Tamer Ajaj

Berlin, Germany
tamer.ajaj@gmail.com
tamerajaj.com/projects-list
linkedin.com/in/tamerajaj
github.com/tamerajaj

Summary

Experienced Data Scientist with a background in Computational Neuroscience and a passion for cutting-edge AI technologies. Proficient in MLOps and Machine Learning Engineering, ensuring automation, reproducibility, and model monitoring. Skilled in Python, TensorFlow, Keras, and developing predictive models for medical, neural, and behavioral applications. Collaborative team-player contributing to joint projects. Excels in acquiring new knowledge swiftly, implementing goal-oriented solutions. Diverse background in academia and industry, offering interdisciplinary abilities and strong analytical skills.

Technical Skills

Machine Learning: TensorFlow, Keras, Scikit-learn, NumPy, Pandas, SciPy, Matplotlib.
MLOps and Cloud: Docker, Flask, MLflow, Google Cloud Platform (GCP).
Data Engineering: PySpark, DBT Cloud.
Monitoring and Automation: MLflow, CI/CD, FastAPI, Prometheus, Grafana, Prefect, GitHub Actions.

Practical Experience

Machine Learning Engineering Boot Camp
6/2023 – 7/2023
neuefische GmbH - School and Pool for Digital Talent — *Hamburg, Germany (Remote)*

- Intensive training boot-camp (180 hours of programming). Topics in four weeks: Software Engineering, Data Engineering, ML Model Deployment, Monitoring, and Alerting, as well as Machine Learning Testing. The training aims to bridge the gap between ML modelling and building real-world ML systems.
- **Software Engineering Project:** Refactored Jupyter Notebooks into object-oriented programming (OOP) programs using Python, built a pipeline for data cleaning and feature engineering, developed a FastAPI app for CRUD operations on house data, and deployed the app in a Docker container. Repo: https://github.com/tamerajaj/w1d5-mle-refactoring-project.
- **Data Engineering Project:** Developed a comprehensive data pipeline project involving ETL and ELT processes, PySpark transformations on Google Cloud Dataproc, loading data to BigQuery, orchestrating with Prefect, and using DBT cloud for modeling. Repo: https://github.com/tamerajaj/w2d5-mle-data-pipeline-project.
- **Machine Learning Engineering Project:** Developed and deployed a machine learning model using scikit-learn's random forest regressor to predict trip durations on the yellow taxi dataset, incorporating API deployment with FastAPI, tracking with MLFlow, and deploying on Google Compute Engine VMs. Repo: https://github.com/tamerajaj/w3d5-mle-model-deployment-project.
- **MLOps Project**: Implemented a structured MLOps workflow for model training, deployment, and monitoring using GitHub Actions, Docker, and Google Cloud Platform. Deployed the model through FastAPI and monitored its performance using Prometheus, Evidently, and Grafana, ensuring efficient automation and seamless integration throughout the process. Repo: https://github.com/tamerajaj/w4d5-mle-mlops-project.

Data Scientist and Project Manager as PhD (TV-L E13 100%)
10/2021 – 12/2022
Max Planck Institute for Human Development — *Berlin, Germany*

Data Scientist | *Django, JavaScript, React, Python, GPT-J, Pandas, R, Matplotlib/Seaborn, oTree.*

- Developed a gamified website using oTree, Django, and JavaScript for a social game with AI. Collected post-game questionnaires and gathered open-ended opinions on participant experience. [Link accessible here.]
- Led the design and deployment of an online experiment website with over 1000 participants to investigate the impact of machine learning design on moral decision-making and honesty.
- Trained a language model to train machine learning algorithms using users' text commands.
- Conducted data collection, analysis, and visualization utilizing Python, R, Matplotlib, and Seaborn.
- Effectively communicated results through presentations and publications.

Project Manager

- Managed a Django and React web development team for a project examining the impact of blurring participants' faces during video conference calls on empathy and sharing behavior.
- Ensured alignment between psychology/sociology researchers' requirements and the software's technical capabilities in a product management/owner role.
- Facilitated smooth collaboration and information exchange as the primary point of communication between web developers and cross-departmental researchers.
- Developed comprehensive requirements and a detailed timeline for effective planning and execution of different project stages.

Software Developer for Machine Learning in Neural Data 02/2017 – 06/2020
Machine Learning Group, Fraunhofer-HHI (Part time) *Berlin, Germany*

- *Tools: Python, MNE-Python, sklearn, TensorFlow, Keras, LabStreamingLayer, VR, Neural recordings, LSTM, Data visualization.*
- Successfully developed a novel software solution for synchronizing data recordings between multiple systems of neural and movement recordings, and virtual reality streams.
- Developed a patented algorithm for VR video compression using head movement prediction and neural data.
- Conducted data collection, analysis, and visualization and communicated results in writing, presentations, and publications.

Education

Max Planck School of Cognition Germany
PhD Candidate *2020 – 2022*
- Focus: Human-AI interaction, Statistics and Experimental Design, AI.

HU Berlin, TU Berlin Berlin, Germany
M.Sc. in Computational Neuroscience *2016 – 2019*
- Focus: ML for Neural Signal analyses, Image Perception.

Damascus University Damascus, Syria
B.Sc.Eng. in Electronics and Telecommunication *2009 – 2016*
- Focus: Signal processing.

Training

Machine Learning Engineering (neuefische GmbH) 06/2023 – 07/2023
- Link: neuefische.de/en/bootcamp/machine-learning-engineeringp

MLOps Zoomcamp (DataTalksClub) 05/2023 – Present
- Link: github.com/DataTalksClub/mlops-zoomcamp

ML Eng. for Production (MLOps) Specialization (Coursera) 03/2023 – 05/2023
- Courses:
  Introduction to Machine Learning in Production.
  Machine Learning Data Lifecycle in Production.

Deep Learning Specialization (Coursera) 2021

Projects (Full List: academic | Github)

Python Telegram bot (Personal Project) | *Python, Telegram, Docker, CI/CD, Git.* 05/2023
- A telegram bot to communicate the availability between two team members.
- Used Git Workflow for linting and deployment.
- Deployed on Fly.io with persistent memory.

Face detection with eye tracking and brain recordings (PhD Internship) 09/2021 – 12/2021
- *Python, LabStreamingLayer, Eye tracking (PupilLabs), video processing.*
- Implemented an algorithm that detects face recognition events using a combination of eye tracking, brain recordings, and image processing.

Evolution of trust under ambiguity (PhD Internship) | *Python, JavaScript, jsPsych.* 09/2020 – 12/2020
- Deployed, designed and implemented an online experiment to examine the impact of ambiguity on trust development.
- Analyzed data to understand the effects of ambiguity on trust-building behaviors.
- Demonstrated skills in experimental design, data collection, and analysis in the context of trust and ambiguity.

EEG-based Image Quality Assessment (Thesis) | *Python, sklearn, TensorFlow, EEG.* 06/2019 – 10/2019
- Conducted a master's thesis project analyzing EEG neural data to objectively evaluate visual perceptual quality of complex natural images.
- Utilized Python, LabStreamingLayer, Brainvision EEG recordings, and ML libraries (sklearn, TensorFlow) for data analysis and modeling.
- Demonstrated the potential of EEG as a tool for assessing objective image quality.
- Presented research findings in a master's thesis, paper, and presentation.
- Presented the research findings in a master's thesis, paper, and large-scale presentation.
- Demonstrated proficiency in signal processing, machine learning, and data analysis techniques relevant to time-series data.

Languages

English: Fluent
German: Conversational
Arabic: Native