

Arnav Sharma

arnav-sharma2.com | ags7726@psu.edu | linkedin.com/arnav-sharma2 | github.com/Arnavsharma2 | 408-806-6951

Education

Penn State University
Computer Science, B.S. (Engineering)
Minor Artificial Intelligence

GPA: In Progress
Expected December 2028

Experience

Wefire

Hayward, CA

Software Engineering Intern

January 2025 – Present

- **AI-Powered Reddit Post Analyzer ([Github](#))**: Developed a robust scraping tool in **Python** to collect and analyze up to **5,000** posts from targeted subreddits, automating the process of gathering user sentiment. Integrated the Google Gemini API to perform complex **NLP tasks**, automatically classifying each post by financial domain (e.g., investing, retirement) and question type, and generating a concise 2-3 sentence summary. Engineered a data processing pipeline using **Pandas** to systematically structure the raw scraped data and AI-generated insights.
- **Monitor + Notification Tool ([Github](#))**: Engineered an automated bot to monitor Reddit posts in real-time by streaming data from specified subreddits using the **PRAW** library. Implemented a keyword-matching system to instantly identify relevant posts by searching titles and body content, triggering automated alerts. Developed a notification pipeline that sends detailed email alerts to a designated recipient via **SMTP**, ensuring immediate awareness of keyword mentions.

Projects

End-to-End Predictive Modeling Implementations ([Github](#))

2025

Python, Scikit-learn, Pandas, XGBoost, Matplotlib, Plotly, PyCaret, TensorFlow, Keras, Google

- Developed a portfolio of machine learning models to solve regression, classification, and time-series challenges, including predicting house prices, forecasting stock returns, and identifying customer churn with the best models projecting over **98%** prediction accuracy.
- Engineered comprehensive **data preprocessing** pipelines to systematically handle missing values, perform feature scaling, and encode categorical data for diverse datasets.
- Architected, trained, and evaluated a variety of algorithms including **Linear Regression, Ridge, Lasso, XGBoost, deep learning-based LSTM neural network**, to select the optimal model for each specific case.

Interactive Resume Chatbot ([Github](#))

2025

Python, LangChain, OpenAI API, Vector Databases (ChromaDB/FAISS), Flask, Google Gemini API

- Architected a Retrieval-Augmented Generation (**RAG**) pipeline using **LangChain** to create a conversational AI agent capable of answering detailed questions about my professional and academic background.
- Engineered a data ingestion process to parse my resume, automatically chunking the text and storing it as vector embeddings in a **ChromaDB** database for efficient, context-aware information retrieval.
- Developed and integrated a Large Language Model (**LLM**) via the **OpenAI API** to generate natural and accurate responses, creating an interactive chatbot that effectively simulates the initial stages of a technical interview.

Full-Stack Campus Dining Menu Analyzer ([Github](#)) ([Website](#))

2025

React.js, Node.js, Express.js, Cheerio, HTML, CSS, Google Gemini API

- Developed a full-stack web application to scrape, parse, and display daily menu data from the Penn State University dining website, providing students with a centralized, user-friendly interface to view nutritional information.
- Engineered a backend server with **Node.js** and **Express.js** to manage data scraping logic using **Cheerio**, implementing a caching system that reduced redundant requests by **95%** and improved data delivery speed. Constructed a dynamic and responsive frontend with **React.js**, creating an intuitive user interface with search and filtering capabilities that allow users to instantly find specific food items across multiple dining halls.

Technical Skills

Languages: Python, JavaScript, Java, C++, HTML, CSS

Technologies: React.js, Node.js, Express.js, Flask, LangChain, TensorFlow, Docker, Scikit-learn, Pandas, NumPy, XGBoost

Concepts: Data Structures and Algorithms, Frontend, Backend, Systems, Debugging, Web Development, Data Science, Data Analysis, Agile Methodologies, Management, Testing, System Design and Development