

Ethan Waxman

ejw99@cornell.edu | (202) 604-8256 | github.com/ethan-wax | ethanwaxman.com

Machine Learning Engineer — PyTorch · Apache Airflow · AWS SageMaker · ETL Pipelines · Python · Production ML

EDUCATION

Cornell University, College of Engineering — B.S. Computer Science & Applied Physics, GPA: 3.53

Honors: 4-Year Varsity Heavyweight Rower (balanced 20+ hrs/week training with rigorous double major)

Certifications: AWS Solutions Architect — Associate (Jun. 2026)

Relevant Courses: Foundations of AI, Computational Linguistics, Neural Networks, Compilers, Computer Architecture, Operating Systems, Discrete Structures (Honors), Functional Programming

ML & DATA ENGINEERING EXPERIENCE

FI Consulting, Arlington, VA Software Developer Jun. 2024 – Present

- Led design of 3 Graph Neural Network (GNN) architectures for recommendation and churn prediction; achieved 82% prediction accuracy identifying \$100M+ in at-risk revenue; collaborated with clients to iterate and orient the project.
- Architected and optimized GNN training pipelines using PyTorch to process 1M+ nodes/users in under 5 minutes, overcoming high-dimensional data scaling bottlenecks.
- Owned an ML entity resolution service on AWS SageMaker from development to deployment, utilizing fuzzy matching and probabilistic record linkage to remove 10,000+ duplicates.
- Engineered high-throughput ETL pipelines using Apache Airflow to migrate multiple terabytes of data into MongoDB for downstream model consumption.
- Developed an executive analytics dashboard using PySpark and Databricks, enabling real-time distributed querying of 15M+ customer records.
- Deployed a Jenkins pipeline to automate CI/CD; dropped build and deployment times by 75%.

Machine Learning Intern Jun. 2023 – Jun. 2024

- Reduced inference latency for a RAG-based search application by 45%, implementing Redis caching to ensure high reliability and sub-second response times for 50+ concurrent users.

RESEARCH AND ML PROJECTS

Large Language Model (LLM) Cognition Research — Cornell University

- Designed an experimental framework to evaluate LLM interpretability using controlled syntactic ambiguity tasks; analyzed correlation between next-token probabilities and human repair preferences.
- Authored a 7-page technical paper analyzing how model architecture affects the processing of syntactic ambiguity.

Iterative Chess Engine — Python, PyTorch, Reinforcement Learning

- Developed 4 engine iterations, evolving from Minimax with Alpha-Beta pruning and MCTS to a hybrid deep learning evaluation model.
- Trained a Policy-Value Neural Network in PyTorch using 100,000 Grandmaster positions (100 epochs) and reinforced through thousands of self-play iterations.

Full-Stack Academic Compiler — Java

- Led a 3-person team to engineer a 10,000+ line compiler supporting 15+ language features including recursion, arrays, and lexical scoping.

TECHNICAL SKILLS

ML Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, NLTK

Data Tools: Apache Airflow, PySpark, Databricks, MongoDB, Redis, ChromaDB

Cloud & DevOps: AWS (Solutions Architect – Associate, SageMaker, Certified Cloud Practitioner), Docker, Neo4j, Terraform, Jenkins

Languages: Python (Advanced), Java (Advanced), OCaml (Advanced), Rust (Intermediate), C++ (Intermediate)