Taha Binhuraib

@taha.huraib@gmail.com |
 \fbox LinkedIn | \diamondsuit GitHub |
 \diamondsuit Google Scholar | \diamondsuit Atlanta, Georgia

EDUCATION

Georgia Institute of Technology	Atlanta, Georgia
P.h.D. in Psychology	2024-
Bahcesehir Univeristy	Istanbul, Turkey
B. Sc. in Industrial Engineering; GPA: 3.80/4.00 , Graduated with high honors as the second-ranking student in the department	2017 - 2021
Work Experience	
Language, Intelligence and Thought (LIT) Lab	Atlanta, Georgia Aug 2024 – Current
Developing interpretable models for encoding language in the brainUnderstanding core algorithmic design choices in Topographic Language Models	
Vision Lab, Istanbul Technical University	Istanbul, Turkey
Research Assistant	Mar 2023 – June 2024
• Published the paper titled Protodiffusion: Classifier-free diffusion guidance with protot	ype learning at ACML 2023
• Developed model demonstrated superior performance in the early stages of training commethod.	mpared to the baseline
Novus Technologies	Istanbul, Turkey
Machine learning engineer	Oct 2021 - July 2024
• Novus is a Boston-based start-up that conducts research on Large Language Models (I applications.	LMs) and builds industrial
• Worked on projects for companies using open-source LLMs and techniques such as SFT experimenting with state-of-the-art research papers.	Γ and RAG. Implementing and
• Worked on translation for low-resource languages using LLMs.	
• Built a mini langchain that supports different inference platforms such as vLLM.	
Gama Energy	Istanbul, Turkey
Data Scientist	Sep 2021 - Feb 2022
• Developed a fully automated pipeline for different regression and classification tasks.	-
• Provided data-driven statistical analyses on how different factors affect different energy	v markets.
Towards Data Science	
Technical Writer	Sep 2021 - Feb 2022

- Diffusion Models Demystified
- The Essence of RNNs
- NLP with CNNs
- An Introduction to Linear Algebra for Deep Learning

From Language to Cognition

• B AlKhamissi, G Tuckute, Y Tang, **T. Binhuraib**, A Bosselut, M Schrimpf, "From Language to Cognition: How LLMs Outgrow the Human Language Network", **CCN 2025 (Proceedings Submission)**

TopoLM

• N Rathi, J Mehrer, B AlKhamissi, **T. Binhuraib**, NM Blauch, M Schrimpf, "TopoLM: brain-like spatio-functional organization in a topographic language model", **ICLR 2025(Oral)**

Topoformer

• T. Binhuraib, Tuckute, G., NM Blauch, "Topoformer: brain-like topographic organization in Transformer language models through spatial querying and reweighting", ICLR Re-Align 2024

Protodiffusion: Classifier-free diffusion guidance with prototype learning

• G. Baykal, H. F. Karagoz, **T. Binhuraib**, and G. Unal, "Protodiffusion: Classifier-free diffusion guidance with prototype learning", **ACML**, **2023**

Proposing a Novel Artificial Neural Network Methodology for Forecasting Risk of the COVID-19 Pandemic

• **T.Binhuraib**, GT Temur, "Proposing a Novel Artificial Neural Network Based Methodology for Forecasting Risk of Covid-19 Pandemic" **Proceeding of IDS 2020, 2023**

Studying the Neurological Differences between Winning and Losing a Gamble | GitHub

- A study in which fMRI data was used to study the neurological effects of gambling.
- Generated Statistical mappings of the brain using generalized linear models.
- Developed logistic regression and deep learning encoding models.

FastAPI Application with Sentence Transformers and Pinecone | GitHub

- Enabled users to perform similarity searches on the stored embeddings using Pinecone's search functionality.
- Developed an application that utilizes pre-trained models from Sentence Transformers to encode text.

Price Forecasting and System Direction Determination in the Balancing Power Market. | GitHub

PyTorch GANs | GitHub

• Implemented the original Generative Adversarial Networks paper using PyTorch.

BART Morphology | GitHub

• Trained a BART language model from scratch for an inflection task.

Skills

Programming: Python, C, C++, Rust, SQL, Go
Technologies: Git, Docker, AWS, Gradio, Streamlit, Google Cloud, Azure
Web Development: Flask, Django, FastAPI
Machine Learning Scikit-learn, PyTorch, TensorFlow, Jax, Tranformers(HF)
Languages: Arabic (Native), English (Fluent), Turkish (Native)