

AVINASH PANDEY

aopandey24@gmail.com | 317-998-3501 | Indianapolis, IN | <https://www.linkedin.com/in/avinashopandey>

TECHNICAL SKILLS

Scripting Languages and Packages: Python, R, SQL, PyTorch, Scikit-learn, XGBoost, Pandas, NumPy, SciPy

AI-Development Frameworks and Tools: Databricks, TensorFlow, MLflow, Spark MLlib, Git, GitHub

AI & Data Science Applications: Predictive Analytics, Time Series Forecasting, Anomaly Detection, Customer Segmentation, NLP (Text Classification, Sentiment Analysis, Named Entity Recognition, Topic Modeling)

Deep Learning Research: Multi-Model Learning, Transfer Learning, Representation Learning, Human-in-the-Loop Learning, Model Optimization

PROFESSIONAL EXPERIENCE

Purdue University

Feb 2025 – Present

Research Assistant

Indianapolis, IN

- Design and implement a robust multi-LLM ensemble pipeline integrating Google Gemini 2.0, GPT-4, Anthropic Claude, and BioGPT to generate concise, fact-checked summaries of biomedical literature.
- Develop an asynchronous pipeline using PubMed's Entrez API to retrieve comprehensive metadata from biomedical literature—scalable to tens of thousands across a database of 35+ million citations.
- Engineer dynamic, user-centric relevance ranking algorithms by integrating interactive questionnaires to personalize literature synthesis and improve information retrieval.
- Implement cross-model error correction to reduce hallucinations and enhance summary accuracy.
- Author detailed documentation and maintained rigorous evaluation metrics throughout development, paving the way for a paper submission and future grant proposals.

Aider Ventures

July 2024 – Nov 2024

Machine Learning Engineering Intern

Indianapolis, IN

- Designed and implemented automated data pipelines utilizing FAISS embeddings and ChromaDB, enabling efficient data processing and retrieval for summarizing 2,600+ research papers from ICML 2024.
- Developed and optimized data classification models leveraging the Gemini API, improving the categorization and structuring of research data, leading to a 40% increase in system efficiency.
- Applied metadata filtering algorithms in LangChain to enhance data analysis accuracy, increasing precision of extracted research insights by 20%.
- Built an interactive dashboard for exploratory data analysis, providing trend analysis on 10,000+ papers, including contributions and key research areas for non-technical professionals to anticipate potential startup opportunities.

Legislative Services Agency

Dec 2023 – Mar 2024

Business Analyst Intern

Indianapolis, IN

- Conducted data analysis and validation on Indiana General Assembly datasets, ensuring 98% accuracy of legislative bill and its report through data verification processes.
- Developed and optimized SQL-based queries to extract and transform large legislative datasets, enhancing data accessibility and analytical reporting.
- Utilized Tableau and Power BI to create visual reports and dashboards, enabling a cross-functional team to derive actionable insights from complex legislative data.
- Provided data-driven insights to agency staff, optimizing decision-making processes by refining data reporting and analytical tools.

PROJECTS

AI-Driven Diabetes Prediction Pipeline | Python, Scikit-learn, Deep Learning

Aug 2024 – Nov 2024

- Built a machine learning pipeline using Random Forest, SVM, and Neural Networks for predicting diabetes.
- Enhanced model performance by 15% using batch normalization, L2 regularization, and SMOTE for class imbalance.

EDUCATION

Purdue University

December 2024

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

Coursework: Data Science, Artificial Intelligence, Introduction to Statistics

LEADERSHIP & AWARDS

- Dean's List x5, 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.