Amirhossein Abaskohi *September 19, 2000*

Senior Undergraduate Student Majoring in Computer Engineering

Website | amirhossein.abaskohi@gmail.com | AmirAbaskohi

amirhossein-abaskohi | Google Scholar | live:.cid.11d29e5a6ca6fb23

Education

University Of Tehran — September 2018 - April 2023

B.Sc in Computer Science — *Ranked top 15% in class*

Department of Electrical and Computer Engineering

- **Cumulative:** 17.94/20.00 (3.8/4) [Last Two Years: 18.42 (4/4)] (Faculty Average: 15.10/20.00)
- **ImportantCourses:**
  - Machine Learning (18.6/20)
  - Artificial Intelligence(20/20)
  - Design and Analysis of Algorithms (20/20)
  - Data Structures (19.8/20)
  - Database Design (18.3/20)
  - Foundation Models in Natural Language Processing (19.1/20)

Motahari High School — 2015-2018

Diploma in Mathematics and Physics

GPA: 19.81/20 Written GPA (19.66 Full GPA)

Research Interests

- Natural Language Processing
- Computer Vision for Medical Imaging
- Cognitive Computing
- Multi-modal Models
- Machine Learning for Health
- Human Centered AI

Research Experience

Undergraduate research assistant at the University of Tehran — Jul 2022 - March 2023

Under Supervision of Prof. Yadollah Yaghoobzadeh and Dr. Sascha Rothe

I am working on the influence of data augmentation methods for contrastive-based prompt learning in the RoBERTa language model.

Undergraduate research assistant at the University of Tehran — Nov 2021 - Jun 2022

Under Supervision of Prof. Azadeh Shakery and Prof. Yadollah Yaghoobzadeh

I was working on creating pseudo-translation documents for the pre-training stage of a multilingual transformer-based language model. We mapped sentences from one language to another by using bilingual dictionaries with distinct monolingual corpora. The quality of mapped sentences is improved by training a model reordering to enhance the quality of sentences for a language.

Undergraduate research assistant — Sept 2021 - Jun 2022

Under Supervision of Prof. Pedram Rooshenas

I developed a VAE-based generative model to simultaneously produce a picture and its caption. We had two goals: 1) Checking the effect of text and image fusion in the latent space on the quality of generative models, and 2) Using the model to create artificial image captioning datasets.

Undergraduate research assistant at the University of Tehran — Aug 2021 - Present

Under Supervision of Prof. Behnam Bahrak

I am working in the Data Analytics Lab of the University of Tehran. I have done several projects, including binary sarcasm classification and Persian emotion detection. I recently started a project on detecting people suffering from BPD based on social media data.

Undergraduate research assistant at the University of Tehran — Nov 2020 - Sept 2021

Under Supervision of Prof. Hadi Moradi

We developed a cognitive test for preschoolers to determine whether or not they are ready for school. At first, as our system was an AI-powered test-taking system, I was working on the website's back end. After that, for the ASR system in speech-based test, we suggested a new pre-train objective for the Wav2Vec model, and we reached a state-of-the-art model for automatic speech recognition.

Teaching Experience

- **Artificial Intelligence** Prof. Hakimeh Fadaie and Yadollah Yaghoobzadeh — Fall 2021, Spring 2022
- **Design and Analysis of Algorithms** Prof. Javad Dousti and Prof. Hamid Mahini — Spring 2021, Spring 2022
- **Engineering Probability and Statistics** Prof. Behnam Bahrak — Fall 2020
- **Operating systems** Prof. Mehdi Kargahi — Spring 2022, Fall 2022
- **Programming languages and Compilers (Chief TA)** Prof. Fatemeh Ghasemi — Spring 2021, Fall 2021, Spring 2022, Fall 2022
- **Data Structures** Prof. Heshaam Faili — Fall 2020, Fall 2021
- **Discrete mathematics** Prof. Siamak Mohammadi — Spring 2020, Fall 2020
- **Computer Architecture** Prof. Saeed Safari — Fall 2021, Spring 2022

