ALIREZA SALEMI

alireza.salemi@ut.ac.ir | Alireza Salemi | alirezasalemi7 | +98 939 774 5850

Personal Data

General Information:

First Name: **Alireza**
Last Name: **Salemi**
Date of Birth: **8 Jan. 1999**
Place of Birth: **Bushehr, Iran**

Research Interests:

- **NLP**
- **Machine Learning**
- **Neural Networks**
- **Computer Vision**
- **Data Science**
- **Mathematics**

Education

University of Tehran [website] Tehran, Iran
B.Sc. student in Computer Engineering *Sep. 2017 – present*

- Ranked 1^{st} among 103 Computer Engineering students
- GPA: 19.67/20
- Relevant Course Works:
  - Artificial Intelligence (20/20)
  - Analysis of Algorithms (20/20)
  - Data Structures and Algorithms (20/20)
  - Probability and Statistics (20/20)
  - Advanced Programming (20/20)
  - Engineering Mathematics (20/20)

Imam Khomeini High School Bushehr, Iran
Diploma in Mathematics and Physics Discipline *Sep. 2013 – Jun. 2016*

- Ranked 1^{st} among 63 Mathematics and Physics Discipline students
- GPA: 19.83/20

Publications

[1] *ARMAN: Pre-training with Semantically Selecting and Reordering of Sentences for Persian Abstractive Summarization*
Alireza Salemi, Emad Kebriaei, Ghazal Neisi Minaei and Azadeh Shakery
To appear in proceedings of **EMNLP-2021**

[2] *UTNLP at SemEval-2021 Task 5: A Comparative Analysis of Toxic Span Detection using Attention-based, Named Entity Recognition, and Ensemble Models [paper]*
Alireza Salemi, Nazanin Sabri, Emad Kebriaei, Behnam Bahrak and Azadeh Shakery
Proceedings of **SemEval-2021**

Research Experiences

Pre-training Language Models with Focus on Summarization UT, Iran
Under Supervision of Dr. Azadeh Shakery *Jan. 2021 – Present*

- This research aimed to develop new Transformer-based language models that perform specifically well in summarization. We suggested three novel pre-training objectives and a new abstractive summarization dataset for the Persian language. Furthermore, we tested our models in few-shot and zero-shot situations too. Our models get SOTA results in all available Persian summarization datasets and many NLU tasks.

Toxic Span Detection [repository] [paper] UT, Iran

Under Supervision of Dr. Behnam Bahrak and Dr. Azadeh Shakery *Aug. 2020 – Feb. 2021*

- This research aimed to develop new machine learning models to annotate toxic words of a tweet. We used statistic-based and keyword-based methods as traditional methods of detecting toxicity and compared them with new neural techniques like attention-based and NER-based models.

Model Learning in Software Product Lines [repository][website] TeIAS, Iran

Under Supervision of Prof. Mohammad Mousavi and Dr. Hossien Hojjat *Jul. 2020 – Dec. 2020*

- Ensuring software correctness is an essential discipline of software engineering. Many quality assurance techniques require a model describing the system's behavior. In this research, we survey various methods of extracting behavioral models from software systems, focusing on software product lines.

Decentralized Enforcement in Message-Based Systems [repository] UT, Iran

Under Supervision of Dr. Fatemeh Ghassemi *Jun. 2020 – Oct. 2020*

- In message-based systems, particular ordering of some messages may violate the desired properties such as confidentiality. To make such systems safe, we propose a confidentiality-based runtime enforcement decentralized algorithm that, given an automata-based specification of unwanted message sequences, prevents specific unwanted sequences messages from being sent.

