

Weihan Xu

Durham, NC | weihan.xu@duke.edu | +1 734-548-7537 | <https://wx83.github.io/>

Education

Duke University

Master of Science in Computer Science (AI/ML track)

GPA: 3.83/4.0

Coursework: Deep Learning(A+), Theory and Algorithm for Machine Learning(A+), Introduction to Medical Robotics, Natural Language Processing, User Security and Privacy, AI Security and Privacy

Durham, NC

Expected May 2025

University of Michigan-Ann Arbor

Bachelor of Science in Computer Science (with Honors) and Data Science

GPA: 3.87/4.0

Honors Thesis: *Recurrent Neural Network on Predicting Intensive Care Transfers and Other Unforeseen Events*

Coursework: Deep Learning for Computer Vision, Machine Learning, Applied Regression Analysis, Database, Data Structures and Algorithms, Theoretical Statistics, Probability and Statistics, Linear Algebra, Introduction to Computer Organization, Introduction to Natural Language Processing

Ann Arbor, MI

April 2023

Publications

Published:

2025:

- **TeaserGen: Generating Teasers for Long Documentaries**
Weihan Xu, Paul Pu Liang, Haven Kim, Julian McAuley, Taylor Berg-Kirkpatrick, Hao-Wen Dong
The Thirteenth International Conference on Learning Representations (ICLR)

2024:

- **A New Dataset for Tag- and Text-Based Controllable Symbolic Music Generation**
Weihan Xu, Julian McAuley, Taylor Berg-Kirkpatrick, Shlomo Dubnov, Hao-Wen Dong
The 25th International Society for Music Information Retrieval - Late Breaking Demo
- **A New Dataset, Notation Software, and Representation for Computational Schenkerian Analysis**
Stephen Hahn, *Weihan Xu*, Jerry Yin, Rico Zhu, Yue Jiang, Simon Mak, Cynthia Rudin
The 25th International Society for Music Information Retrieval (ISMIR)
- **Smart Tools, Smarter Concerns: Navigating Privacy Perceptions in Academic Settings**
Yimeng Ma, *Weihan Xu*, Hongyi Yin, Yuxuan Zhang, Pardis Emami-Naeini
The Twentieth Symposium on Usable Privacy and Security (SOUPS Poster)
- **SentHYMNent: An Interpretable and Sentiment-Driven Model for Algorithmic Melody Harmonization**
Stephen Hahn, Jerry Yin, Rico Zhu, *Weihan Xu*, Yue Jiang, Simon Mak, Cynthia Rudin
The 30th SIGKDD Conference on Knowledge Discovery and Data Mining - Applied Data Science Track

2023:

- **Equipping Pretrained Unconditional Music Transformers with Instrument and Genre Controls**
Weihan Xu, Julian McAuley, Shlomo Dubnov, Hao-Wen Dong
IEEE Big Data 1st Workshop on AI Music Generation with AI Music Competition

Preprints:

2025:

- **Generating Symbolic Music from Natural Language Prompts using an LLM-Enhanced Dataset**
Weihan Xu, Julian McAuley, Taylor Berg-Kirkpatrick, Shlomo Dubnov, Hao-Wen Dong
Under Review
- **Video-Guided Text-to-Music Generation Using Public Domain Movie Collections**
Haven Kim, Zachary Novack, *Weihan Xu*, Julian McAuley, Hao-Wen Dong
Under Review

Research Experience

Professor Julian McAuley's Research Lab

San Diego, CA

Project: Conditional Music Generation

May 2023 - Sep 2023

- Developed a system that generates symbolic music aligned with user-specified genre and instrument inputs by fine-tuning a pretrained music transformer with prepended tokens
- First-authored a workshop paper accepted by 1st Workshop on AI Music Generation at IEEE Big Data 2023 and **won AI Music Innovation Award**
- Collected and constructed a large-scale dataset of nearly one million entries by extracting metadata tags from the MuseScore platform and generating captions with a large language model, resulting in a first-authored paper accepted at ISMIR-Late Breaking Demo
- Successfully built two conditional music generation systems: one that generates symbolic music based on multiple tags and another that generates music using free-form text inputs, resulting in a first-authored preprint
- Developing a text-to-symbolic music model by expanding the vocabulary of the large language model (BLOOM) and fine-tuning it with paired captions and music tokens

