

Amirhossein Kazemnejad

6650 Rue Saint-Urbain, Montreal, QC H2S 3G9

✉ amirhossein.kazemnejad@mila.quebec • 🌐 Website • 📄 kazemnejad • 🎓 Google Scholar

Research Interests

Compositional Generalization, Length Extrapolation, Computational Power of Transformers, Positional Encoding, Representation Learning for Sequential Data

Education

McGill University / Mila MSc (Thesis-based) in Computer Science Supervisor: Prof. Siva Reddy	2021 - 2023 (Expected) GPA: 4/4
Iran University of Science and Technology (IUST) BSc in Computer Engineering Supervisor: Prof. Mohammad Taher Pilehvar Thesis: <i>Analyzing The Effects of Semantic Violations on Computational Language Models</i>	2015 - 2020 GPA: 17.90/20

Publications

The Curious Case of Absolute Position Embeddings Findings of Empirical Methods in Natural Language Processing (EMNLP), Koustuv Sinha*, Amirhossein Kazemnejad*, Siva Reddy, Joelle Pineau, Dieuwke Hupkes, Adina Williams	2022
Paraphrase Generation by Learning How to Edit from Samples Association for Computational Linguistics (ACL), Amirhossein Kazemnejad, Mohammadreza Salehi, Mahdieh Soleymani	2020

*equal contribution

Research Experience

Present Jan 2021	Mila - Québec AI Institute <i>Graduate Research Assistant Supervisor: Prof. Siva Reddy</i> <ul style="list-style-type: none">Working on improving human-like generalization and adaptive computation in Transformers. Collaborated with teams from AI2, Meta AI, and IBM Research.Focus: Out-of-Distribution Generalization, Transformer Architecture, Positional Encoding
Present Jul 2022	Noah's Ark Lab - Huawei Canada <i>Part-time Research Intern @ NLP Team Supervisor: Dr. Prasanna Parthasarathi</i> <ul style="list-style-type: none">Investigating the gap between world knowledge acquisition and utilization in pre-trained language models.Focus: Knowledge Extraction, Pre-Trained Language Models, Downstream Knowledge Transfer
2020 2018	Machine Learning Lab - Sharif University <i>Undergraduate Research Assistant Supervisor: Prof. Mahdieh Soleymani</i> <ul style="list-style-type: none">Worked on equipping seq2seq Transformers with a retrieval-based editor module to improve conditional text-generation.Focus: Paraphrase Generation, Conditional Text Generation, Retrieval-based Editor Models

Selected Research Projects

Length Generalization in Transformers <i>Advisor: Prof. Siva Reddy</i> <ul style="list-style-type: none">Evaluating the impact of positional encoding on length generalization and its interaction with scratchpad (chain-of-thought).Analyzing the surprising effectiveness of Transformers without positional encoding and how they encode positions. Work is being done in collaboration with IBM Research.	Sep 2022 - Present
Knowledge Triggers in Pre-Trained Language Models <i>Advisor: Dr. Prasanna Parthasarathi</i> <ul style="list-style-type: none">Identifying the key determinants and inductive biases that trigger the effective application of knowledge within PLMs for reasoning in downstream tasks through designing a behavioral probing dataset.Investigating the claim that knowledge probing serves as a lower bound for measuring the world knowledge in PLMs.	Jul 2022 - Present
The Shift Invariance Property in Absolute Position Embeddings <i>Advisor: Prof. Siva Reddy</i>	Mar 2022 - Jul 2022

- Studied absolute position embeddings (APEs) in Transformers and their ability to capture relative position information.
- Demonstrated over-reliance on positional shortcuts and poor performance on non-zero positions of models trained with APEs.
- Provided extensive experimental results over 8 finetuning and prompting tasks. Work done in collaboration with Meta AI and published at **EMNLP'22** Findings. Also, presented at **BlackboxNLP'22**

Contrastive Learning For Structured Prediction

Jan 2021 - Dec 2021

Advisor: *Prof. Siva Reddy*

- Proposed a contrastive learning framework to enforce consistency of local structures in the output space.
- Applied the framework to the representations of the Transformer's decoder to improve compositional generalization.
- Improved out-of-domain performance in SCAN and a few semantic parsing tasks with no modifications to the model. Work done in collaboration with AI2.

Improving Editor-based Text Generation Architecture

2018 - 2020

Advisor: *Prof. Mahdieh Soleymani*

- Incorporated the retrieve-and-edit framework into Transformers with a novel module for a fine-grained editing process.
- Utilized the framework in paraphrase generation for improved quality and data augmentation for improved diversity.
- Demonstrated the generation improvements in terms of automatic metrics and human evaluation. Work published at **ACL'20**.

Teaching Experience

Teaching Assistant , McGill, COMP 330 Theory of Computation (Prof. Prakash Panangaden)	Fall 2021
Teaching Assistant , McGill, COMP 204 Computer Programming (Prof. Yue Li)	Winter 2021
Teaching Assistant , Sharif University, CE719 Deep Learning (Prof. Mahdieh Soleymani)	Winter 2020
Teaching Assistant , IUST, Deep Learning (Prof. Mohammad Taher Pilehvar)	Winter 2019
Teaching Assistant , IUST, Intro. to NLP (Prof. Sauleh Eetemadi)	Winter 2019
Teaching Assistant , IUST, Intro. to AI and Expert Systems (Prof. Mohammad Taher Pilehvar)	Fall 2018

Professional Activities

Organiser	GenBench Workshop (website), EACL 2021 (website), Virtual Conference of EMNLP 2020 (website)
Program Committee	ACL 2023, EMNLP 2022, ACL Rolling Review Dec. 2021

Services and Volunteer work

Open-Source Contribution

- **TensorFlow 2.0** (GitHub links: [#375](#), [#503](#), [#511](#), [#534](#), [#546](#), [#535](#), [#673](#), [#603](#), [#335](#))
Contributed to Seq2Seq module features, bug fixes, and documentation.
- **Jekyll Academic Template** ([GitHub Repo](#))
Feature-rich Jekyll template for academic courses with 44 stars and 82 forks on GitHub.
- **PLM Research Codebase** ([GitHub Repo](#))
PyTorch/Huggingface codebase, Seq2Seq tasks, decoder-only architectures, support for dependency injection, and SLURM.

Blog Posts

- **Transformer Architecture: The Positional Encoding** ([link](#))
Top Google result for *Positional Encoding* with nearly 5K monthly views. Referenced by [Stanford](#), [CMU](#), and [MIT](#) courses.
- **How to do Deep Learning research with absolutely no GPUs** ([link to parts #1 and #2](#))
- **TensorFlow 2.0 Tutorial** ([link](#))

Honors and Awards

- (2021) Awarded **Graduate Excellence Fellowship** by McGill University.
- (2020) Ranked **2nd** among Computer Engineering students of the 2015 batch.
- (2017-2019) Recipient of **Outstanding Student** Award by Iran University of Science and Technology.
- (2015) Top **99.2nd** percentile in national university entrance exam among nearly 200,000 participants.

Skills

Programming Languages	Python, Java, C++, Bash
Frameworks & Libraries	TensorFlow, Keras, PyTorch, PyTorch Lightning, Numpy, HuggingFace, WandB, Matplotlib, Plot9
Other Tools	Git, Docker, Singularity, SLURM, GitHub Actions
Languages	Persian (Native), English (Fluent), Arabic (Limited)