Mohsen Fayyaz

CS Ph.D. Student · University of California, Los Angeles

UCLA Computer Science Department, 404 Westwood Plaza, Engineering VI, Los Angeles, CA 90095-1596

✉ mohsenfayyaz@cs.ucla.edu | 🏠 mohsenfayyaz.github.io | ⬛ mohsenfayyaz | 🔗 mohsenfayyaz | 🎓 Mohsen Fayyaz

Education

University of California, Los Angeles *Los Angeles, CA*
Ph.D. in Computer Science *Sep. 2023 - Present*
- advisor: Nanyun (Violet) Peng

University of Tehran *Tehran, Iran*
Master's in Artificial Intelligence, Computer Engineering *Sep. 2021 - Aug. 2023*
- GPA: 20/20 (4/4)
- advisor: Yadollah Yaghoobzadeh

University of Tehran *Tehran, Iran*
B.Sc. in Computer Engineering *Sep. 2017 - Sep. 2021*
- GPA: 19.08/20
- Ranked 3rd among 103 Computer Engineering students

Research Interests

- Natural Language Processing
- Deep Learning

Publications

7. **Evaluating Human Alignment and Model Faithfulness of LLM Rationale** *arXiv preprint 2024*
- **M Fayyaz**, F Yin, J Sun, N Peng

6. **SkipPLUS: Skip the First Few Layers to Better Explain Vision Transformers** *CVPR Workshops 2024*
- F Mehri, **M Fayyaz**, MS Baghshah, MT Pilehvar

5. **DecompX: Explaining Transformers Decisions by Propagating Token Decomposition** *ACL 2023*
- A Modarressi*, **M Fayyaz***, E Aghazadeh, Y Yaghoobzadeh, MT Pilehvar
- (* equal contribution)

4. **BERT on a Data Diet: Finding Important Examples by Gradient-Based Pruning** *ENLSP @ NeurIPS 2022*
- **M Fayyaz***, E Aghazadeh*, A Modarressi*, MT Pilehvar, Y Yaghoobzadeh, SE Kahou
- (* equal contribution)

3. **GlobEnc: Quantifying Global Token Attribution by Incorporating the Whole Encoder Layer in Transformers** *NAACL 2022*
- A Modarressi*, **M Fayyaz***, Y Yaghoobzadeh, MT Pilehvar
- (* equal contribution)

2. **Metaphors in Pre-Trained Language Models: Probing and Generalization Across Datasets and Languages** *ACL 2022*
- E Aghazadeh*, **M Fayyaz***, Y Yaghoobzadeh
- (* equal contribution)

1. **Not All Models Localize Linguistic Knowledge in the Same Place: A Layer-wise Probing on BERToids' Representations** *BlackboxNLP @ EMNLP 2021*
- **M Fayyaz***, E Aghazadeh*, A Modarressi, H Mohebbi, MT Pilehvar
- (* equal contribution)

Research Experience

Graduate Student Researcher at University of California, Los Angeles *Los Angeles, CA*
Under supervision of Prof. Nanyun (Violet) Peng *Fall 2023 - Present*
- Ph.D. Researcher at PLUS Lab

Remote Collaborator at École de technologie supérieure

Montreal, Canada

UNDER SUPERVISION OF PROF. SAMIRA EBRAHIMI AND PROF. YADOLLAH YAGHOOBZADEH

Spring 2022

- 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

2021 - 2023

- 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.

Work Experience

Data Scientist

Tehran, Iran

DIVAR

Jun. 2023 - Aug. 2023

- Developed NLP services including various ML tasks and models at large scale.

Data Scientist

Tehran, Iran

SOTOON

Jul. 2022 - Jun. 2023

- 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

2024	**CS263 - Natural Language Processing**, Prof. Nanyun (Violet) Peng	*UCLA*
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

Nanyun (Violet) Peng, violetpeng@cs.ucla.edu		*UCLA*
Mohammad Taher Pilehvar, mp792@cam.ac.uk		*Cardiff University / TeIAS*
Yadollah Yaghoobzadeh, y.yaghoobzadeh@ut.ac.ir		*University of Tehran*
Hossein Hojjat, hojjat@ut.ac.ir		*University of Tehran*