

Arnav Sharma

Contacts:

- **Personal Website:** arnav-sharma2.com
- **Email Address:** aqs7726@psu.edu
- **LinkedIn:** linkedin.com/arnav-sharma2
- **Github:** github.com/Arnavsharma2
- **Phone Number:** 408-806-6951

Professional Summary A motivated and proactive Computer Science student at Penn State with a passion for applying AI and machine learning to solve real-world problems. Experienced in full-stack development, data analysis, and building end-to-end predictive models. Seeking to leverage my skills in software engineering and AI to contribute to innovative projects in the tech industry.

Education Penn State University

- **Major:** Computer Science, B.S. (Engineering)
- **Minor:** Artificial Intelligence
- **GPA:** First Semester of Fall 2025 in progress
- **Class Standing:** Sophomore (by credits)
- **Graduation Timeline:** Accelerated (Expected December 2028)
- **Relevant Coursework:** Physics: Mechanics, Programming and Computations I, Intro to Psychology, Communication Arts, History of the Vietnam War

Experience Wefire | Hayward, CA *Software Engineering Intern* | January 2025 – Present

- **AI-Powered Reddit Post Analyzer** (Github: <https://github.com/Arnavsharma2/Reddit-Scraper-and-AI-Analysis>): 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: <https://github.com/Arnavsharma2/SubReddit-Monitor>): 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: <https://github.com/Arnavsharma2/Predictive-Modeling-Implementations>) | 2025

- *Technologies: Python, Scikit-learn, Pandas, XGBoost, Matplotlib, Plotly, PyCaret, TensorFlow, Keras*
- 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.
- Designed, trained, and evaluated a variety of algorithms including Linear Regression, Ridge, Lasso, XGBoost, and a deep learning-based LSTM neural network to select the optimal model for each specific case.
- **Interactive Resume Chatbot** (Github: <https://github.com/Arnavsharma2/Chat-With-my-Resume>) | 2025
 - *Technologies: Python, LangChain, OpenAI API, Vector Databases (ChromaDB/FAISS), Flask, Google Gemini API*
 - Designed 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.
- **Full-Stack Campus Dining Menu Analyzer** (Github: <https://github.com/Arnavsharma2/PSUMenuAnalyzerWebsite>) (Website: psumenu.com) | 2025
 - *Technologies: 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.
 - 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.

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 & Algorithms, Algorithmic Problem-Solving (LeetCode), Frontend, Backend, Systems, Debugging, Web Development, Data Science, Data Analysis, Agile Methodologies, Management, Testing, System Design and Development

Personal Interests & Hobbies

- **Fitness:** I enjoy hitting the gym and currently bench 190 lbs. My philosophy is that consistency trumps all other factors for gaining muscle. My advice for beginners is to stay consistent and focus on nutrition.
- **Food & Cooking:** I love cooking, staying healthy by eating whole foods, and trying new dishes. My favorite food is goat curry with naan or rice.
- **Travel & Outdoors:** My favorite hike was Mission Peak in California. I have also visited Virginia, Hawaii, India, and Mexico.

- **Movies & TV:** My favorite movies include The Shawshank Redemption, Shutter Island, and Nightcrawler. My favorite actor is Jake Gyllenhaal, and my favorite TV show is Snowfall.

Previous Education

- **Monta Vista High School** | Cupertino, CA | 2021 - 2025
 - *Activities:* CS Club, Chess Club
 - *Relevant Coursework:* APCS A, AP CALC-AB, AP PHYS-A, AP Gov, AP STATS, AP MacroEcon
- **De Anza College (Dual Enrollment)** | Cupertino, CA | 2023 - 2024
 - *Relevant Coursework:* Precalculus, Python Programming, Environmental Science, Spanish 3