Saman Khamesian

Graduate Research Associate at Arizona State University (ASU)

Tempe, Arizona, U.S.
in/samankhamesian
s.khamesian@asu.edu
github.com/SamanKhamesian

Education

Arizona State University (ASU) — Tempe, U.S.
Ph.D. Computer Science - Machine Learning — *Present*

Shahid Beheshti University (SBU) — Tehran, Iran
M.Sc. Computer Engineering - Artificial Intelligence and Robotics — *2018 – 2021*

- GPA: 3.88/4.0 (17.59/20) | Total Credits: 32/32
- Master's Thesis Grade: 19.33/20 (Outstanding)

Shahid Beheshti University (SBU) — Tehran, Iran
B.Sc. Computer Engineering — *2014 – 2018*

- GPA: 3.64/4.0 (17.16/20) | Last year GPA: 3.83/4.0 | Total Credits: 142/142
- Ranked 4th in B.Sc. GPA Ranking

Publications

[1] Khamesian, S., Malek, H. Hybrid self-attention NEAT: a novel evolutionary self-attention approach to improve the NEAT algorithm in high dimensional inputs. Evolving Systems (2023).

Status Published in Evolving Systems Journal
Preprint Link Springer
Source Code github.com

Honors and Awards

Ranked 198th in National M.Sc. Universities Entrance Examination — 2018
Among more than 19'000 contestants — *Tehran – Iran*

Qualified for the 19th ICPC Regional (Asia Region) Programming Contest — 2017
Held by Sharif University of Technology — *Tehran – Iran*

Ranked 12th in the 18th ICPC Regional (Asia Region) Programming Contest — 2016
Held by Sharif University of Technology — *Tehran – Iran*

Research Experience

My research interests lie primarily in the area of **Machine Learning**. I also have some experience in:

- Investigating how to improve and optimize ANNs with Evolutionary Algorithms like NEAT - **Master's Thesis**
- Working with **Evolutionary Algorithms** for solving challenging vision-based tasks like playing Atari Games

Work Experience

Mofid Brokerage — Tehran, Iran
Machine Learning Engineer — *Oct 2021 - Jun 2023*

My main focus was on **Time Series Problems**. I also gained experience in:

- Stock Volume Prediction (Time-Series) based on **LSTM networks** and Financial Indicators
- Spoof Detection in **eKYC** services based on Anomaly Detection methods and **Image Processing** techniques

Technical Skills

- **Programming Languages**
  Python | Java | C/C++ | Android | MATLAB
- **Operating Systems**
  Windows (XP/Vista/Seven/Eight/Ten)
  macOS (High Sierra/Mojave/Catalina)
  Linux (Ubuntu/CentOS)
- **Others**
  LaTex | Git | Parallel Programming
- **Machine Learning**
  - Ability to work with popular libraries and frameworks in Python (e.g. TensorFlow-Keras, pandas, torch, numpy, scikit-learn, matplotlib, etc.)
  - Familiar with popular models (e.g. Classifiers, Neural Networks, Deep Learning, Decision Trees, Evolutionary Algorithms, etc.)
  - Familiar with data pre-processing and model evaluation techniques (e.g. Curve Analysis, PCA, Error Metrics, etc.)

Teacher Assistant Experience

Data Structures - **Dr. Abin**	2016 – 2019
Discrete Mathematics - **Dr. Safaei**	2016 – 2018
Algorithms Design - **Dr. Ghavamizadeh**	Spring 2017
Advanced Programming - Java - **Dr. Vahidi**	Spring 2015

Selected Courses

- **Master's Courses**

Advanced Data Mining	18.25/20
Fuzzy Sets and Systems	19.00/20
Machine Learning	18.25/20
Neural Networks	17.10/20
Pattern Recognition	17.00/20
Evolutionary Computing	17.70/20
Digital Speech Processing	18.25/20

- **Bachelor's Courses**

Signals and Systems	20.00/20
Language and Automata Theory	17.75/20
Algorithms Design	20.00/20
Data Structures	20.00/20
Engineering Mathematics	18.80/20
Engineering Statistics and Probability	18.20/20
Advanced Programming	17.00/20