Project: Teaser Generation

May 2024 - Sep 2024

- Constructed a dataset of 1,000+ multimodal documentary videos (avg. 30 minutes each) by scraping raw data from YouTube, annotating paired teaser and main videos, and separating audio-visual components. Extracted text via transcription and separated audio tracks into speech, music, and sound effects
- Proposed and implemented an interval-based model and a learning-based model for generating automatic teasers under 3 minutes for long documentaries, designing new evaluation metrics to complement subjective testing, resulting in a first-authored accepted by ICLR 2025
- This project led to a NVIDIA grant with 21K A100 40GB hours (PI: Hao-Wen Dong)

Interpretable Machine Learning Research Lab

Durham, NC

Project: Interpretable music composition with Schenkerian analysis

Dec 2023 - May 2024

- Contributed to the development of a Schenkerian Analysis dataset by annotating music depth and integrating the music data into software for computational recognition
- Proposed a novel clustering approach using a graphical neural network (GNN) that eliminates the need for predefined cluster numbers in each layer, specifically designed for music hierarchical structures
- Developed a codebase from scratch using PyTorch Geometric (PyG) and co-authored an accepted ISMIR paper
- Reimplemented a baseline model with updated libraries, obtaining results for an accepted KDD paper

Professor Sardar Ansari's Research Lab

Ann Arbor, MI

Project: Time Series Prediction in Intensive Care Unit

Nov 2021 - May 2023

- Preprocessed raw medical data to develop a Recurrent Neural Network and Transformer-based model to improve accuracy in predicting patient deterioration rates
- Explored replacing an existing XGBoost model with a deep learning approach for time series data, contributing valuable insights into model performance
- Completed an Honor Thesis under the supervision of Professor Sardar Ansari and Professor Kayvan Najarian

Projects

Diagnostic Classification in Educational Measurement

Ann Arbor, MI

Data Science Major Capstone(Advised by Prof. Gongjun Xu)

Jan 2022 - May 2022

- Proposed a new algorithm to predict psychological attributes with educational assessment data with Cognitive Diagnosis Models package in R
- Overcame computational difficulty of estimating the correlation between answering specific exam questions correctly and psychological attributes with machine learning tools (Restricted Boltzmann Machines)
- Experimented with 4 different methods of determining the correlation matrix, resulting in a research report

Accessibility Tools for the Visually Impaired

Ann Arbor, MI

Computer Science Major Capstone

May 2022 - Jun 2022

- Collectively (in a group of four) built a website that allows visually impaired people to upload photos of

surroundings and describes the photos for them

- Designed various features into the website, such as giving warnings if photo includes roadblocks or other obstacles
- Constructed the code and user interface for the website with one other team member
- Tested the website and made a demo describing the function of the website

Teaching Experience

Electrical and Computer Engineering Department, Duke University

Durham, NC

Course: Introduction to Deep Learning (Taught by Prof. Vahid Tarokh)

Sep 2024 - Dec 2024

- Assist in teaching a graduate-level deep learning course
- Hold weekly office hours to guide students on mathematical proofs and coding assignments
- Mentor nine project teams (four members each) through the design, implementation, and presentation of their final projects
- Develop exam questions

Leadership & Activities

Shanghai Adream Charitable Foundation

Shanghai, China

Project: Bridging Education Gaps in Underrepresented Areas

Jun 2020 - Oct 2020

- Managed the materials group, coordinating the procurement and organization of 14 types of materials for an educational conference with over 150 expert participants
- Supported the Project Manager in preparing the project proposal and delivering reports to donors
- Analyzed the operational effectiveness of over 500 Adream classrooms in Shandong Province and formulated actionable project plans

Professional Service

Journal Reviewer: International Journal of Computer Vision(IJCV)

Skills

Technical skills: C; C++; Python(PyTorch & Tensorflow); Java; SPSS; R; MongoDB; SQL; HTML; CSS; LaTeX

Language: English, Mandarin

Instruments: Piano(Advanced), French horn(Beginner), Violin(Beginner)