Avinash Pandey

aopandey@purdue.edu | (317) 998-3501 | Indianapolis, IN 46204 | linkedin.com/in/avinashopandey/ | github.com/Aopandey | avinashpandey.streamlit.app/

Education

Bachelor of Science in Computer Science Anticipated Graduation: December 2024
Purdue University, IN
Minor: Mathematics
Cumulative GPA – 3.2

Student & Professional Organization

- **Team Leader**, Boiler Gold Rush (BGR) April 2024 – August 2024
  - Facilitated the orientation of 800 new students to university life, promoting academic success and personal growth
- **Vice President**, Computer Science Club (CS Club) May 2023 – May 2024
  - Revitalized the CS Club after a 2-year hiatus, led a team in organizing tech events, increasing club engagement by 100%
  - Coordinated hackathons, attracted 200+ participants and secured industry sponsorships, increased club funding by 50%
- **Peer Mentor**, International Peer Mentoring Program (IPMP) May 2022 – May 2024
  - Mentored 50+ international students to achieve personal & academic goals by fostering a diverse and inclusive environment

Work Experience

Machine Learning Engineering Intern July 2024 – Present
Aider Ventures Indianapolis, IN

- Develop automated pipelines using ChromaDB and FAISS Index Embeddings to fetch and summarize trending research papers, processing and summarizing over 2000 papers, ensuring efficient data handling and information retrieval from icml.cc
- Engineer scalable data pipelines leveraging Gemini API, ZenML, and Prefect Python to analyze and classify over 10,000 research papers, enhancing the categorization process of AI and LLM datasets
- Implement metadata filtering algorithms in LangChain, improving the precision of summarized research outputs by 20%
- Focus on building scalable, DAG-based pipelines, employing advanced techniques to manage and process large datasets, reducing processing time by 40%, improving the efficiency of data workflows, and enabling rapid data analysis and decision-making
- Deploy a Streamlit-based application showcasing the developed application enhancing user engagement and accessibility

Business Analyst Intern – Office of Technology Services December 2023 – March 2024
Legislative Services Agency Indianapolis, IN

- Collaborated with software developers to support & troubleshoot critical software, achieving a 95% issue resolution rate
- Executed advanced data verification and timeliness assurance processes for the Indiana General Assembly, leveraging data analytics and quality control methodologies to maintain unparalleled data fidelity and promptness with a 98% accuracy rate
- Facilitated strategic interactions and technical support with Indiana lawmakers and Legislative Services Agency staff, optimizing business processes through problem resolution and the development of tailored applications and tools

Projects

Mutual Learning Algorithm for News Classification (Senior Capstone Project) August 2024 – Present

- Implement a mutual learning framework that improves Naïve Bayes accuracy by 14% and SVM by 5% in classifying the BBC news dataset. Develop a text preprocessing pipeline, reducing noise by 20% and boosting precision by 15%. Achieve 96.83% accuracy with a non-linear SVM on 2,225 news articles, across 5 categories, optimizing efficiency in resource-constrained environments

Weather Station Data Hub (Python) March 2024

- Developed a Flask and Pandas-powered web service delivering precise weather data, station information, and year-filtered analytics through a streamlined API interface, ensuring real-time updates and 99% uptime

Workout Planner App (Python) January 2024

- Implemented a Workout Planner in Python featuring multi-interface (CLI, GUI via PySimpleGUI, Web via Streamlit) architecture, advanced file I/O for persistent data storage, and cross-platform support to optimize user interaction

Counts of Bushels Project (SQL) October 2023

- Architected a 5NF normalized database for Indiana's corn and soybean yields, utilizing SQL for relational database implementation, improving data retrieval efficiency by 40%, and authored a comprehensive technical report detailing the database design

Technical Skills

- **Languages:** Python, SQL, R/R-Studio, Java, C++, C, JavaScript, CSS
- **Framework and Libraries:** Gemini API, Pandas, PyTorch, TensorFlow, LangChain, Haystack, ZenML, Prefect, Tableau, PowerBI, Flask, Django, PyPDF2, Fitz, GitHub, AWS, Microsoft Azure, React, Redux, Express, PySimpleGUI, Paperetl, GROBID
- **Database Management:** MySQL, Microsoft SQL Server, Docker, Dagster, MongoDB, Azure Data Studio