

Khawaja Abaid Ullah

Creator Teras | Machine Learner | Fulbright Scholar



khawaja.abaid@gmail.com

in KhawajaAbaid



khawajaabaid.github.io

Computer Science graduate specializing in machine learning through self-learning. I've created Teras, which is a Keras-based unified deep learning library for tabular data.

I also occasionally blog about Machine Learning on medium. [Check them out!](#)

SKILLS

- **Programming:** Python, C, JavaScript, C++, C#
- **Data Wrangling:** Pandas, Numpy, Scipy, Statsmodels, NetworkX
- **Visualization:** Matplotlib, Seaborn
- **Machine Learning:** JAX, Flax, Keras, Tensorflow, PyTorch, Scikit-Learn, XGBoost, CatBoost, LightGBM
- **NLP:** AllenNLP, NLTK, Spacy, Hugging Face
- **Database:** SQL, MongoDB
- **Cloud:** AWS EC2
- **Web Scraping:** Requests, BeautifulSoup, Selenium, Scrapy
- **Others:** Django, PyQt5, Git, OpenCV, Photoshop, MS Office

EDUCATION

University of Narowal

Bachelor of Science in Computer Science (GPA: 3.69/4.0) (Topper)

2017-2021

CERTIFICATES

Harvard University's CS 50

CS50x 2022 Intro to Computer Science ([Official Certificate Link](#))

EXPERIENCE

Open-Source Developer / Self Employed

May 2022 – Present

- **Teras:** A Unified Deep Learning Library for Tabular Data ([Github](#))
- **Keras:** Contributor to the development of Google's deep learning library
- **Prewardict:** An NLP based nerdy game. (<https://khawajaabaid.github.io/prewordict/>)
- **machine-learning-in-c:** Machine learning from absolute scratch in c. ([Github](#))
- **TweetsCloudBot:** A twitter bot that generates word clouds of tweets. ([Bot Twitter Link](#))

Junior Python Developer (Remote) / Comundo, Denmark

Feb 2023 – May 2023

- Developed semi-supervised, robust web scrapers for data curation
- Aggregated the scraped data to estimate CO2 emissions of clients.

Freelance Software Developer

Oct 2021 – Sept 2023

- Web Scraping
- Task Automation

PROJECTS

Teras – A Unified Deep Learning Library for Tabular Data / Open-Source Project

- Provides state-of-the-art deep learning architectures for tabular data proposed by researchers under a uniform API.
- Supports popular backends including JAX, PyTorch and TensorFlow.
- Implements classification, regression, data generation, imputation architectures and much more!

USED: Keras, JAX, PyTorch, TensorFlow, TensorFlow-Probability, Scikit-Learn

Repo: <https://github.com/KhawajaAbaid/teras>

machine-learning-in-c – Machine Learning from Absolute Scratch in C (*Under development*)

- Implements all the backpropagation, linear algebra operations and everything else from scratch without using any third-party library.
- Currently provides implementations of modes like MLP and GAN, but new architectures are coming.

USED: C

Repo: <https://github.com/KhawajaAbaid/machine-learning-in-c>

Keras – Google’s Deep Learning Library

- Implemented Beta and Binomial distribution functions in three different backends namely TensorFlow, JAX and PyTorch, and wrapped them to be used by a single interface.
- Fixed some bugs and improved some code

USED: Keras, JAX, PyTorch, TensorFlow

jax-machine-learning – Machine Learning from scratch in JAX (*Under development*)

- Aims to implement models from simple MLPs to state-of-the-art generative models
- Currently provides implementations of modes like MLP and GAN, but new architectures are coming.

USED: JAX

Repo: <https://github.com/KhawajaAbaid/jax-machine-learning>

Prewardict / A fun little nerdy game (Inspired by Wordle)

- Processed 180,000 medium articles and clustered the similar words
- Generated word clouds from these clusters. Given a word cloud, user has to predict one word that best belongs in the word cloud.

USED: Scikit-Learn (TFIDF, SVD, KMeans), WordCloud, Numpy, Pandas, Matplotlib, Kaggle

Play it here: <https://khawajaabaid.github.io/prewordict/>

Crowd Detection and Behavior Analysis / Final Year Project, BS-CS

- Worked on Computer Vision to train Deep Learning models for Person Detection as well as Detection of Certain (similarly attired) Groups among crowd.
- Developed custom algorithms for Chaos Detection and Crowd Detection

USED: TensorFlow Object Detection API, OpenCV, Matplotlib, PyQt5, Google Colab

Gamer Assistant / Addresses the unaddressed needs of gamers / CS50X Project

- A website based on Django and consisting of three main components as highlighted below.
- **FPS Predictor:** Predicts the expected FPS for a user specified PC for a specified game
 - Scraped and curated FPS and games data from sites like Steam & Techpowerup.
 - Used XGBoost to build the FPS Predictor

- **Games Recommender:** Recommends games based on user's specified PC
 - **PC Recommender:** Recommends PC components based on user's favorite games
- USED:** Django, XGBoost, Pandas, Scikit-Learn, Beautiful Soup, Requests

HONORS, AWARDS

- **Fulbright Scholar** MS in Artificial Intelligence / Rochester Institute of Technology Sep 2023
(I'm supposed to attend RIT in Fall 2024, but due to VISA delay, my scholarship may be revoked)
- **Topper** Computer Science Class of 2021 / University of Narowal July 2021
- **Kaggle Expert** / Kaggle Feb 2023

TWITTER ACHIEVEMENTS

- "You can" reply by **Yann LeCun** ([tweet link](#)) Nov 2023
- Replied by **Yann LeCun** June 2023

ACTIVITIES

I'm always in the pursuit of learning. I occasionally participate in Kaggle Competitions and blog about Machine Learning on medium. I love books, coding, gaming, F1 and jogging.