Pouyan Nahed

+1 (702) 351 4042 | nahed@unlv.nevada.edu | pouyannahed.com | github.com/pouyan9675 | linkedin.com/in/pouyan-n-508a4586 My main interests are in field of Machine Learning and Natural Language Processing

Experience

Ancestry, Data Scientist Intern/Co-Op | Lehi, UT

- Collaborated with the Handwriting Recognition team to develop AI solutions for processing ancient handwritten documents. extracting key genealogical information.
- Designed and optimized Vision-Language Models (VLMs) for end-to-end structured information extraction, achieving a 25% reduction in inference cost and time.
- Gained hands-on experience in application development and deploying deep learning models on cloud platforms such as AWS, designing and implementing scalable solutions for large-scale handwritten document processing workflows.

University of Nevada - Las Vegas, Research Assistant (Prof. Kazem Taghva) | Las Vegas, NV

- Designed and developed an open-source data analysis platform called Tri-AL: VisuAL ClinicAL TriAL using the Django framework to visualize and keep track of Alzheimer's Disease trials posted on clinicaltrials.gov
- Large Language Model quantization and serving to extract key values from trials and use these entities to fine-tune a text gener**ation** model for trial summarization.
- Currently, working clinical trials summarization using Gen-AI and Large Language Models and their potential usage in Drug Repurposing.

University of Nevada - Las Vegas, Teaching Assistant | Las Vegas, NV

 Worked as a Teaching Assistant for several courses: Advanced Natural Language Processing, Analysis of Algorithms, Data Structures, Database Management Systems and Automata and Formal Languages

Skills

Programming	Python, Java, CUDA, Git, Scripting (Bash), LaTeX, HTML, JavaScript, Debugging
Machine Learning	Data Visualization, Reinforcement Learning, Experiment Tracking, Quantization, Statistics
Tools	Tensorflow, Pytorch, Docker, OpenCV, Huggingface, AWS SageMaker, Google Cloud Platform, NumPy, Pandas

Education

4.0/4.0 PhD in Computer Science, University of Nevada, Las Vegas | Nevada, USA

15.6/20 BS in Computer Engineering, University of Guilan | Guilan, Iran

Achievements: UNLV Access Grants (2020-23) | Member of the Scientific Association for Computer Science @ University of Guilan Courses: Graph Data Mining | Advanced Machine Learning | Advanced Analysis of Algorithms | Introduction to Natural Language Processing

Publications

Tri-AL: An open source platform for visualization and analysis of clinical trials

Information Systems

- Developed an open-source platform to address functionality gaps in Clinical Trials.gov, including data collection, visualization, reporting, and historical analysis tools.
- Designed a programmable module for incorporating machine learning models to extract disease-specific data from unstructured trial reports.
- Demonstrated the use of Tri-AL for Alzheimer's disease reporting and trial participation analysis by sex, gender, race, and ethnicity.

Enabled target users (scientists, researchers, pharmaceutical companies) to explore over 440,000 medical studies more effectively. Enhancing Clinical Trial Summarization: Leveraging Large Language Models and Knowledge **Graphs for Entity Preservation**

International Congress on Information and Communication Technology

- Developed a summarization model using Encoder-Decoder Transformers to generate concise clinical trial summaries, improving efficiency for researchers and healthcare professionals.
- Proposed a clean high quality dataset that improve model performance, achieving up to a 23% improvement in ROUGE scores outperforming baseline methods.
- Used a graph-based evaluation metric to assess entity preservation in clinical trial summaries, ensuring critical medical entities remains intact

A Recommendation Model for Predicting Alzheimer's Drugs' Mechanism of Action

Intelligent Sustainable Systems - Selected Papers of WorldS4

- Developed a machine learning model using BioBERT and a deep neural network to classify the mechanism of action (MoA) for Alzheimer's drugs, achieving a 97% F1-score.
- Built a structured dataset from clinical trials and ALZFORUM therapeutics data, enabling efficient drug development analysis and decision-making.

Alzheimer's disease drug development pipeline: 2022

Alzheimer's & Dementia: Translational Research & Clinical Interventions

- Analyzed the drug development pipeline for Alzheimer's Disease, highlighting key trends and novel therapeutic approaches.
- Collaborated with a multidisciplinary team of clinicians to provide insights into translational research and clinical interventions.

Clinical text classification of Alzheimer's drugs' mechanism of action

International Congress on Information and Communication Technology

Developed a interpretable clinical text classification system for identifying mechanisms of action in Alzheimer's drugs.

2022

June 2023 - Present

Jan 2020 - Present

Jan 2020 - Present

2025

2018

2025

2025

2022

Projects

Context Extraction Attack in RAG Systems

Private & Secure AI Course Project

- Developed and evaluated **adversarial attacks** on Retrieval-Augmented Generation (RAG) systems, targeting context extraction to expose system prompts and sensitive embedded information.
- Synthesized a custom minimal dataset using generative models to simulate scenarios and test attacks, achieving improved context retrieval metrics (e.g., ROUGE and entity sensitivity levels).
- Demonstrated potential vulnerabilities in RAG systems, highlighting the critical importance of **trust and safety** in machine learning models and proposing potential mitigations to reduce risks of unauthorized data access and manipulation.

Clinical Trials Summarization

Research Project

- Trained a **generative model** with clinicaltrials.gov data to summarize clinical trials effectively and summarize information by preserving entities within the text.
- Enhanced summarization accuracy through filtering and cosine similarity-based tf-idf vectorization during extractive summarization model development.

Minimal Search Engine

Personal Project

- Implemented an efficient search engine powered by **BERT**-based document indexing, with vector embeddings stored in **REDIS** for swift and precise document retrieval, optimizing search performance.
- Designed and implemented a simple UI to list files and show the results with a graphical interface.
- Designed a **Retrieval-Augmented Generation (RAG)** module to provide a summary of the results using a Small Language Model (SLM).

Tri-AL: VisuAL ClinicAL TriALs

Research Project

- Developed a **real-time visualization framework** for clinicaltrials.gov data and help clinical trial researchers to explore trials across the world and keep track of their status.
- Utilized **Django** framework along with Python visualization tools such as Plotly to build different types charts.
- Designed and implemented an **extendable module** for the system to enable Computer Scientist integrate their Machine Learning models into the system for data analysis and missing value predictions.

Deep FlaPyBird

Personal Project

- Implemented Deep Reinforcement Learning using PyTorch in Python to train an agent for playing the Flappy Bird game.
- Developed a custom simulation environment within the **OpenAI Gym** framework, utilizing it to define states and actions for training the model.
- Designed a computer vision Deep Neural Network architecture optimized for the Flappy Bird environment, allowing the agent to learn complex policies for navigating the game successfully.

Jan 2021 - May 2022

Sep 2023 - Ongoing

Sep 2024

Jan 2021