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

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**

- Developed a machine learning pipeline combining Random Forest, SVM, Decision Tree, and Feedforward Neural Networks, achieving 97% recall and 95% ROC AUC in predicting diabetes.
- Addressed class imbalance using SMOTE and SMOTE-Tomek, improving prediction accuracy and model generalizability.
- Identified HbA1c and blood glucose as key predictors and enhanced deep learning model performance by 15% using batch normalization and L2 regularization, demonstrating scalable and interpretable healthcare AI solutions

Mutual Learning Algorithm for News Classification | *Python, NLP, Machine Learning* **Aug 2024 – Dec 2024**

- Developed a mutual learning framework integrating Multinomial Naive Bayes, SVMs, and MLP Neural Networks, achieving 98% accuracy.
- Engineered text preprocessing techniques (tokenization, lemmatization) to enhance text representation and model performance, leading to faster convergence and reduced over-fitting.
- Improved classification accuracy by 3.21% through a student-teacher knowledge-sharing approach, enabling efficient model training in resource-constrained environments.

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
- Secured Industry Sponsorship and organized hackathons for 200+ students as the Vice-President of the Computer Science club resulting in doubled club engagement.