Notable Course Projects

AirHockey, an online multiplayer game for android [repository] Spring 2021

- Course: Cyber Physical Systems
- Tools: *Java, Android SDK, Python*
- Description: *AirHockey is an online multiplayer distributed android game written using java and android SDK. This application uses Bluetooth to connect devices and simulates a real air hockey game.*

LOGHMEH, an online food delivery website [frontend][backend] Spring 2020

- Course: Internet Engineering
- Tools: *Java, Spring, Maven, Javascript, react, Docker, kubernetes, mysql*
- Description: *LOGHMEH is an online food delivery website written using Java and Spring for backend and javascript and React-Web for frontend. Also, Docker and Kubernetes helped to increase the portability of this application.*

Acton, an actor based compiler [repository] Fall 2019

- Course: Programming Languages and Compilers
- Tools: *Java, Gradle, Antlr, Jasmin*
- Description: *Acton is an actor-based programming language written with Java and produces Java classes using Jasmin that are runnable with JRE. This is a powerful tool to simulate parallel systems.*

FPU, a floating-point processing unit for division and multiplication [repository] Fall 2019

- Course: Computer Aided Design
- Tools: *Verilog, Python, Modelsim, Vivado*
- Description: *FPU is a floating-point processing unit with division and multiplication commands for single and double floating-point numbers. This module that was written with Verilog and synthesized with Vivado could be used as co-processors.*

Awards

The Best Thesis Award in Computer Engineering UT, Iran

B.Sc. thesis was selected as the best thesis in the spring 2021 semester *Spring 2021*

F.O.E (Faculty of Engineering) Award UT, Iran

Ranked 1^{st} among all of 103 Computer Engineering students in 2018 and 2019 *Fall 2018, 2019*

University of Tehran Scholarship UT, Iran

Received scholarship from the UT Sponsors Foundation as an exceptional talent *Fall 2017 - 2020*

University Entrance Examination Iran

Ranked as top students in at national entrance examination to universities in 2017 *Fall 2017*

- Ranked 217^{th} (national) and 59^{th}(regional) among more than 148k candidates

Teaching Assistantship

Artificial Intelligence

- Instructor: Dr. Hakimeh Fadaei
- Semesters: Fall 2020, Spring 2021
- Role: Supervisor, responsible for course projects

Database Design

- Instructor: Dr. Azadeh Shakery
- Semesters: Fall 2020
- Role: TA, responsible for Homework about normal forms in database

Programming Languages and Compiler Design

- Instructor: Dr. Fatemeh Ghassemi
- Semesters: Fall 2020, Spring 2021
- Role: Chief TA, responsible for course projects and Homeworks

Design and Analysis of Algorithms

- Instructor: Dr. Hamid Mahini
- Semesters: Spring 2020
- Role: TA, responsible for Homework about graphs and related algorithms

The Theory of Formal Languages and Automata

- Instructor: Dr. Hossien Hojjat
- Semesters: Spring 2020, Fall 2020, Spring 2021
- Role: TA, responsible for Homework about parsing algorithms and normal forms

Engineering Mathematics

- Instructor: Dr. Mahdi Tale Masouleh
- Semesters: Spring 2020
- Role: TA, responsible for Homework about mapping and its applications in solving problems

Skills & Qualities

Academic Skills

Skills that are related to my education and work

- Programming Languages and frameworks: Python, Java, C/C++, R, Dart, Javascript, Verilog HDL, Tensorflow, Keras, Pytorch, Flutter, Numpy, Pandas, Seaborn, Scikit-Learn, Matplotlib, React-web, Express, Spring
- Tools: NLTK, Spacy, Gensim, Transformers, Modelsim, Quartus, Vivado, Multisim, Proteus, MySQL, PostgreSQL, Neo4j, Redis, Elastic Search, Git, Android Studio, LaTeX, Antlr4, Docker, Kubernetes, Maven, Gradle

Personal Qualities

Qualities that are related to my personal abilities

- Organized
- Team Player
- Punctual
- Fast Learner
- Diligent
- Problem Solver
- Creative
- Ethical
- Flexible
- Reliable

Languages

Persian: Native

English: Fluent (I will take the TOEFL test on October 16)

Arabic: Familiar

References

Available upon request.