Aviv Gelfand

◆ Jerusalem, Israel ✉ avivgelfand@gmail.com ☎ 054-300-7257 in aviv-gelfand ⌂ AvivGelfand

Summary

- 1st year MSc candidate in Data Science (HUJI); Thesis: Dimensionality reduction of Deep Learning problem space through statistical methods and graph theory. BSc in Statistics and Data Science (HUJI; GPA 89).
- Dedicated to long-term growth in a cutting-edge company with a positive atmosphere; available for 80%-100% time role as a Data Scientist / ML or AI Researcher / Engineer / SW Dev.
- Demonstrated success in AI and software development through academic/private projects and roles, with a track record of implementing machine learning models to solve complex problems requiring strong collaboration and rapid self-learning skills, consistently delivering efficient and accurate models.

Experience

NLP Data Scientist / Research Assistant *Oct 2023 – May 2024*
Hebrew University of Jerusalem - Dr. Adi Leibovitch's Quantitative Legal Research lab

- Successfully executed a Supervised-Fine-Tuning training of large language models to classify defendant arguments in trials for a broader statistical legal research project; reduced labeling time and errors, replacing manual tagging with LLM-based tagging.
- Drove strategic alignment by delivering situational analyses and action plans to US-based stakeholders, effectively prioritizing objectives and focusing efforts to meet deadlines.
- Dealt with significant imbalance using embedded cluster stratified sampling methods. Corrected errors in the labeled data. Reached over 90% accuracy and 80% F1 scores in most target labels. Substantially reduced model complexity and training time using efficient sampling. Presented findings at a prominent statistics and data science conference in May 2024.

Statistics & Machine Learning Instructor *Oct 2022 – Sept 2023*
Upscale Analytics

- Built dozens of ML hands-on projects for a Machine Learning and Statistical Inference course, with constant feedback from a Senior Data Scientist (and head lecturer), Michael Winer.
- Instructed data analysts through their first steps in statistical inference and classic machine learning: linear and logistic regressions, ensemble learners such as random forests & XGBoost, and clustering methods.
- Advanced from teaching assistant to instructor, then to a partner role. 3 successful semesters of weekly TA sessions with a significant presence of students.

Research & Development Assistant *Apr 2022 – Mar 2023*
Hebrew University of Jerusalem, Renana Peres's Computational Social Science Lab

- Self-taught and programmed smart-contracts with a C++-based programming language; developed simulations interface using React.js. Designed and analyzed online surveys; conducted literature and academic peer reviews. Web-scraped data monthly.

Data Analyst & Research Team Leader *Mar 2015 – Dec 2020*
Israel Defense Forces (Volunteered, Lieutenant)

- Conducted research, assessments, and analytics for high-ranked decision-makers. Triplicated, updated, and maintained strategic databases while collaborating with programmers and economists.
- Led cross-organizational projects resulting in several two in-depth research papers with an unprecedented scale of collaboration between two research branches.

Education

The Hebrew University of Jerusalem *Oct 2023 – July 2026*
MSc in Data Science — LLMs-focused Thesis, guided by Prof. Ariel Jaffe.

The Hebrew University of Jerusalem *Oct 2021 – Aug 2024*
BSc in Statistics & Data Science (GPA 89)

Skills and Technologies

Scripting Languages and tools: Python, Java, Solidity (C++ based), R, SQL, JS

Programming Libraries: Pandas, PyTorch, Transformers, SKlearn, LLM's SDK (OpenAI, Groq, etc) and more

GenAI Tools: Prompt engineering (COT, etc), Crusur AI, Claude, ChatGPT, Perplexity, a 'Langtalks' fan

EDA & ETL: Basic and complex Visualizations, Statistical Inference and modeling, Cleaning, Encoding / Embedding

ML: Efficient Sampling, Feature Engineering, Parameters Optimization, Evaluation Metrics

CI/CD & Other: GIT, Linux, SLURM Cluster system, GitHub workflows, shell scripts, Big Query (well experienced in SQL), Data-Bricks, Apache Spark

Notable Projects

Fine Tuning Large Language Models *GitHub Repo* ↗
- This repository presents my (Aviv Gelfand) work on applying supervised fine-tuning (SFT) methods to build classifiers for labels as part of a broader project in statistical legal research collaboration led by Dr. Adi Leibovitch and Sharon Levy from The Hebrew University of Jerusalem Faculty of Law, Prof. J.J.Prescott and Grady Bridges from the University of Michigan Law School.

De Grandmizer Shtick! GenAI Hackathon *GitHub Repo* ↗
- De Grandmizer Shtick is a Python application that generates conversations using GPT-3 and converts them into audio files using the Google Cloud Text-to-Speech API.
- The project is designed to help create the illusion of an occupied home to intimidate burglaries and was created during HUJI Generative AI Hackathon 2023.

Driver Behavior Detection Using Smartphone Signals *GitHub Repo* ↗
- A Signal Detection project applying ML to identify risky driver behavior, based on academic publications and open source data-sets.

Alfred - University Tasks Personal Assistant *GitHub Repo* ↗
- A deployed personal assistant software that scrapes the Hebrew University Moodle Site and notifies me via Telegram bot about every change in task deadlines.

For over 30 more projects, check out my GitHub profile. *github.com/AvivGelfand* ↗