Ehsan Aghazadeh

Education

- **University of Massachusetts, Amherst** Feb. 2023 – Feb. 2028 (Expected)
  Masters/Ph.D. student in Computer Science Massachusetts, Amherst
  Advisors: Hossein Pishro-Nik, Hedyeh Beyhaghi

- **University of Tehran** Sep. 2017 – Aug. 2022
  B.Sc. in Computer Engineering Tehran, Iran
  GPA: 17.04/20

Interests

Natural Language Processing ML Theory Deep Learning

Publications

8. CGES: Confidence-Guided Early Stopping for Efficient and Accurate Test-Time Scaling
E. Aghazadeh, A. Ghasemi, H. Beyhaghi, H. Pishro-Nik
Under submission

7. A Comprehensive Analysis for Visual Object Hallucination in Large Vision-Language Models
L. Jing, H. Chen, **E. Aghazadeh**, X. E. Wang, X. Du
KnowledgeFM @ ACL 2025

6. From RAGs to rich parameters: Probing how language models utilize external knowledge over parametric information for factual queries
H. Wadhwa, R. Seetharaman, S. Aggarwal, R. Ghosh, S. Basu, S. Srinivasan, W. Zhao, S. Chaudhari, **E. Aghazadeh**
BlackboxNLP @ EMNLP 2024

4. DecompX: Explaining Transformers Decisions by Propagating Token Decomposition
A. Modaressi, M. Fayyaz, **E. Aghazadeh**, Y. Yaghoobzadeh, M. T. Pilehvar
ACL 2023

3. BERT on a Data Diet: Finding Important Examples by Gradient-Based Pruning
M. Fayyaz*, **E. Aghazadeh***, A. Modaressi*, Y. Yaghoobzadeh, M. T. Pilehvar, S. E. Kahou
*Equal contribution *ENLSP @ NeurIPS 2022*

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

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

Experience

- **Graduate Research Assistant at University of Massachusetts Amherst** Feb. 2025 – Now
  University of Massachusetts Amherst, Supervisors: Dr. H. Pishro-Nik, Dr. H. Beyhaghi Amherst, MA
  CGES: Confidence-Guided Early Stopping for Efficient and Accurate Test-Time Scaling in Large Language Models

  - Proposed a Bayesian confidence-based framework (CGES) for test-time scaling that leverages per-response confidence scores to guide early stopping.

- Theoretically proved CGES converges to the correct answer and can match or exceed majority voting performance while using fewer LLM calls.
- Demonstrated substantial improvements in efficiency across multiple reasoning benchmarks without sacrificing accuracy, using both reward-model and probability-based confidence estimates.

- **Graduate Research Assistant at University of Massachusetts Amherst** March 2024 – Sep. 2024
  University of Massachusetts Amherst Amherst, MA
  What Causes Visual Hallucination? A Comprehensive Analysis for Hallucination in Large Vision-Language Models

  - Conducted coarse-grained model analysis by probing the language decoder, vision encoder, and projector module using various probing methods.
  - Demonstrated that hallucinations are more likely to originate from the vision encoder and projector module.
  - Performed fine-grained analysis of the vision encoder, identifying specific parts of the encoder block responsible for hallucinations.

- **Undergraduate Research Assistant at University of Tehran** Sep. 2022 – Jan. 2023
  University of Tehran, Supervisors: Dr. Mohammad Taher Pilehvar Tehran, Iran
  Introducing a new approach for measuring attribution scores in language models

  - Introducing DecompX, an explanation method based on propagating decomposed token vectors up to the classification head
  - Incorporating all the encoder layer components including non-linear functions, propagated the decomposed vectors throughout the whole model

- **Undergraduate Research Assistant at IPM** Sep. 2022 – Dec. 2022
  IPM, Supervisors: Dr. Mohsen Fayyaz, Dr. Mohammad Sabokrou Tehran, Iran
  Proposing a new Transformer-based Block for Novelty Detection Task and Preventing Catastrophic Forgetting

- **Remote Research Assistant at École de technologie supérieure** Feb. 2022 – Nov. 2022
  École de technologie supérieure, Supervisor: Dr. Samira Ebrahimi Kahou Montreal, Canada
  BERT on a Data Diet: Finding Important Examples by Gradient-Based Pruning

  - Adapting GraNd and EL2N metrics to the language domain to identify important examples in a dataset
  - Showing that in contrary to the results in computer vision, early score computation steps are not sufficient for finding a proper subset of the data in NLP
  - Observing that pruning a small fraction of the examples with the highest EL2N/GraNd scores will result in better performance and in some cases even better than fine-tuning on the whole dataset

- **Undergraduate Research Assistant at University of Tehran** Apr. 2021 – Jan. 2022
  University of Tehran, Supervisor: Dr. Yadollah Yaghoobzadeh Tehran, Iran
  Metaphor Detection Exploration

  - Applied probing methods for metaphorical knowledge in pre-trained models for the first time and showed that PLMs do encode it
  - Pointed out that metaphorical knowledge is encoded in the middle layers than other layers
  - Quantified how metaphorical knowledge is generalizable across four languages and four dataset sources

- **Undergraduate Research Assistant at Tehran Institute for Advanced Studies** Jan. 2021 – Aug. 2021
  Tehran Institute for Advanced Studies, Supervisor: Dr. Mohammad Taher Pilehvar Tehran, Iran
  Investigation of Contextualized Representations

- Pointed out that the weight mixing evaluation strategy in edge probing does not lead to reliable conclusions in layer-wise cross model analysis studies
- Carried out a probing analysis across three commonly used pre-trained models by relying on an information-theoretic probing method
- Extended our probing experiments to fine-tuned representations to examine how linguistic information changes during fine-tuning
- Provided complementary results to validate our findings, we also employed RSA to measure the amount of change in the representations after fine-tuning

Academic Service

Program Committee / Reviewer for 20+ papers in EACL 2024, NAACL 2024, ACL 2024, ACL Rolling Review (since 2023)

Extra Curricular & Leadership

Member of Scientific Student Chapter Jun. 2019 – Aug. 2020
University of Tehran, School of Electrical and computer Engineering *Tehran, Iran*

- We held three high-quality schools for three high tech topics: Data Science, Internet of Things, and Signal Processing.
- We held over 20 other events.
- I was the leader of "Data Science Winter School" executive team.