

Zhechao Jin

97zjin@gmail.com | (480) 692-0734 | San Jose, CA | LinkedIn: [ZcJin](#) | GitHub: [ZekeJin97](#)

EDUCATION

Northeastern University, San Jose, California Sep 2023 - Present
M.S. in Computer Science, GPA: 3.86 / 4.0 *Expected Graduation: Sep 2025*
Relevant Coursework: Natural Language Processing, Machine Learning, Data Mining, Algorithms, Social and Statistical Network Analysis, Object-Oriented Design, Discrete Structures, Cloud Computing, Pattern Recognition and Computer Vision

University of Pennsylvania, Philadelphia, Pennsylvania Aug 2021 - May 2023
M.S.Ed in Higher Education, GPA: 3.93/ 4.0

Arizona State University, Tempe, Arizona Aug 2017 - May 2021
B.A. in Communication, GPA: 3.48/ 4.0

TECHNICAL SKILLS

Programming Languages: Python, SQL, R, JavaScript, TypeScript, Java, C

Tools & Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, spaCy, NLTK, OpenAI API, LangChain, Scikit-learn, Seaborn, Matplotlib, NumPy, Pandas, Next.js, React, FastAPI, AWS, Docker, OpenCV

PROJECTS

Multilingual AI Advisor for International Students | [GitHub](#) Mar 2025 - May 2025

- Built a full-stack RAG system (FastAPI, FAISS, Sentence-Transformers, GPT-4) for real-time advising on U.S. college admissions and F-1 visa compliance in English, Spanish, and Mandarin.
- Developed a serverless-ready FastAPI backend for automatic PDF ingestion, chunking, embedding (MiniLM-L6-v2), vector storage, and multilingual LLM querying.
- Created a Next.js + Tailwind frontend for document upload, live Q&A, language switching, and admin controls with <1s response time.
- Implemented soft/hard delete workflows, metadata versioning, and health/stats APIs to support scalable deployments.
- Designed modular RAG pipelines to enable easy expansion to new domains (e.g., UK Tier 4, Canadian study permits).

Spanish Dialect Detection | [GitHub](#) Jan 2025 - Apr 2025

- Fine-tuned wav2vec2-spanish for multi-class dialect classification (Spain, Mexico, Chile, Argentina, DR) using Hugging Face's Trainer API with 92.1% accuracy.
- Preprocessed .parquet audio with dynamic range normalization, silence removal, and truncation; used Wav2Vec2Processor for consistent sampling & input formatting.
- Enforced speaker-disjoint train/test splits to reduce speaker bias.
- Compared text-based classifiers: TF-IDF + Logistic Regression, TF-IDF + SVM (linear kernel), and fine-tuned multilingual BERT; found lexical cues more informative than semantic embeddings.

HigherEd Dropout Prediction | [GitHub](#) Jan 2025 - Apr 2025

- Predicted student dropout risk using Logistic Regression, Random Forest, and XGBoost with soft voting ensemble.
- Engineered features from enrollment history, GPA progression, financial status, and academic performance.
- Visualized feature importance using SHAP values (model interpretability tool) to surface interpretable dropout risk factors.
- Performed 5-fold stratified cross-validation with high F1-scores (0.91) across folds, demonstrating model stability.

WORK EXPERIENCE

University of Pennsylvania, Philadelphia, Pennsylvania Sep 2021 - May 2023
Special Projects Coordinator - Education Entrepreneurship Program

- Assisted program manager with day-to-day operations of running the graduate degree program, including project management, event planning, scheduling, and logistics.
- Managed social media accounts (Twitter, LinkedIn, Facebook), leveraging AI tools to optimize program marketing materials to improve social media engagement.
- Provided on-site support for program-intensive weekends, including IT support for faculty and students, and site setup.

Arizona State University, Tempe, Arizona May 2020 - May 2021
International Admissions Representative

- Assisted international prospective and admitted students in the admissions process, leveraging strong communication skills for recruitment.

INTERESTS AND HOBBIES

- Game Modding, Spanish Learning, Karaoke