Yonatan Yishak Yifat

Addis Ababa | yonatanyishak111@gmail.com | Linkedin.com/in/yonatanyishakyifat | GitHub.com/YonatanBest | Leetcode.com/u/YonatanBest

EDUCATION

Addis Ababa University, Addis Ababa ***Expected Graduation: Aug 2027***

Bachelors of Science in Information Systems

SKILLS

Languages ***Intermediate*** Python(2yrs) | ***Intermediate*** C++(2yrs) | ***Intermediate*** JavaScript(1yrs) | ***Beginner*** Java(1yrs)

Software Docker | Jupyter Notebook | Git | VSCode | Linux | GitHub | LaTex

Frameworks and Libraries Langchain | Tensorflow | Flask | Pytorch | React | Scikit Learn | Numpy | Seaborn | Pandas | ngc-learn

EXPERIENCE

Machine Learning and NAC Intern ***Jan 2025 - Present***

iCog Labs *Addis Ababa, AA*

- Integrated AMP training using PyTorch amp, reducing model training time by **~30%** without sacrificing performance.
- Implemented **Bayesian hyperparameter tuning** with Optuna for PC-Transformer models, including dynamic learning rate scheduling and warm-up configuration, leading to faster convergence in training trials.
- Replaced **word-level tokenization with subword-level BPE**, resolving uncertainty in Transformer outputs and improving model generalization on rare tokens.
- Discovered and resolved device mismatch in PC-Transformers by refactoring for device-agnostic execution on CUDA.
- Documented research findings and presented technical summaries of **5+ academic papers**

Machine Learning and AI Research Intern ***Oct 2024 - Dec 2024***

ThinkingBeyond Skåne County, Lund

- Earned a spot in the highly competitive program with an acceptance rate of only 10.7%
- Built and trained **6+ CNN models** using TensorFlow on a 10,000-image dataset for binary classification (Cats vs. Dogs)
- Conducted comparative analysis of SGD, Adam, AdaGrad, AdaDelta, RMSprop & Nadam across 4 metrics
- Optimized learning rates and batch sizes, achieving **79.09% accuracy and 79.33% F1-score** with AdaGrad
- Led dataset preprocessing, CNN architecture design, and reproducibility pipeline for peer-reviewed experimentation
- Collaborated on a paper titled **A Comparative Analysis of Optimization Algorithms for Classification with CNN**
- Designed a research poster, which was selected as a finalist

Research and Data Analysis Intern ***Jun 2024 - Aug 2024***

Ethioware *Addis Ababa, AA*

- Designed and deployed targeted surveys to collect high-quality, relevant data for ongoing AI research initiatives
- Engineered and cleaned over **200 raw datasets**, improving data integrity and analysis reliability
- Built and validated multiple AI models using **19 selected features**, extracting actionable insights for internal use
- Achieved **10% more accuracy** for the model from the previous trained model

PROJECTS

AI Keystroke Predictor with Typing Sound | Flask | Random Forest | Librosa *(~30 hours)* - AI Keystroke Predictor ***Sep 2024 - Dec 2024***

- Engineered ML system to predict keystrokes from audio, targeting vulnerabilities in acoustic side-channel attacks
- Processed **5,531** labeled audio samples (2,585 spacebar, 2,946 other keys) using Librosa to extract relevant acoustic features
- Extracted audio features with Librosa and trained Random Forest classifier, achieving **98.55%** accuracy
- Deployed a Flask-based web app to simulate and demonstrate real-world inference risks in low-noise environments

CodeAI – AI-Powered Code Generation Tool | Flask | IBM Watson AI | Prompt Engineering *(~10 hours)* - CodeAI ***Oct 2024 - Oct 2024***

- Developed web app that uses IBM Watson's foundation models to generate code snippets from user-defined specs
- Designed a user-friendly Flask interface with secure prompt customization and dynamic language/framework support

Time Series Prediction with LSTM and S4 (SSM) | PyTorch | NumPy | Matplotlib *(~12 hours)* - Time Series ***Apr 2025 - May 2024***

- Developed two deep learning models **(5-layer LSTM and 5-layer S4)** for multivariate climate forecasting using the Daily Delhi Climate dataset
- Built cyclic temporal features and applied sequence modeling to predict mean temperature, humidity, and wind speed
- Achieved strong forecasting performance with S4, improving LSTM's **MAE** from **2.07°C to 1.60°C**
- Visualized prediction trends vs. actual climate patterns; implemented scaling, sequence slicing, and temporal embeddings for effective modeling

Activities

Reboot The Earth International Hackathon - Finalist in hackathon after building AI model for climate farming suggestions ***Jun 2024***

Ethiopian Collegiate Programming Contest - Received Honorable Mention for problem-solving and algorithmic excellence ***Oct 2024***

IBM TechXchange Pre-Conference watsonx Hackathon - Top 100 project score from 4000+ submissions ***Sep 2024 - Oct 2024***

Ethiopian Statistical Service Technician - Contributed to the 2nd Agricultural Census with technical support ***Aug 2024 - Sep 2024***