Soudeh Nilforoushan

+1 6476553294 Toronto, Canada
nilforoushan.soodeh@gmail.com or snilforo@uwo.ca
https://www.linkedin.com/in/soudeh-nilforoushan/ https://github.com/soodeh-nilforoushan

Research Interest

Data Management
Data Mining
Natural Language Processing
Deep Learning
Machine Learning

Education

Western Ontario University, London, Canada **January 2021 - September 2022**
MSc in Computer Science — fully-funded Scholarship *88.5/100 GPA*

- Thesis Title: Improving Deep Entity Resolution by Constraints
- Supervisor: Dr. Mostafa Milani

Amirkabir University of Technology (Tehran Polytechnic), Iran **September 2014 - October 2019**
BSc in Computer Engineering *16.80/20 GPA*

- Project Title: Implementing an object tracking software
- Supervisor Name: Dr. Ahmad Nickabadi

Publications

S. Nilforoushan, Q. Wu, and M. Milani, "Entity matching with AUC-based fairness", *2022 IEEE International Conference on Big Data*, IEEE, 2022, pp. 5068– 5075.

Research & Work Experience

Software Engineer **September 2023 – Jan 2025**
Pelypay International *Toronto, Canada*

- Collaborated with data scientists and engineers to design and implement machine learning pipelines for fraud detection, reducing fraudulent transactions by 30%.
- Designed and developed a fraud detection system with a React-based front-end and a Spring Boot back-end, integrating RESTful APIs for real-time transaction monitoring.
- Optimized SQL queries for a fraud detection project, cutting fraudulent transactions by 30% and enhancing security.

Software Engineer **June 2022 – September 2023**
The Home Depot *Toronto, Canada*

- Achieved a notable 50% enhancement in the runtime speed of a data pipeline project by strategically migrating an App Engine project to Google Cloud Platform (GCP).
- Boosted project efficiency by 20% through Airflow, GCP, and Python in a data engineering team.

Machine Learning Researcher **January 2021 – September 2022**
Western Ontario University *London, Canada*

- Applied fairness constraints and introduced a new fairness metric on a *deep learning entity matching* problem, which tended to be biased towards marginalized classes.
- Enhanced a Large Language Model (LLMs) by leveraging state-of-the-art NLP techniques, including BERT.
- Fine-tuned the model's parameters and architecture to significantly improve accuracy and performance in identifying and eliminating duplicate data.
- Presented more than 15 papers regarding *data cleaning*, *data privacy*, and *machine learning* to my research group.

Researcher & Data Analyst **January 2019 – June 2020**
Brain Engineering Center, Institute for Research in Fundamental Sciences (IPM) *Tehran, Iran*

- Gathered information with EyeTracker for the movement of human pupils and analyzed data.
- Launched a psycho-physics task to enhance human attention to an advertisement video by making imperceptible changes, such as adding a light-colored border around the advertisement target.

Technical Skills

Languages: Python, Java, Scala, C++
Frameworks, platforms, and libraries: LangChain, Hugging Face, LLaMA, PyTorch, Airflow, Pandas, Scikit-learn, Matplotlib, NLTK, NumPy, Transformers, TensorFlow, Git, GCP, AWS, Spring Boot.
Database systems: MySQL and SQL Server, Data Management
Operating Systems: Linux, MacOS, Windows.

Teaching Assistantships

Western University

- Advance Programming (Java), Winter 2021- Spring 2021-Fall 2022
- Data Analysis, Winter 2022

Amirkabir University

- Database Management Systems, Fall 2018
- Computer Architecture, Fall 2017

Courses (Western University)

Course Title	Instructor(s)	Grade
Advanced Machine Learning	Charles Ling	A
Introduction to Data Science	Charles Ling	A+
Advanced Database Management Systems	Mostafa Milani	A+
Information Visualization	Kamran Sediq	A

Related Projects (Github)

- Developed a chatbot using large language models for robotic grasping and manipulation under the supervision of Dr. Kasaei.
- Developed a chatbot using the LangChain library and the LLaMA model. The chatbot is designed to answer general questions and analyze text from various file formats, including PDFs and Word documents.
- Completed retail data analysis with Walmart data and loan data analysis projects during the **Simplilearn** bootcamp using Python, Keras, and Tensorflow.
- Implemented Twitter airline sentiment analysis to evaluate and compare different Machine Learning and deep learning techniques.
- Implemented the basics of the LangChain library using Ollama, OpenAI, and Hugging Face to develop a natural language processing application that can generate text, understand context, and interact with users more effectively.
- Built language model using transfer learning.

Honours and Awards

- Accepted with full funding ($50000) for MSc studies at Western University. 2021
- Recognized as an *Exceptional Talent* of the Amirkabir University of Technology. 2015
- Ranked *top 0.05%* in the National Entrance Exam among all Iranian Students in Mathematics & Physics (approximately 230000 applicants). 2014