Edwin Chacko

Github | Portfolio Website | LinkedIn | edwin.chacko@mail.utoronto.ca | 647-540-5048

Education

University of Toronto Sep. 2022 – May 2027
BASc. in Engineering Science, Machine Intelligence Option. *Toronto, Canada*
Relevant Coursework: Machine Learning, Data structures, Algorithms, Matrix Algebra and Optimization, Computational Linguistics, Natural Language Computing, Computer Architecture, Probability & Statistics

Technical Skills

Languages: Python, C/C++, CUDA, SQL, JS, Assembly, Verilog, MATLAB, React, Django
Libraries: TensorFlow, PyTorch, scikit-learn, keras, NumPy, pandas, Matplotlib, Hugging Face Transformers
Developer Tools: Git, Docker, VS Code, Visual Studio, Linux (Ubuntu), Shell/Bash, HDF5

Experience

Machine Learning Researcher May 2024 – Present
McMaster University - ChemAI Lab *Hamilton, ON*
- **Founded** and leading the Spectro project at McMaster University, exploring novel applications of AI in chemistry.
- Designed a multi-modal molecule prediction model, **Spectro**, achieving an **accuracy of 93%** (see projects).
- Collaborated with Dr. Kylie Luksa and other domain experts to inform model development and behaviour.
- Co-authored and submitted a paper to **AI4Mat-NeurIPS**, currently pending approval.

Calibrations Engineering Intern May 2023 – Sep. 2023
VACS Calibrations *Toronto, ON*
- Calibrated electronic and mechanical equipment, following the **IEE and ISO17025** standards.
- Performed **statistical analysis**, including standard deviation, uncertainty propagation, and regression, to validate calibration accuracy and assess reliability.
- Developed insights from calibration data, improving process efficiency by **10%** and measurement precision by **15%**.

Projects

Spectro | *TensorFlow, PyTorch, Hugging Face Transformers, JS, Linux* May 2024 – Present
- **Led development** of a multimodal molecule elucidation model using IR and NMR data, achieving **93% accuracy**
- Preprocessed and augmented datasets, employing data normalization and, **r**educing class imbalance by **30%**.
- Fine-tuned a modified ResNet for functional group prediction from IR images, achieving an **f1 score of 91%**.
- Designed and tuned a decoder **(RNN with LSTM)**, achieving **93% accuracy** in molecule prediction.
- Implemented a **complete ML pipeline** in TensorFlow with: data augmentation and oversampling, **SMOTE**, cross-validation, a custom learning rate scheduler, **distributed training**, joint training, on-the-fly dataloading.

Chess Engine | *C++, CUDA, Docker, Postman, JavaScript* Jun. 2023 – Apr. 2024
- C++ chess engine (rated 1800) with hardware optimizations like BitBoards and compile-time optimizations.
- Implemented multithreaded Negamax search with alpha-beta pruning, **reducing search time by 45%**.
- Integrated Zobrist hashing, a transposition table, and quiescence search, to improve search accuracy and eliminate redundant calculations.
- Utilized custom **CUDA** kernels for move generation and evaluations, significantly accelerating parallelizable tasks to achieve around **35 million nodes per second** in performance test.

Chess NNUE (Efficiently Updatable Neural Network) | *PyTorch, C++, SQL, HDF5* May 2024 – Present
- Developing chess static evaluation with NNUE in PyTorch, currently achieving **80% accuracy**.
- Augmented and preprocessed **83 million data points**, exploring utilizing a custom C dataloader.
- Bridging the **int_8 quantized** variant into my C++ chess engine, **reducing inference time by 70%**.

Remote ML Server | *Linux (Ubuntu), Bash, Docker, Git* Jun. 2024
- Configured a retired GPU cryptocurrency mining rig into a remote server using docker, achieving **10x faster training time** for ML models.

Publications

[1] (Pending) Chacko, Sondhi, et al. *"A Multi-modal Approach for Molecule Elucidation Using IR and NMR Data."* AI4Mat-NeurIPS 2024. December 2024.