

Bingchen Yin

yingbingchen@gmail.com • www.linkedin.com/in/bingchenyin • www.bingchenyin.com

EDUCATION

University of Illinois at Urbana-Champaign | Champaign, Illinois

August 2024 - Present

B.S. in Information Science + Data Science | GPA: 3.9/4.0

SKILLS

Languages: Python, SQL, HTML, CSS, JavaScript, C++, R

Tools: Microsoft Office, Power BI, Git, GitHub, Jupyter Notebook, MySQL, Unreal Engine 5, Visual Studio, Figma

Libraries: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly, SciPy, XGBoost, Statsmodels, PyTorch

EXPERIENCE

SpaceTid | Shanghai, China

June 2025 - Aug 2025

Web Development Intern

- Contributed to front-end website development using HTML, CSS, and JavaScript to design and build pages
- Learning data center structures and systems using CAD tools while also learning about business operations

PROJECTS

Spellbound| Unreal Engine 5, C++, Blueprints

September 2026 - Present

- Developing a single-player platformer with planned exploration, environmental puzzles, and sword-and-magic combat.
- Implemented basic player movement, gameplay mechanics, UI, and HUD.

Personal Portfolio| HTML, CSS, JavaScript, GitHub

June 2026 - Present

- Designed and built a multi-page portfolio showcasing technical projects, skills, education, and experience.
- Implemented project category filtering, smooth transitions between project views, and downloadable project files using JavaScript. Hosted the website on Cloudflare.

Books Data Analysis| Python, Pandas, Matplotlib, Seaborn, Plotly

August 2026

- Cleaned and transformed book metadata and created 24 visualizations exploring different variables.
- Used correlation heatmaps, density plots, and interactive charts to communicate patterns, with written interpretations addressing sampling limitations and potential bias.

Gene Expression Classification | Python, Pandas, Scikit-learn, Matplotlib

May 2026

- Built a 700-tree Random Forest pipeline with median imputation to classify five cancer types using thousands of RNA-Seq gene expression measurements.
- Achieved **100%** accuracy on a stratified held-out test set; evaluated precision, recall, and a confusion matrix while documenting limitations of the controlled dataset.

Housing Price Prediction | Python, Pandas, Scikit-learn, Matplotlib

April 2026

- Developed a histogram-based gradient boosting regression pipeline with missing-value imputation and categorical encoding to predict housing sale prices.
- Achieved an **8.54%** mean absolute percentage error, with **93.39%** of test predictions within 20% of actual prices; examined predicted versus actual values.

Credit Card Fraud Detection | Python, Pandas, Scikit-learn, Matplotlib

April 2026

- Trained a 200-tree Random Forest classifier on 54,276 transactions, with fraudulent transactions representing approximately 0.58% of training observations.
- Achieved **98.48%** precision and **82.28%** recall for fraud detection on 13,569 test transactions; evaluated false positives and missed fraud cases.

MLB Pitch Classification | Python, Pandas, Scikit-learn, Matplotlib

February 2026

- Built a K-Nearest Neighbors pipeline with missing-value imputation, numerical scaling, and categorical encoding to classify MLB pitch types.
- Trained on 2024 tracking data and achieved **98.34%** accuracy on 1,806 test pitches from 2025.

INTERESTS

Personal Interest: PC Hardware, Gaming, Sports(Golf, Basketball, Soccer)

Content Creation: Create content for multiple accounts across YouTube, Instagram, TikTok, and Bilibili. 5M+ total views