Mohsen Fayyaz

RESEARCH ASSISTANT · UNIVERSITY OF TEHRAN

Department of Electrical and Computer Engineering, Faculty of Engineering, University of Tehran, Iran

✉ mohsen.fayyaz77@gmail.com | 🏠 mohsenfayyaz.github.io | ⊙ mohsenfayyaz | 🔗 mohsenfayyaz | 🎓 Mohsen Fayyaz

Education

School of Electrical and Computer Engineering, University of Tehran   *Tehran, Iran*

MASTER'S IN ARTIFICIAL INTELLIGENCE, COMPUTER ENGINEERING   *Sep. 2021 - Present*

- TOTAL GPA: **20/20 (4/4)**
- Relevant Course Works:
    - Natural Language Processing (20/20)   - Machine Learning (20/20)
    - Neural Networks (20/20)   - Statistical Inference (20/20)

School of Electrical and Computer Engineering, University of Tehran   *Tehran, Iran*

B.SC. IN COMPUTER ENGINEERING   *Sep. 2017 - Sep. 2021*

- TOTAL GPA: **19.08/20**
- Ranked 3rd among 103 Computer Engineering students
- Relevant Course Works:
    - Artificial Intelligence (20/20), Advanced Programming (20/20), Design and Analysis of Algorithm (20/20), Parallel Programming (19.2/20), Programming Languages and Compilers (20/20), Operating Systems (19/20)

National Organization for Development of Exceptional Talents   *Tehran, Iran*

DIPLOMA IN MATHEMATICS AND PHYSICS DISCIPLINE   *Sep. 2014 - Jun. 2017*

- GPA: 19.68/20

Research Interests

- Natural Language Processing   - Deep Learning

Publications

BERT on a Data Diet: Finding Important Examples by Gradient-Based Pruning   *ENLSP @ NeurIPS 2022*

- **Mohsen Fayyaz***, E. Aghazadeh*, A. Modarressi*, MT Pilehvar, Y. Yaghoobzadeh and S. Ebrahimi
- (* equal contribution)

GlobEnc: Quantifying Global Token Attribution by Incorporating the Whole Encoder Layer in Transformers [paper] [code] [video]   *NAACL 2022* *(Oral Presentation)*

- Ali Modarressi*, **Mohsen Fayyaz***, Yadollah Yaghoobzadeh and Mohammad Taher Pilehvar
- (* equal contribution)

Metaphors in Pre-Trained Language Models: Probing and Generalization Across Datasets and Languages [paper] [code] [video]   *ACL 2022*

- Ehsan Aghazadeh*, **Mohsen Fayyaz*** and Yadollah Yaghoobzadeh
- (* equal contribution)

Not All Models Localize Linguistic Knowledge in the Same Place: A Layer-wise Probing on BERToids' Representations [paper] [code] [blog]   *BlackboxNLP @ EMNLP 2021*

- **Mohsen Fayyaz***, Ehsan Aghazadeh*, Ali Modarressi, Hosein Mohebbi and Mohammad Taher Pilehvar
- (* equal contribution)

Research Experience

Remote Research Assistant at École de technologie supérieure   *Montreal, Canada*

UNDER SUPERVISION OF PROF. SAMIRA EBRAHIMI AND PROF. YADOLLAH YAGHOOBZADEH   *Spring 2022 - Present*

- BERT on a data diet: Finding important examples by gradient-based pruning (ENLSP @ NeurIPS 2022)

Graduate Research Assistant at University of Tehran   *Tehran, Iran*

UNDER SUPERVISION OF PROF. YADOLLAH YAGHOOBZADEH AND PROF. MOHAMMAD TAHER PILEHVAR   *Fall 2021 - Present*

- Quantifying global token attribution by incorporating the whole encoder layer in Transformers. (NAACL 2022)
- Metaphors in pre-trained language models: Probing and generalization across datasets and languages. (ACL 2022)

NLP Research Intern at Tehran Institute for Advanced Studies (TeIAS) *Tehran, Iran*
UNDER SUPERVISION OF PROF. MOHAMMAD TAHER PILEHVAR *Summer 2021*
- Layer-wise probing on BERT, XLNet, and ELECTRA representations for linguistic knowledge. (BlackboxNLP @ EMNLP2021)

Undergraduate Research Assistant at University of Tehran *Tehran, Iran*
UNDER SUPERVISION OF PROF. AZADEH SHAKERY *Spring 2021*
- Investigating toxicity detection knowledge in contextual language models and possible biases.

