-model aggregation and validation steps to compare generated summaries, reduce unsupported claims, and improve reliability for user-facing literature review workflows.
- Evaluated the system with ROUGE, BERTScore, ablation studies, and a 10-user study; documented methods and findings in a first-author biomedical AI preprint.

Legislative Services Agency

Dec 2023 – Mar 2024

Business Analyst Intern

Indianapolis, IN

- Designed and managed ETL pipelines to validate and preprocess Indiana General Assembly datasets, achieving 98% accuracy across legislative bill and report datasets.
- Developed and optimized SQL-based workflows, improving data extraction, transformation, and loading efficiency within SQL Server.
- Collaborated in a cross-functional team with Business Analysts and Software Developers to enhance business intelligence workflows, leveraging Tableau and Power BI for improved data reporting and analysis.
- Provided technical support to Indiana Lawmakers and troubleshooted for data integration issues, ensuring seamless access to structured legislative data sources.

PROJECTS

ICML 2024 Research Paper Analyzer | *Personal Project, LLM/RAG Pipeline*

Jul 2024 – Nov 2024

- Built an end-to-end AI pipeline to collect, parse, retrieve, summarize, and analyze 2,600+ ICML 2024 research papers using PDF extraction, vector search, and LLM-based summarization.
- Implemented retrieval workflows with FAISS and ChromaDB, enabling paper-level question answering and topic exploration over a large research corpus.
- Used Gemini 1.5 Flash and LangChain metadata filtering to classify research areas, extract structured paper insights, and support targeted analysis by topic, author, and institution.
- Developed dashboard visualizations to surface research trends, top contributors, and major topic areas for non-technical users exploring AI research and startup opportunities.

LEADERSHIP & AWARDS

- Dean's List 5 terms, International Jaguar Excellence Award, ALDPES Honors Society.
- Led orientation of 800+ students fostering academic success and community engagement as a Team Leader.
- Doubled CS club engagement by organizing hackathons & securing sponsorships as the Vice-President.