Tel Aviv
055 - 6604122
Idan.arbiv@gmail.com
https://www.linkedin.com/in/idan-arbiv
https://github.com/IdanArbiv

Idan Arbiv

Machine Learning Engineer with interests in LLMs, Generative AI, and Deep Learning.

Education

M.Sc. Computer Science/ Ben Gurion University/ GPA - 97

Deep Learning Researcher under the guidance of Dr. Omri Azencot, I focus on **Deep Learning** in the areas of **Sequential Modeling, Generative Modeling, and Representation Learning** utilizing technologies such as **PyTorch, Keras, Scikit-learn, Hugging Face, Optuna, Pandas, spaCy, OpenCV, FastAI and Neptune.**

Publications

- *N. Berman*, I. Naiman*, **I. Arbiv** *, G. Fadlon *, and O. Azencot. Sequential Disentanglement by Extracting Static Information From A Single Sequence Element. International Conference on Machine Learning (**ICML**), 2024*
- *N. Berman*, I. Naiman*, **I. Arbiv**, G. Fadlon, Itai Pemper, and O. Azencot. Utilizing Image Transforms and Diffusion Models for Generative Modeling of Short and Long Time Series. Neural Information Processing Systems (**NeurIPS**), 2024*

B.Sc. Computer Science/ Ben Gurion University /GPA – 93 (Summa Cum Laude, Dean's List)

Experience

AUG 23 –PRESENT

Algorithm Engineer– Machine Learning & Data Science/ SONY Semiconductor

- Designed and developed a cutting-edge **LLM-based** web application using **advanced RAG** techniques to address the challenges of searching and extracting 3GPP **technical documents**, providing a scalable solution for over **100 engineers.**
- Implemented **fusion-based indexing** using **BM25** and **FAISS** vector stores for efficient and accurate document retrieval, integrating **deep learning encoders** and **Cohere reranker** for enhanced accuracy.
- Leveraged **LangChain** and **Graph RAG** with a knowledge graph to connect and contextualize different document chunks.
- Utilized the **ELK stack** to manage and monitor document indexing, search performance, and overall system health.
- Integrated **RAGAS** for evaluation pipelines, deployed with **Streamlit**, and automated **CI/CD** processes using **TeamCity**.
- Enhanced system functionality with **GenAI agents**, adding **conversational memory**, multi-question **routing** with LLMs.
- Created a Python package for **interactive plot visualization**, leveraging **Matplotlib** to streamline data analysis and exploration.

OCT 23 –PRESENT

Teaching Assistant/ Ben Gurion University

- CS 202.1.2051 - Principles Of Programming Languages (Spring 2023, 2024)
- CS 202.1.5391 - Distributed Systems Programming (Winter 2023, 2024)

MAY 22 –AUG 23

Software Engineer (Student)/ SONY Semiconductor

- Developed scalable, testable, and maintainable code for **cellular IoT desktop applications** using **OOP, SOLID principles**, and **design patterns** in **Java**, managing all stages from design to production with **Jira, Git, and Bitbucket**.
- Developed new software functionalities using **Java and Spring Boot framework** on the server side, as well as **databases management**, leveraging my understanding of various algorithms.
- Worked on the client-side using **React and Redux**, creating dynamic and responsive UI components that enhanced user experience while effectively interacting with **RESTful** APIs and employing analytical tools for optimal results.

NOV 15 – AUG 20

Communications (C4I) Officer - Captain / IDF

- Led **communication** and technology systems for combat and special units, **commanding 40+ soldiers, 15 officers, and 5 noncommissioned officers**. Demonstrated strong leadership in diverse operational missions, including special assignments abroad, earning an **excellence certificate** for exceptional performance.