David Guo

davidmy.guo@mail.utoronto.ca • guo-david.github.io • guo-david

RESEARCH INTERESTS

Exploring data-driven modeling, reasoning, and interaction. Prior experience includes applied machine learning, analysis, and research-oriented development across multiple domains. Aiming to deepening my expertise and contribute to collaborative research in machine learning and applied science.

EDUCATION

University of Toronto **Sep 2021 – Apr 2026**
B.A.Sc. in Engineering Science (Machine Intelligence Option)
CGPA: 3.83/4.0
Honours: UofT Scholar (2021), Rowe Innovation Award (2022), Dean's List (2022–2025)
Selected Coursework: Deep Learning, Probabilistic Models, Optimization, Computer Vision, Reinforcement Learning.

RESEARCH EXPERIENCE

Data-Driven Decision Making Lab (D^3M), University of Toronto **May 2025 – Present**
Research Assistant | Supervisor: Prof. Scott Sanner – Supported by NSERC USRA.

- Designed and ran a comprehensive human–human study for conversational recommendation with visual catalogs spanning 900 turns and 720 feedback instances.
- Built a dataset and annotation pipeline integrating LLM-assisted intent tagging with human verification, enabling reproducible analysis of dialogue alignment.
- Developed a stage-aware dialogue analysis framework and evaluated GPT-based recommendation models for alignment and user satisfaction prediction.

Acceleration Consortium and Matter Lab, University of Toronto **Sep 2025 – Present**
Research Assistant (Volunteer) | Supervisor: Dr. Nasim Abdollahi

- Developed and integrated a graph attention network model for molecular solubility prediction within the Consortium's generative
- Engineered feature extraction and batching pipelines for large-scale molecular graphs, optimizing data throughput and model stability.
- Designed evaluation protocols and ablation studies to benchmark GAT variants against baseline architectures, improving predictive

Sunnybrook Research Institute (Orthopaedic Biomechanics Lab) **Sep 2025 – Present**
Capstone Research Project | Principle Investigator: Dr. Michael Hardisty

- Modeled healthy spinal geometry via a bi-directional quantile regression framework using XGBoost, predicting normative vertebral volume and inter-centroid distance from demographic and biomarker features.
- Constructed a population-level Statistical Shape Model of healthy vertebrae using Generalized Procrustes Analysis to characterize shape variance and spatial correlation across levels.
- Integrated the regression and SSM modules into a unified Python library and interactive visualization dashboard to enable clinician-facing assessment of spinal metastasis–related deformities.

PUBLICATIONS & PREPRINTS

Guo, D.[*], Sun, M.[*], Jiang, Y.[*], Liang, J., & Sanner, S. *VOGUE: A Multimodal Conversational Fashion Recommendation Dataset.* Submitted to **ACM WebConf 2026.** arXiv:2510.21151
[*]Equal contribution

INDUSTRY EXPERIENCE

ModiFace (L'Oréal) **May 2024 – May 2025**
Software Developer Intern – Augmented Reality

- Re-architected the client-facing rendering layer of an AR makeup try-on platform, improving runtime stability and responsiveness across real-time facial tracking pipelines.
- Designed a unified exception-handling and telemetry framework for debugging and user-side diagnostics, enhancing observability and iteration efficiency.
- Collaborated on modular API design and data instrumentation for production analytics, supporting integration with experimental vision models.

DealerVu **May 2023 – Aug 2023**
Software Developer

- Built end-to-end modules in React and Azure .NET for a large-scale dealership management platform, integrating transactional and analytics pipelines.
- Implemented interactive reporting dashboards with time-series visualization and predictive summaries, supporting data-driven forecasting for dealer performance.
- Designed reusable localization and data-layer components to standardize UI behavior and streamline feature deployment across regions.

SELECTED PROJECTS

Pancreas Lesion Segmentation and Classification (nnU-Net V2) **Winter 2024/2025**
PyTorch | *GitHub*
Extended nnU-Net V2 with a multi-task classification head for lesion subtype prediction, exploring shared-representation learning between segmentation and classification tasks.
Achieved a Dice score of 0.92 on pancreas segmentation and macro-average precision of 0.82 across lesion subtypes.

GAN-Augmented Silicon Wafer Classification **Fall 2024**
PyTorch | *GitHub*
Applied Wasserstein GANs with gradient penalty to synthesize minority wafer defect classes, investigating generative augmentation for imbalanced visual datasets.
Improved downstream classification accuracy by approximately 6%, demonstrating measurable gains in defect detection generalization.

Music Generation Using Autoencoders and Transformer Mixture Models **Spring 2024**
TensorFlow | *GitHub*
Explored generative sequence modeling for symbolic music using a hybrid VAE–Transformer architecture.
Demonstrated improved melodic consistency and rhythmic coherence across benchmark MIDI datasets.

AWARDS & HONOURS

NSERC USRA – Undergraduate Student Research Award **Summer 2025**
Rowe Innovation Award – Best Design Solution, Praxis Course – $4,000 **Apr 2022**
UofT Scholar – University of Toronto Scholars Entrance Award – $7,500 **2021**

RESEARCH & TECHNICAL SKILLS

Research: Experimental design and data curation; human–LLM annotation pipelines; model evaluation, ablation, and error analysis; statistical modeling (SciPy, statsmodels); reproducible pipeline development (Git, Docker, LaTeX, W&B).

Programming: Python, C, C++, MATLAB, JavaScript
ML Frameworks: PyTorch, TensorFlow, scikit-learn
Data Tools: Pandas, NumPy, Matplotlib, Seaborn, OpenCV
Other: Git, Docker, Linux, LaTeX, Weights & Biases