Sarah Alnegheimish

32 Vassar Street, MIT LIDS, Cambridge, MA 02139

smish@mit.edu | sarahmish.github.io | sarahalnegheimish | sarahmish

Education

2019 - present **Massachusetts Institute of Technology** | GPA: 5.0/5.0
S.M. in Electrical Engineering and Computer Science
S.M. in Computational Science and Engineering

2013 - 2017 **King Saud University** | GPA: 5.0/5.0
B.Sc. in Computer Science
First Class Honors

Experience

Sep 2019 - present Research Assistant, **MIT**
Cambridge, MA, USA.

Jun 2021 - Aug 2021 Software Engineer Intern, **DataCebo**
Worked on synthetic data generation using deep learning, with a primary focus on generating time series data.
Cambridge, MA, USA.

Dec 2017 - Aug 2019 Research Specialist, **Center for Complex Systems at KACST and MIT**
Riyadh, Saudi Arabia.

May 2017 - Dec 2017 Data Analyst, **Mozn**
Developed solutions for challenging problems on a national scale, including: policy auditing, fraud detection, and optical character recognition using machine learning.
Riyadh, Saudi Arabia.

Sep 2016 - Feb 2017 Junior Teaching Assistant, **King Saud University**
Courses: Programming I using Java
Riyadh, Saudi Arabia.

Publications

- Liu, D., **Alnegheimish, S.**, Zytek, A. and Veeramachaneni, K., *MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series.* ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), 2022 - [accepted].
- Alhasoun, F. and **Alnegheimish, S.**, *Probabilistic Programming Bots in Intuitive Physics Game Play. 35^{th} AAAI Conference on Artificial Intelligence*, 2021.
- Geiger, A., Liu, D., **Alnegheimish, S.**, Cuesta-Infante, A., Veeramachaneni, K., *TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks.* IEEE Conference on Big Data, 2020.
- **Alnegheimish, S.**, Alrashed, N., Aleissa, F., Althobaiti, S., Liu, D., Alsaleh, M. and Veeramachaneni, K., *Cardea: An Open Automated Machine Learning Framework for Electronic Health Records.* IEEE Conference on Data Science and Advanced Analytics (DSAA), 2020.
- **Alnegheimish, S.**, Alnuhait, F., Albrahim, H., Al-Mogherah, S., Alrajhi, M. and Hosny, M., *An Intelligent Bio-Inspired Algorithm for the Faculty Scheduling Problem.* International Journal of Advanced Computer Science and Applications, 2018, 9(5).

Working Manuscripts

- **Alnegheimish, S.**, Liu, D., Sala, C., Berti-Equille, L. and Veeramachaneni, K., *Sintel: An Overarching Ecosystem for End-to-End Time Series Anomaly Detection* - [under review].
- **Alnegheimish, S.**, Alsuwailem, A., Frank, M., Alabdulkareem, A. and Rahwan, I., *Analyzing the Network Structure for the Arab Standard Occupational Classification.*

Honors & Awards

- Graduate Scholarship, King Abdulaziz City for Science and Technology (KACST), 2019 - 2026.
- Google's CS Research Mentorship Program (CSRMP), 2020 - 2021.
- Graduate Fellow, MiSK, 2019 - 2022.
- 2nd place at AEC annual best graduation project, 2017.
- Best graduation project in the college of Computer and Information Sciences, King Saud University, 2017.
- Best poster at the 9th Undergraduate Research Conference, Zayed University, 2017.
- Dean's list for outstanding students at King Saud University, 2015 and 2016.

Research Projects

Time Series Anomaly detection Sep 2019 - present
Supervised by Kalyan Veeramachaneni *MIT*

- Leading *Sintel*, an ecosystem for anomaly detection, annotation, and feedback integration.
- Developed *Orion*, a python library for end-to-end time series anomaly detection.
- Designed algorithms for time-based anomaly scoring.
- Created a benchmark suite for comprehensive quality and computational evaluations.
- Built a generative adversarial network for reconstruction-based anomaly detection.

Cardea Sep 2018 - Aug 2019
Supervised by Kalyan Veeramachaneni *Center for Complex Systems*

- Developed *Cardea*, an automated machine learning library to solve health related prediction problems.
- Integrated HL7's *Fast Healthcare Interoperability Resources* standard as a representation for data.
- Automated the data ingestion, organization, and featurization components of the framework.

Job Space Dec 2017 - May 2019
Supervised by Ahmad Alabdulkareem and Iyad Rahwan *Center for Complex Systems*

- Constructed an occupation network by discovering job-job relationships through their underlying skills.
- Utilized the job space as a function to predict individual career trajectories.
- Analyzed the impact of automation on job mobility and the education hierarchy.

Class Projects

Gender Bias in Language Models Spring 2021
6.864 Natural Language Processing

- Identified model components (neurons, self-attention heads) that contribute to biased output.
- Proposed mitigation strategies to reduce the discrimination severity of the model output.

Intuitive Physics for Game Play Bots Fall 2019
9.660 Computational Cognitive Science

- Built a probabilistic program to infer actions performed by an agent in Newtonian settings.
- Complemented the program with a model-free approach for efficient sampling.

Presentations & Workshops

Network Science for Data Analysis Workshop 27th Feb, 2019
Women in Data Science (WIDS), Riyadh, Saudi Arabia.

Cadea Platform for Smart Health Analytics 30th Nov, 2018
MIT Hacking Medicine, Riyadh, Saudi Arabia.

S^3: Saudi Skill Space 14th Nov, 2018
Misk Global Forum, Riyadh, Saudi Arabia.

Deep Learning for Image Recognition Workshop 5th Mar, 2018
Women in Data Science (WIDS), Riyadh, Saudi Arabia.

Skills

	Applied Machine Learning, Data Science, Software Engineering
Programming Languages	Python, Julia, R, C, C++, MATLAB, Java, JavaScript, HTML, SQL
	English (native), Arabic (native)