Khawaja Abaid Ullah

khawajaabaid.github.io | khawaja.abaid@gmail.com | in/KhawajaAbaid | Rochester, NY

Objective

I'm driven by the strong desire to work on the development of innovative and revolutionizing machine learning algorithms, especially those that have the promise to improve the power-efficiency of AI models.

Education

Master of Science in Artificial Intelligence   2024-2026
Rochester Institute of Technology   Rochester, NY
- Fully Funded by US Dept. of State under Fulbright Scholarship

Bachelor of Science in Computer Science   2017-2021
University of Narowal   Narowal, Pakistan
- Graduated as the Topper of Computer Science Class of 2021

Work Experience

Open Source Developer   May 2022 - Present
Self Employed   Pakistan
Teras (*GitHub*)
- Developed Teras, which is a Unified Tabular Deep Learning library built on top of Keras 3.
- Implemented state of the art Deep Learning architectures for Tabular data proposed by researchers, from classification, regression, to data imputation as well as synthetic data generation.
- It supports all popular Machine Learning backends, namely JAX, PyTorch and TensorFlow.

Keras
- Contributed to the development of Keras which is the most popular Deep Learning framework used by millions every day. My contribution includes the development of Beta and Binomial random sampling functions that support all Keras backends, namely JAX, PyTorch and TensorFlow.

micrograd_c (*GitHub*)
- Implemented Andrej Karpathy's micrograd library from scratch in C. micrograd is a basic scalar value autograd engine with a neural network library built on top that is similar to PyTorch API.

Junior Python Developer (Remote)   Feb 2023 - May 2023
Comundo   Cophenhagen, Denmark
- Developed semi-supervised, robust web scrapers for energy consumption data curation from approved client sources.
- Cleaned and aggregated the scraped data to estimate CO_2 emissions for clients as well as for downstream training of machine learning models for future forecasts.

Freelance Software Developer   Oct 2021 - Sep 2023
Self Employed / Fiverr   Pakistan
- Developed web scraping tools for clients using Python libraries like requests, BeautifulSoup, selenium etc.
- Developed automation tools to automate various repetitive digital tasks like file conversions and data movement in bulk.

Recent Projects

codriver – *An efficient LLM powered contextualized code generation tool*   December 2024
- Inspired by GitHub copilot, attempted to make an extremely efficient contextualized code generation tool using the Gemma family of LLMs open sourced by Google Deepmind.
- Ported Gemma 2, from Flax linen to NNX API. Developed a custom Gemma 260M variant.
- Using knowledge distillation, trained the custom Gemma 260M variant first using Gemma 2 2b and then CodeGemma 2b as teacher models, using text and code datasets respectively for the model to learn lexical as well as code structure.

Others

Achievements: Fulbright Scholar. Kaggle Expert.
Certificates: Harvard University's CS50X.