Selected Projects

Time Series Classification for Human Activity Recognition — February 2020

Abstract Human activity recognition is the problem of classifying sequences of accelerometer data recorded by smartphones into known well-defined movements. Classical approaches to the problem involve hand-crafting features from the time series data based on fixed-sized windows and training machine learning models, such as decision trees' ensembles. The difficulty is that this feature of engineering requires strong expertise in the field. Recently, deep learning methods such as recurrent neural networks like LSTMs and variations that use one-dimensional convolutional neural networks or CNNs have been shown to provide state-of-the-art results on challenging activity recognition tasks with little or no data feature engineering instead of using feature learning on raw data. Approximate accuracy is about %88 to %92 in this method.

Project Type Neural Network Final Project

Dataset Human Activity Recognition Using Smartphones Dataset

Language Python (numpy, pandas, keras, and TensorFlow)

The Winton Stock Market Challenge — January 2020

Abstract I implemented a simple LSTM network as a sample solution for solving the Winton stock market challenge. In this competition, the challenge is to predict the return of a stock, given the history of the past few days. Organizers provide 5-day time windows, days D-2, D-1, D, D+1, and D+2. We are given returns in days D-2, D-1, and part of day D, and we are asked to predict the returns in the rest of day D and days D+1 and D+2.

Project Type Advanced Data Mining Final Project

Dataset Provided by Kaggle

Language Python (numpy, pandas, scikit-learn, keras, and matplotlib)

Imputation of Missing Values — July 2019

Abstract I implemented my version of the following method. Missing values in datasets should be extracted from them or estimated before they are used for classification, association rules, or clustering in the pre-processing stage of data mining. The paper uses a fuzzy c-means clustering hybrid approach that combines support vector regression and a genetic algorithm. In this method, the fuzzy clustering parameters, cluster size, and weighting factor are optimized, and missing values are estimated.

Project Type Fuzzy System Final Project

Dataset UCI Machine Learning Repository (Glass - Hamberman - Iris - Musk - Wine and Yeast)

Language Python (numpy, pandas, scikit-learn, and scikit-fuzzy)

Credit Card Fraud Detection — July 2019

Abstract I implemented my version of the following method for detecting credit card fraud. The authors model the sequence of operations in credit card transaction processing using a hidden Markov model (HMM) and show how it can be used for the detection of fraud. An HMM is initially trained with the normal behavior of a cardholder. If an incoming credit card transaction is not accepted by the trained HMM with sufficiently high probability, it is considered to be fraudulent. At the same time, they try to ensure that genuine transactions are not rejected.

Project Type Machine Learning Final Project

Dataset Not officially released. Available on my GitHub repository.

Language Python (hidden-markov, numpy, pandas, and scikit-learn)

Two Level Genetic Algorithm for Clustered TSP — January 2019

Abstract The clustered traveling salesman problem (CTSP) seeks to compute the shortest Hamiltonian tour that visits all the vertices, in which the vertices of each cluster are visited consecutively. A two-level genetic algorithm (TLGA) was developed for the problem, which favors neither intra-cluster paths nor inter-cluster paths, thus realizing integrated evolutionary optimization for both levels of the CTSP. A large-scale traveling salesman problem (TSP) can be converted into a CTSP by clustering so that it can then be solved by the algorithm. In the paper, test results demonstrate that the clustering TLGA for large TSPs is more effective and efficient than the classical genetic algorithm. I implemented my version of this technique.

Project Type Evolutionary Computing Final Project

Language C++/C

Music Genre Classification of Audio Signals — January 2019

Abstract Automatic musical genre classification can assist or replace the human user in this process and would be a valuable addition to music information retrieval systems. In addition, automatic musical genre classification provides a framework for developing and evaluating features for any type of content-based analysis of musical signals. The author of this project uses a support vector classifier (SVC), and I implemented my version of this method.

Project Type Digital Speech Processing Final Project

Language Python (pandas, librosa, numpy, joblib, and scikit-learn)

Languages

English — Fluent

TOEFL iBT Score: 93/120 (Reading 20/30, Listening 26/30, Speaking 23/30, Writing 24/30) — *Oct. 2022*

Persian (Farsi) — Native

▸ NOTE: Hyperlinks are colored blue.