

Mory Gharasuie

Norfolk, VA, USA / mmoha014@odu.edu / +1 757 287 1602

<https://www.linkedin.com/in/mory-gharasui-53415258/> / <https://github.com/mortezamg63>

Technical Skills

Languages & databases: python, Java, C++, ASP Webform, C#, SQL, MySQL, HTML

Libraries & APIs: Tensorflow, Keras, PyTorch, OpenCV, Scikit-learn, NLP toolkit, HuggingFace, Pandas, Matplotlib, Seaborn, LangChain, Pandas, Dask, BeautifulSoup

Development tools: Anaconda, Jupyter Notebook, Google Colab, Visual Studio, Git, Docker

Operating Systems: Windows, Linux, Mac OS X

Experience

Royan Communication Company | Qom, Iran

2013 - 2019

Software Developer

- Developing websites for small and medium-sized enterprises.
- Customizing web-based administration interfaces for applications on Linux server machine (such as chat server, FreeRadius server, and Elastix).
- Enhancing the panels with support for multiple languages and designing them to be more intuitive and user-friendly, tailored to meet the specific needs and preferences of the customers.

Old Dominion University | Norfolk, USA

Aug 2019 – Present

• *Research Assistantship*

- Developing applications for mobile and serverless domains by leveraging ML, DL and CV.
- Doing research on Improving the performance of ML and DL models on classification problems for tabular data in SSL setting.
- research on mitigating the impact of bias in imbalanced data in training ML and DL models in Image and tabular domains.

• *Teaching Assistantship*

- Programming with C/C++ and Java (CS150, CS250, CS251)
 - Teaching Labs and recitations
 - assignment Development
 - Grading

Medical Aid | Norfolk, USA

Summer 2024

- Collaboration in Developing a ChatBot utilizing Large Language Models (LLMs) and Retrieval-Augmented Generation.
- Medical data extraction with reference to papers or resource
- Presentation of results based on standard medical format
- Providing relevant questions or considerations from recent papers for better diagnosis

Education

Old Dominion University | Norfolk, USA

Aug 2019 – present

Research Interests: Machine Learning (ML), Computer Vision (CV),

GPA: 3.84/4.0

Natural Language Processing (NLP), Semi-supervised Learning (SSL), Tabular Data Processing

University of NabiAkram | Tabriz, Iran

Master of Science in Computer Engineering

University of Shamsipoor | Tehran, Iran

Bachelor of Science in Computer Engineering

Projects

Data Science and Machine Learning Projects | *Tensorflow, PyTorch, TorchVision, OpenCV, Scikit-learn, Seaborn, Matplotlib, Pandas, Git, BeautifulSoup*

[GitHub](#)

Created and managed a repository featuring data science and machine learning projects. These projects involve working with various datasets, addressing challenges suitable for beginner to intermediate levels in data science. The repository includes:

- Data processing and preprocessing
- Exploring data analysis (EDA)
- Feature engineering and selection
- Machine learning development and evaluation
- Data visualization and interpretation

Large Language Models | *HuggingFace, Pytorch, Pandas, Dask, LangChain, PEFT, RAY, OpenAI APIs*

[GitHub](#)

In this repository, I investigate the realm of retrieval-augmented generation (RAG) systems, embedding models, prompt engineering, and fine-tuning large language models. Throughout this journey, I am learning about Generative AI and leveraging some embedding models for practical applications. The repository showcases my ongoing exploration in this exciting field and will be continually updated with new insights and findings.

Pricer (Agentic LLM) | *HuggingFace, Pytorch, LangChain, Chormadb, BeautifulSoup, Scikit-learn*

[GitHub](#)

An autonomous price estimation framework using LLMs (GPT-4o, Claude, Llama 3.1-8B) and traditional ML models (Random Forest, SVM, Word2Vec). GPT-4o-mini's performance (average price difference) with RAG, improves from 80.9 to 55.57. I fine-tuned Llama3.1-8B and achieved 46.67 average error. I developed an agent that creates an ensemble model combining RAG+GPT-4o-mini, fine-tuned Llama, and Random Forest, achieving 54.62 error. I use a Gradio-based UI and integrated Pushover API for real-time deal alerts from DealNews.Com.

SAWTAB: Tabular Data and Semi-Supervised Learning (SSL) | *Pytorch, Tensorflow, Tensorboard, Scikit-learn, Matplotlib, Seaborn*

This research project focuses on addressing two significant challenges in machine learning with large tabular datasets: class imbalance and the difficulties associated with one-hot encoding for high-cardinality categorical features. To overcome these challenges, a target encoding method is proposed for a Semi-Supervised Learning (SSL) setting. The method improves target encoding by leveraging unlabeled data through pseudo-labels and adaptively adjusting their influence to minimize the impact of noisy pseudo-labels. Experimental results on various datasets, compared against multiple frameworks, demonstrate that the approach achieves superior generalization performance.

Publications

PAKDD 2024: SAWTab: Smoothed Adaptive Weighting for Tabular Data in Semi-Supervised Learning | [Link](#)

ICKG 2022: Progressive Feature Upgrade in Semi-supervised Learning on Tabular Domain | [Link](#)

BodySys 2021: Performance Monitoring for Exercise Movements using Mobile cameras | [Link](#)

MVIP 2015: An efficient run-based method for connected component labeling | [Link](#)

MVIP 2013: Real-time dynamic hand gesture recognition using hidden Markov models | [Link](#)