

# SINA BAGHERI NEZHAD

sina.bagherinezhad@pdx.edu • sinaai.github.io

## EDUCATION

---

<b>PhD in Computer Science</b> <i>Portland State University</i>	Sep 2022 – Dec 2025 (anticipated)
<b>Masters in Computer Engineering - Artificial Intelligence</b> <i>Iran University of Science and Technology</i>	Sep 2019 – Dec 2021
<b>Masters in Technology Management</b> <i>University of Tehran</i>	Feb 2014 – Feb 2017
<b>Bachelor in Industrial Engineering</b> <i>Karaj Azad University</i>	Sep 2007 – Jul 2012

## RESEARCH EXPERIENCE

---

<b>Graduate Research Assistant</b> <i>PortNLP Lab &amp; Department of Computer Science @ Portland State University</i>	Aug 2023 - Now
<ul style="list-style-type: none"><li>• Conduct research on multilingual language models, focusing on the correlation between language characteristics and the performance of multilingual models. Investigate the impact of model scaling on seen and unseen language performance.</li><li>• Research fairness in cross-lingual and multilingual summarization and question-answering (QA) systems, focusing on ensuring equitable model performance across languages and fair representation of diverse groups in summaries.</li><li>• Develop and analyze fair clustering algorithms, contributing to the development of a novel constant-approximation algorithm for fair sum-of-radii clustering.</li><li>• Supervised by Prof. Ameeta Agrawal and Prof. Sayan Bandyopadhyay.</li></ul>	
<b>Researcher</b> <i>Data mining lab @ Iran University of Science and Technology</i>	Sep 2020 - Apr 2021
<ul style="list-style-type: none"><li>• Research and develop a method for searching and retrieving information from a vast collection of multilingual documents, particularly those containing both Arabic and Persian content, using Elasticsearch.</li><li>• Research and develop a method for providing search suggestions to users based on the search history of other users in a search engine.</li><li>• Supervised by Prof. Behrouz Minaei-Bidgoli</li></ul>	
<b>Master's Thesis</b> <i>Iran University of Science and Technology</i>	Apr 2021 - Dec 2021
<ul style="list-style-type: none"><li>• Using machine learning methods to determine factors affecting mortality rate and hospitalization length of COVID-19 patients.</li><li>• Supervised by Prof. Nasser Mozayani</li></ul>	
<b>Master's Thesis</b> <i>University of Tehran</i>	Jan 2016 - Feb 2017
<ul style="list-style-type: none"><li>• Determining innovation risks in research and development phases: an application of qualitative research in the pharmaceutical industry</li><li>• Supervised by Prof. Nima Garoosi Mokhtarzadeh</li></ul>	

## TEACHING EXPERIENCE

---

<b>Graduate Teaching Assistant</b> <i>Portland State University</i>	Sep 2022 - Now
<ul style="list-style-type: none"><li>• TA in Ethics in AI, Instructed by Dr. Ameeta Agrawal</li><li>• TA in Discrete Structures I and II, Instructed by Dr. Sarah Mocas</li></ul>	
<b>Graduate Teaching Assistant</b> <i>Iran University of Science and Technology</i>	Sep 2020 - Feb 2021
<ul style="list-style-type: none"><li>• TA in Neural Networks, Instructed by Dr. Nasser Mozayani</li></ul>	
<b>Graduate Teaching Assistant</b> <i>University of Tehran</i>	Sep 2014 - Feb 2015
<ul style="list-style-type: none"><li>• TA in Non-parametric statistic, Instructed by Dr. Ezzatollah Asgharizadeh</li></ul>	
<b>Teaching Assistant</b> <i>Karaj Azad University</i>	Sep 2010 - Jun 2012
<ul style="list-style-type: none"><li>• TA in Operation Research, Instructed by Dr. Sadollah Ebrahimnejad</li><li>• TA in Engineering statistics, Instructed by Dr. Sadollah Ebrahimnejad</li></ul>	

## OTHER RELEVANT EXPERIENCE

---

### Head of Software Development

Sep 2021 - Aug 2022

*Basir Andishan Bina Tadbir*

- Implementing agile software development methods within the team's culture
- Introducing and integrating a code review process into the software production workflow
- Designing and configuring Jira, Bitbucket, and Confluence to optimize team collaboration and project management

## PUBLICATIONS

---

<b>Beyond Data Quantity: Key Factors Driving Performance in Multilingual Language Models</b>	2024
Sina Bagheri Nezhad, Ameeta Agrawal and Rhitabrat Pokharel	
Under-review	
<b>Constant-Approximation Algorithms for Fair Sum-of-Radii Clustering</b>	2024
Sina Bagheri Nezhad, Sayan Bandyapadhyay and Tianzhi Chen	
Under-review	
<b>Investigating the Impact of Model Scaling on Seen and Unseen Language Performance</b>	2024
Rhitabrat Pokharel, Sina Bagheri Nezhad, Ameeta Agrawal and Suresh Singh	
Under-review	
<b>Fair Summarization: Bridging Quality and Diversity in Extractive Summaries</b>	2024
Sina Bagheri Nezhad, Sayan Bandyapadhyay and Ameeta Agrawal	
Algorithmic Fairness through the Lens of Metrics and Evaluation @ NeurIPS 2024	
<b>Evaluating Multilingual Long-Context Models for Retrieval and Reasoning</b>	2024
Ameeta Agrawal, Andy Dang, Sina Bagheri Nezhad, Rhitabrat Pokharel and Russell Scheinberg	
The 4th Multilingual Representation Learning (MRL) Workshop @ EMNLP 2024	
<b>What Drives Performance in Multilingual Language Models?</b>	2024
Sina Bagheri Nezhad and Ameeta Agrawal	
The Eleventh Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2024) @ NAACL 2024	
<b>Exploring the Maze of Multilingual Modeling</b>	2023
Sina Bagheri Nezhad and Ameeta Agrawal	
arXiv preprint arXiv:2310.05404	
<b>Smoking prevalence in Covid-19 patients</b>	2021
Sina Bagheri-Nezhad, Nasser Mozayani, Elham Abdi, Setareh Rostami	
arXiv preprint arXiv:2107.13599	
<b>Investigation of the smoking prevalence among critically ill patients with COVID-19</b>	2021
Sina Bagheri-Nezhad, Elham Abdi, Nasser Mozayani, Marzieh Mehrafza	
Journal of Control 14 (5), 33-37	

## PROFESSIONAL INVOLVEMENT

---

PC member or Reviewer at following conferences:

- Computational Intelligence journal
- The workshop of Secure and Trustworthy Large Language Models (SET LLM @ ICLR 2024)
- The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)
- 10th Annual International Conference on Algorithms and Discrete Applied Mathematics (CALDAM 2024)
- The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)
- Fourth Workshop on Simple and Efficient Natural Language Processing (SustaiNLP @ ACL 2023)

## CERTIFICATES

---

### Deep Learning Specialization

July 2020

*Issued by deeplearning.ai on Coursera*

### Reinforcement Learning Specialization

September 2020

*Issued by University of Alberta and amii on Coursera*

### TensorFlow Developer Professional Certificate

September 2020

*Issued by deeplearning.ai on Coursera*

## TECHNICAL & COMMUNICATION SKILLS

---

- Programming: Python, Bash, SQL
- NLP & ML Libraries: PyTorch, TensorFlow, Hugging Face, Scikit-learn
- Data Analysis: Numpy, Pandas, SciPy, Matplotlib, Plotly
- Tools: Git, Jira, Bitbucket, Confluence
- Languages: English, Persian