Member of NLP Study Group *Tehran, Iran*
UNDER SUPERVISION OF PROF. MOHAMMAD TAHER PILEHVAR *Summer 2020*
- Laid the foundations for carrying out research. Initiated nlpdataset.ir for tracking NLP datasets in Farsi.

Work Experience

Data Scientist *Tehran, Iran*
SOTOON *Jul. 2022 - Present*
- Developed NLP/AutoML services including various ML tasks and models at large scale.

Machine Learning Engineer *Tehran, Iran*
TAPSELL *Sep. 2021 - Jul. 2022*
- Developed NLP engines and ML models for enhancing targeted advertising.

Honors & Awards

2021	**Enrolled in the graduate school as a top student**, Without passing the entrance examinations	*University of Tehran*
2020	**F.O.E (Faculty of Engineering) Award**, Ranked 1st among all of 103 Computer Engineering students	*University of Tehran*
2019	**F.O.E (Faculty of Engineering) Award**, Ranked 3rd among all of 103 Computer Engineering students	*University of Tehran*
2017	**University of Tehran Scholarship**, Received scholarship from the University of Tehran Sponsors Foundation as an exceptional talent	*University of Tehran*
2017	**Ranked in top 0.4% in Konkur Exam**, Of all 148,429 participants in Iran	*Iran*
2015	**Second place in Khwarizmi Youth Award**, An important national research award given annually by the President of Iran.	*Iran*

Teaching

TEACHING ASSISTANTSHIP

2020-2022	**Artificial Intelligence**, Prof. Yaghoobzadeh, Dr. Fadaei	*University of Tehran*
2019-2020	**Formal Languages and Automata Theory**, Prof. Hojjat	*University of Tehran*
2019-2020	**Advanced Programming**, Prof. Khosravi	*University of Tehran*
2019	**Digital Logic Design**, Prof. Navabi	*University of Tehran*

Professional Development

The following online courses were taken to acquire skills relevant to my research:

- Build Basic Generative Adversarial Networks (GANs), Coursera.org - Feb 2021 - Credential
- Deep Learning Specialization, Coursera.org - Sep 2020 - Credential
- Sequence Models, Coursera.org - Sep 2020 - Credential
- Convolutional Neural Networks, Coursera.org - Aug 2020 - Credential
- Structuring Machine Learning Projects, Coursera.org - Aug 2020 - Credential
- Improving Deep Neural Networks: Coursera.org - Aug 2020 - Credential
- Neural Networks and Deep Learning, Coursera.org - Aug 2020 - Credential

Skills

Programming	Python (PyTorch, sklearn, NumPy, Pandas, Tensorflow, PySpark), C++, Java, Java-script, SQL
Tools and Frameworks	PyTorch, Keras, Tensorflow, HuggingFace, Transformers, Git, React-Native, Android Studio
Languages	• Persian: Native • English: Fluent (TOEFL score: **116/120**, Reading: 29, Listening: 30, Speaking: 28, Writing: 29)

Notable Projects

Natural Language Processing Course Assignments @ University of Tehran   *Winter 2022*
- Including Tokenization, Naive Bayes, RNN, LSTM, Transformers, Machine Translation, QA and NLU. (Pytorch)

Neural Networks Course Assignments @ University of Tehran   *Winter 2022*
- Including CNN, RNN, GAN and RL. (Pytorch, Tensorflow)

Semantic Matching of Queries to Ads   *Summer 2021*
- Building proper embeddings for queries and ads to achieve fast prediction of sponsored search advertising using their proximity in the same representation space. [repository]

Colorizing Grayscale Images Using Inception Network   *Summer 2020*
- Modified an inception model and trained it on 10K images to colorize grayscale images. [repository]

Online Real-time Propeller LED Display   *Spring 2020*
- Embedded Systems Final Course Project. https://www.youtube.com/watch?v=wf1WOrsO_tA

References

Mohammad Taher Pilehvar, mp792@cam.ac.uk	*TelAS*
Yadollah Yaghoobzadeh, y.yaghoobzadeh@ut.ac.ir	*University of Tehran*
Hossein Hojjat, hojjat@ut.ac.ir	*University of Tehran*