

# Weijia Xiao

Boston, MA | [xiao.wei@northeastern.edu](mailto:xiao.wei@northeastern.edu) | [LinkedIn](#) | [GitHub](#) | [ResearchGate](#) | [Google Scholar](#) | [Portfolio](#)  
Availability: January–August 2025, Open to Relocation

## EDUCATION

**Northeastern University**, Khoury College of Computer Sciences Expected Graduation: May 2026  
*Bachelor of Science in Computer Science, Concentration in AI* | (PlusOne) Master of Science in Artificial Intelligence | **GPA: 4.0 / 4.0**  
**Courses:** Machine Learning and Data Mining, Deep Learning, Natural Language Processing, LLM-Integrated Systems, Reinforcement Learning, Mathematics of Data Models (Linear Algebra & Statistics), Data Science Foundations, Database Design, Computer Systems, Biostatistics  
**Master/PhD Courses:** Artificial Intelligence Foundations, Machine Learning with Small Data, Algorithms and Data Structure

## SKILLS

**Languages:** Python (6 years), Java (4 years), C/C++ (2 years), Javascript (2 years), SQL, C#, Go, R, Racket  
**Machine Learning:** PyTorch, Keras, Sklearn, BeautifulSoup, Hugging Face, Vector Database, Retrieval-Augmented Generation  
**Software Development:** Next.js, Vue.js, React, Linux, Git, MySQL Workbench, Firebase, Unity, Maya, Figma, Vercel, Redis

## PEER-REVIEWED JOURNAL PUBLICATIONS

Xue Zhang\*, **Weijia Xiao\***, Brent Cochran, Wangxin Xiao. DeEPsnap: human essential gene prediction by integrating multi-omics data. bioRxiv, 2024. DOI: 10.1101/2024.06.20.599958. **\*co-first author** (equal contribution), currently under review  
Xue Zhang, Wangxin Xiao, **Weijia Xiao**. DeepHE: Accurately predicting human essential genes based on deep learning. PLOS Computational Biology, 2020, 16(9): e1008229. DOI: 10.1371/journal.pcbi.1008229. **47 citations, JCR Q1, impact factor: 3.8**

## EXPERIENCE

**Machine Learning Researcher** [Python, PyTorch, Keras, Sklearn] Jul 2019–Present  
• Published 1 academic paper in a JCR-Q1 (top 11%) peer-reviewed journal, with a total of **2 preprints and 2 journal papers; Got 48 citations**  
• **Proposed a novel snapshot ensemble DNN** (Deep Neural Network), boosted 5 metrics by **1~3%** from baseline DNN for essential gene prediction; experimented and outperformed **GAT** (Graph Attention Networks), **Random Forest**, **AdaBoost**, and **SVM** (Support Vector Machine) up to **15%**  
• Applied and currently improving **Diffusion** model for segmenting medical images, reached **> 80%** mIoU accuracy

**AI Software Engineering Fellow** [Next.js, Firebase, LLM API, Go, MySQL, RAG, Vector Database] Jul 2024–Sep 2024  
*Headstarter* Remote

• **Won Top 3 Finalist in TikTok Hackathon:** coded a social media platform using Go, Redis, and MySQL for backend; designed a recommendation algorithm for ranking the posts and an audience-favored innovative product feature; started from 0 experience in Go, completed within 43 hours  
• Developed an inventory management system with real-time CRUD operations, search functionality, and photo previews; **adapted quickly to all-new full-stack tech stacks** (Material UI, Next.js, React, Firebase, Vercel), **enabled efficient project delivery in a week;** received **>5k** usages  
• Deployed a **GenAI chatbot platform** where user can create and customize their unique chatbots; incorporated CI/CD deployment practices  
• Crafted a **GenAI-powered SaaS** product generating flashcards on users' input topics using Llama 3.1 API; integrated a paywall using Stripe API  
• Built an **interactive support agent** with a custom **RAG pipeline**, responding precisely to users' queries using the knowledge base

**Software Engineer, Game Developer** [C#, Unity, Maya | Java, Java Swing | HTML, CSS, Javascript, Vue.js] Feb 2023–Present  
• Led the development of five 2D/3D games in Unity; **taking flexible roles** (team leader, lead coder, chief graphic and UI/UX designer, 3D modeler) **to complement teammates' strengths for maximizing the outcome as a team; led a team to win Top 11 in a 10,000-attendee game jam**  
• Programmed 4 games in Java using MVC design and AI algorithms (BFS, DFS, A\*, Minimax, Dijkstra); recreated 3 as browser games using Vue.js

## PROJECTS

**AI Game Agents** [Python, Gymnasium] Sep–Oct 2024  
• Constructed a **Q-Learning** agent to play an Action-Adventure game with a **92%** success rate over 100,000 episodes, trained in 2 minutes  
• Coded an agent using the **Simulated Annealing** Local Search algorithm to approximate the optimal solution of the four-color theorem; solved a 15x15 board efficiently in **0.07** seconds, choosing from 9 types of color-filling blocks to minimize the steps usage and ensuring no cell is unfilled

**DevOrbit Chatbot Platform** [Llama 3.1 API, Next.js, React, Firebase, Vercel] Aug 2024  
• Led the design and development of the platform, supporting users to create and customize chatbots, enabling fast deployment in a week  
• Prompt-engineered the **Llama 3.1** LLM, guiding the same model to respond differently according to users' customization for each bot  
• Coded in full stack with Next.js and Firebase for database and authentication; deployed with CI/CD practices for iterative update and deployment

**Image Semantic Segmentation Models Comparison** [Python, PyTorch, torchvision, NumPy, pyplot] Feb–May 2024  
• Preprocessed, visualized, and augmented 7,000+ images with corresponding trimap labels (pet, outline, background) of the Oxford-IIIT Pet dataset  
• Implemented **logistic regression** and **FCN** (Fully Convolutional Network) and curated sophisticated **U-Net** and **DeepLabV3+** models in **PyTorch**  
• Trained, fine-tuned, and tested the models; crafted a detailed report demonstrating a thorough understanding and thinking of the data and models

**Boston House Sale Prediction** [Python, BeautifulSoup, Pandas, NumPy, pyplot, Plotly] Oct–Dec 2023  
• Web-scraped, processed, and visualized **10,000+** Boston recent house sale data from Redfin using a **custom end-to-end pipeline**  
• **Constructed regression models using linear algebra** to fit and predict prices; interpreted the result to aid decisions and discover market trend  
• Coordinated the division of responsibilities among teammates and managed the timeline to ensure the proper completion

## AWARDS

**7 Hackathon / Competition Wins** Aug 2020–Present  
• Won all hackathons attended; served as team leader, product designer and manager, and lead coder; managed clear task division and timeline  
• Excelled at both ideation and coding; proposed novel solutions: wildfire real-time alert using machine learning, coronavirus campus tracker, etc.

**Voxel51 Visual AI Hackathon Intermediate 2nd Place** Sep 2024  
• Designed a wild litter detection and classification system to help preserve wildlife habitats from pollution, thus saving species from endangerment  
• Utilized Transfer Learning techniques; fine-tuned pre-trained **YOLOv8** and **ResNet-50** models, obtained **85%** accuracy in classification

**Northeastern Computer Systems Concurrency Programming Star** Apr 2024  
• Utilized solid data structure knowledge to implement a highly performant concurrent Key-Value store/database in **Linux** using **C**  
• Achieved **>300x** faster than the baseline; speed ranked **2nd** among 200+ students; got a perfect score in **high-concurrency** CRUD tests