Publications

- A.Abaskohi, Y. Yaghoobzadeh, S. Rothe(2023). LM-CPPF: Paraphrasing-Guided Data Augmentation for Contrastive Prompt-Based Few-Shot Fine-Tuning(ACL 2023)
- A.Salemi, A.Abaskohi(equal contribution), S.Tavakoli, Y.Yaghoobzadeh, A.Shakeri (2023). PEACH: Pre-Training Sequence-to-Sequence Multilingual Models for Translation with Semi-Supervised Pseudo-Parallel Document Generation(LoResMT@EACL 2023)
- A.Abaskohi, N.Sabri, B.Bahrak. Persian Emotion Detection using ParsBERT and Imbalanced Data Handling Approaches(Submitted to ACM Transactions on Asian and Low-Resource Language Information Processing)
- A.Abaskohi, T.Zeraati, A.Rasouli, B.Bahrak (2022). UTNLP at SemEval-2022 Task 6: A Comparative Analysis of Sarcasm Detection Using Generative-based and Mutation-based Data Augmentation(International Workshop on Semantic Evaluation)
- A.Abaskohi, F.Mortazavi, H.Moradi (2022). Automatic Speech recognition for Speech Assessment of Persian Preschool Children (Submitted to IEEE/ACM Transactions on Audio, Speech, and Language Processing)

Volunteering

- Artificial Intelligence and Machine Learning Mentor in Summer School 2021 — Summer 2021
- Freelance Content Creation on Medium — Jul 2021-Present
- Working on The Backend API System of Cognitive Tests For Preschool Children Project — Nov 2020-Feb 2021

Extra Courses

- **Natural language processing specialization**
- **Reinforcement Learning Specialization**
- **Machine learning**
- **Advanced Computer Vision with TensorFlow**
- **Intro to Computer Vision and Image Processing**
- **Generative Adversarial Networks**
- **Deep learning specialization**
- **Mathematics for Machine Learning**
- **Cloud Computing Concepts, Part 1**
- **Big data specialization**

Honors and Awards

Supporter Foundation of University of Tehran Scholarship — Sept 2020

Awarded to top 500 students among 35000 students in the university.

Best Undergraduate Project Award — Sept 2022

My project on "Pre-Training Sequence-to-Sequence Multilingual Models for Translation with Semi-Supervised Pseudo-Parallel Document Generation" has been awarded as the best undergraduate project in the 16th Project Day competition of ECE school of the University of Tehran.

Admission to University of Tehran — Sept 2018

Ranked 46th (regional rank), and 1125th (national rank) among 144,437 participants in the Iranian Nationwide University Entrance Exam for Mathematics and Physics discipline.

Iran's Elites Foundation Membership and Financial Support — Nov 2021

After the Automatic Speech Recognition for preschool children project, I have been awarded a 1-year scholarship from this foundation.

Skills

Programming Languages: Python, C\C++, SQL, C#, Java, GNU Octave, MATLAB, Verilog HDL

Technologies: Git, VS Code, Jupyter, Google Colab, Visual Studio, LaTeX

Tools: Tensorflow, Matplotlib, Pytorch, Jupyter, Modelsim, Quartus, CUDA, Django, Flask, GraphQL, Apache Airflow

Operating Systems: Linux (Ubuntu, Parrot, Kali), Windows, Cloudera

---

Notable Projects

Python Code Generator
Python, Huggingface Transformers

Based on the dataset I obtained, I constructed a GPT2 transformer model for this project. I downloaded several python repositories from github and used them to train my model.

Advanced Programming Course Projects
C++

These projects are: famous GUI game: Super Mario, A simple movie network like Netflix, and a simple interpreter.

Artificial Intelligence Projects
Python, Sklearn, Numpy, Pandas, Matplotlib

These project contains: An multi layer neural network implementation from scratch on fashion MNIST dataset, search algorithms, A* algorithm visualization, Naive Bayes sentiment analysis and house price prediction which I participated in the related Kaggle's competition as well.

Sophia Compiler
Java, ANTLR

In this project, I have implemented a complete compiler for a object-oriented language called Sophia. This compiler, checks the code first and then generates the code using Jasmin.

Iranian Music Genre Detection
Python, Tensorflow, Sklearn, Librosa, Flask

In this project, I created used different classifying and clustering models like KNN, SVM, MLP, and KMeans to detect genre of an Iranian music.

Stop Sign Classifier
Python, PyTorch, Pillow

This project was my final project for Intro to Computer Vision and Image Processing course which I use CNN to classify stop sign images.

Premier League Infringing IPs Detection
Python, Apache Airflow, Tensorflow

In this project I created a apache airflow pipeline with different machine learning models to detect the infringing IPs for the premier league matches.

Cafe Bazar Gender Detection
Python, Sklearn

Based on the applications installed, I utilized simple machine learning models to predict the gender of Cafe Bazar (Iranian Android app store) users in this project.