Mansour Saffar

Machine Learning Engineer

+1 (587) 937 0770 | msaffarm.github.io | linkedin.com/in/msaffarm | msaffarmehrjardy@gmail.com

Work Experience

05/22 - Present	 Senior Machine Learning Engineer (Unity Ads) Working as a senior MLE/engineering tech lead in the app-event team (9 members). My duties included sprint planning for engineering, developing and optimizing backend/data/ML services, and writing/reviewing design documents where we developed multiple app-event campaign types for the Unity Ads network. Developed and optimized end2end pipelines including data ingestion pipeline, streaming/batch data processing pipeline, data validation pipeline, model training pipeline, model serving pipeline/microservice, and data/model monitoring pipelines to support app-event campaigns. Technologies: Python, Go, Tensorflow, Kubeflow, Airflow, PySpark, BigQuery, BigTable, Docker, Kubernetes, GKE, Dataproc, Composer, Kafka, Prometheus, Grafana, Terraform, Jira, Jenkins, GitHub Actions
01/22 - 05/22	Lead Machine Learning Engineer AltaML, Toronto • Served as ML Solution Architect and team lead in multiple machine learning projects. • Designed the architecture and lead a team of 5 ML devs to build a customizable object detection solution built on top of Azure.
03/21 - 12/21	 Senior Machine Learning Engineer Lead multiple teams of machine learning engineers (2-3 team members) in multiple projects, including recommender systems and time-series forecasting. Designed the architecture and implemented a machine learning model deployment pipeline built on top of AWS and Kubernetes (EKS). This pipeline is used internally to deploy ML models.
03/18 - 03/21	 Machine Learning Engineer Developed NLU modules (intent recognition and entity detection), extractive text summarization, Ad-Words generation, keyword extraction, and OCR in various NLP projects. Developed an automatic medical report generation system from technician reports using transformer-based models with 87% similarity with radiologists' writings. Developed a human-in-the-loop content-based recommender system to help content editors find the most relevant articles while using the web CMS with an F-score of 80%. Designed and developed multiple task-oriented chatbot projects including synthetic data generation tool for NLU and dialogue management, chatbot to answer complex natural queries from a financial database, customer service and FAQ chatbots. Technologies: Python, PyTorch, Tensorflow (tensor2tensor), Tf-Serving, Rasa, spaCy, NLTK, Gensim, Scikit-learn, MongoDB, Pandas, Dask, PySpark, AWS, Azure, Git, DVC, Docker, Kubernetes, MLflow
05/17 - 08/17	Data Analyst Finning Canada, Edmonton • Developed regression models for rental machinery residual value prediction using ensemble models. • Technologies: Python, C++, Pandas, Scikit-learn, H2O, XGBoost, LightGBM, Azure ML, MySQL
Education	

 2016 - 2019
 M.Sc in Computer Science (GPA: 4/4)
 University of Alberta, Edmonton

 • With a focus on Machine Learning, Natural Language Processing and Chatbots
 University of Alberta, Edmonton

 2011 - 2016
 B.Sc in Electrical Engineering (GPA: 3.67/4)
 University of Tehran, Tehran

 • With a focus on Machine Learning and Medical Image Processing
 University of Tehran, Tehran

Select Projects

09/17 - 11/18	Deep Learning Models for Task-oriented Chatbots	Graduate Research Assistant (Master's Thesis)	
	• Researched usage of self-attentional models for training end-to-end task-oriented chatbots. The results		
	showed faster training with comparable accuracy. [Source Code] [Publication Link]		
	 Developed ChatSim, an architecture agnostic evaluation framework for task-oriented chatbots that can model user characteristics and behaviour in chatbot evaluation. [Source Code] [Publication Link] Technologies: Python, Tensorflow (tensor2tensor), Rasa, spaCy, Git 		
09/16 - 11/16	Retinal Image Segmentation • Developed a segmentation model by applying ensemble and the problem of small dataset size, we achieved good results by	Machine Learning Course SVM models on retinal images. Faced with using bagging methods. [Report Link]	

Technical Skills

• Technologies: Python, MATLAB

Languages	Programming Languages: Python (5+ years), Go (1 year), C++ and Java (familiar)
ML/DL/XAI	Machine Learning, Deep Learning, and Explainable AI: Scikit-learn, H2O, PyTorch, Tensorflow, SHAP
NLP	Natural Language Processing and Conversational AI: Transformers, spaCy, NLTK, Gensim, Rasa (Core & NLU), ParlAI, FairSeq, Tensor2tensor
Data	Big Data Analysis Frameworks and Databases: MySQL, BigTable, BigQuery, Pandas, Dask, MongoDB, Redis, PySpark, Kafka
Cloud	Cloud Computing Platforms: AWS (Certified ML-Speciality, EKS, Lambda, SageMaker), Azure (Azure ML, App Service, AKS), GCP (GKE, BigQuery, BigTable, Composer, Dataproc)
MLOps/DevOps	Machine Learning Model Deployment and MLOps: Docker, Kubeflow, Ariflow, MLflow, Streamlit, Flask, FastAPI, Terraform, Kubernetes, TF-Serving, Grafana, Prometheus, Jenkins, GitHub Actions

Publications

August 2019Mansour Saffar, Amine Trabelsi, Osmar R. Zaiane
• Self-Attentional Models Application in Task-Oriented Dialogue Generation Systems
Recent Advances in Natural Language Processing (RANLP 2019) [Publication Link]February 2019Ghazal Sahebzamani, Mansour Saffar, Hamid Soltanian-Zadeh

Machine Learning Based Analysis of Structural MRI for Epilepsy Diagnosis
 International Conference on Pattern Recognition and Image Analysis (IPRIA 2019) [Publication Link]

Volunteering

October 2020	Public Speaker	Data Science Meetup, Edmonton	
	 Presented a talk, From RNNs to GPT-3, about the progression of deep learning for NLP [Slides] 		
03/19 - 10/19	 Machine Learning Engineer in Ana Project Collaborated with Prof. Zaiane's team in architecture design and intelligent chatbot with emotional capabilities designed to help older 	Joint Project with Amii d developing NLU modules for Ana, an r adults. [Mentioned on CBC News]	
October 2019	Public Speaker	Data Science Meetup, Edmonton	

• Presented a talk about explainable AI (XAI) and its application in industrial ML. [YouTube Video] [Slides]