Yeganeh Kordi

📱 (401)648 1171   •   ✉ yegane.kordi@gmail.com   •   🌐 yeganehkordi.github.io

in yeganehkordi   •   ⌨ yeganehkordi   •   🐦 yeganekordi

EDUCATION

Brown University   2023–Present
PhD Student in Computer Science
Advisor: Stephen Bach

Amirkabir University of Technology   Tehran-Iran
Bachelor of Science (Double degree)
- Electrical Engineering   *2016–2020*
  - Major: Control Systems
  - GPA: 3.84/4
  - Class Rank: 3/40 in Control Group, 14/160 in Electrical Engineering
    Advisors: Mohammad Bagher Menhaj, Mohammad A. Khosravi
- Computer Engineering   *2018–2021*
  - Major: Artificial Intelligence
  - Minor: Computer Networking
  - GPA: 3.78/4
  - Class Rank: Among top 10%
    Advisor: Mohammad Rahmati

RESEARCH INTERESTS

I am broadly interested in AI, Machine Learning, and NLP, particularly, I'm interested in language model alignment, building machines that can follow human instructions, improving the performance of models in zero-shot and few-shot settings, and moving toward generalization and unification.

PUBLICATIONS

- Self-Instruct: Aligning Language Model with Self Generated Instructions
  Yizhong Wang, **Yeganeh Kordi**, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi
  ACL, 2023
- Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ Tasks
  $\diamond$Yizhong Wang, $\diamond$Swaroop Mishra, ♣Pegah Alipoormolabashi, ♣**Yeganeh Kordi**, ..., Chitta Baral, Yejin Choi, Hannaneh Hajishirzi, Noah A. Smith, Daniel Khashabi
  EMNLP, 2022
  $\diamond$ Co-first authors; ♣ Co-second authors, alphabetical order.
- UnifiedQA-v2: Stronger Generalization via Broader Cross-Format Training.
  Daniel Khashabi, **Yeganeh Kordi**, and Hannaneh Hajishirzi
  arXiv preprint arXiv:2202.12359.

RESEARCH EXPERIENCE

- *Allen Institute for AI* and *H2lab at University of Washington*
  Mentors: Daniel Khashabi, Yizhong Wang
  Collaborating on NLP research projects, including   July 2021–July 2023
  - UnifiedQA v2
    - We have created a cross-format QA model using the same process as UnifiedQA, but with more supervision. This results in better transfer between different QA variants and generalization to unseen datasets. In addition, this leads to better in-domain and cross-domain results.
  - Super-NaturalInstructions
    - We introduced the SUPER-NATURALINSTRUCTIONS benchmark, which consists of 1,616 diverse NLP tasks and their expert-written instructions. We also built $\mathrm{T}k$-INSTRUCT, a transformer model trained to follow

various in-context instructions, which outperforms existing instruction-following models on our benchmark.
- Self-Instruct
  · We proposed a self-training approach that uses a Pre-trained LM to generate a large number of instructions and their desired outputs and then fine-tunes the model with the generated data. (in-submission paper)
- *JHU Center for Language and Speech Processing*
  Mentor: Daniel Khashabi
  Collaborating on NLP projects: August 2022–Present
  - Web Instructions
    · We created a benchmark and model for instructions in HTML layout. This can be used as a challenge for language models and optimizing the crowdsourcing process. (paper under preparation)

TEACHING EXPERIENCES

- Teaching Assistant
  - Machine Learning   Fall 2021 - Spring 2023
  - Data Mining   Fall 2021
  - Algorithm Design   Fall 2021 - Winter 2022
  - Cloud Computing   Spring 2021
  - Digital Logic Circuits   Fall 2020-Fall 2021

HONORS

- Ranked 148th in university entrance exam, among more than 70,000 participants.
- Ranked 19th in university graduate entrance exam, among more than 5,000 participants.
- Granted admission from the Talented Student Office of the Amirkabir University of Technology for double degree program.
- Member of Iran's National Elites Foundation.

PROJECTS

- Implementing a Persian search engine from scratch
  - In this project, I have used a dataset of 7000 news articles and scored the documents using TF-IDF. I have also built a k-NN Similarity Search Engine based on this data.
- Implementing of 3D eye tracking method for use in medical experiments.
  - In this project, we set up a proper 3D gaze-tracking project. Then, we designed a headset with eye cameras that are equipped with infrared illuminations. Also, we designed software to collect the hemispherical strabismus scanning and turned it into a real patient test in eye clinics.
- Research paper classification system
- Persian Twitter sentiment analysis

COMPUTER SKILLS

- **Programming Languages:** Python, Java, C/C++
- **ML and NLP:** SpaCy, NLTK, Hugging Face Tokenizers and Transformers, Fairseq, pandas, Sklearn, TensorFlow, Keras, OpenCV, CUDA, Matplotlib, NumPy
- **Web-development:** HTML, CSS, JavaScript, Flask, Django
- **Databases:** MySQL, PostgreSQL
- **Hardware:** ARM, FPGA, VHDL, Verilog, Altium Designer, Proteus Design Suite
- **Cloud Computing:** Hadoop, HAProxy, Kubernetes, Docker
- **Developer Environments:** Jupyter, Colab, IntelliJ, Eclipse, PyCharm, VSCode
- **Other:** Matlab, Git, ROS, Gazebo, Wireshark, OMNeT++, VM VirtualBox

POSITION OF RESPONSIBILITY

- Member of Technical Committee in IranOpen RoboCup Competition 2018
- Member of Student Scientific Association of electrical engineering at Amirkabir University of Technology

❶ References, Further information, and Proofs are available upon Request