Osama Khalid

+1 (737) 300 7173 • osama-khalid@uiowa.edu
www.osamakhalid.org

Research Interests

Machine Learning, Deep Learning, Natural Language Processing (NLP), Computational Linguistics
Geospatial Data Processing, Stylometry, Sensorial Perception

Education

Computer Science — **University of Iowa**
PhD (Advisor: Dr. Padmini Srinivasan) — *2016 – Current*

Linguistics — **University of Iowa**
Master of Arts — *2022 – Current*

Computer Science — **University of Iowa**
Master of Science — *2016 – 2018*

Electrical Engineering — **Lahore University of Management Sciences (LUMS)**
Bachelor of Science — *2009 – 2013*

Professional Experience

SWE Intern — **Google** — **May'23– Aug'23**
- Worked on fine tuning and distilling generative models to detect risk in Google Search.

Student Researcher — **Google** — **Oct'22– Dec'22**
- Worked on deploying systems to evaluate the risks associated with using Google Translate on medical texts.

SWE Intern — **Google** — **May'22– Aug'22**
- Worked on developing methods to evaluate the risks associated with using Google Translate on medical texts.

ML Developer — **Leepfrog Technologies Inc.** — **Jun'18– Jun'19**
- Created and developed a course recommendation system. Worked on creating systems to optimize college student schedules. Deployed models to help college students with Major Selection.

ML Intern — **Leepfrog Technologies Inc.** — **Jun'17 – Aug'17**
- Prototyped and developed a machine learning algorithm which made use of an ensemble of classifiers to predict the majors of students based on their course history, with an accuracy of 81.38%.

Technical skills

- **Languages:** Python, Javascript, C++ MySQL, R
- **Platform & Tools:** PyTorch, TensorFlow, Weka, Matlab, LIWC

Research Projects

- **Sensorial Stylometry in Language**: This project introduces methods that use deep learning language models to capture the stylometric variations in the language used to encode sensorial experiences.
- **Analysing Bias in Taboo Language Detection**: The project uses a BERT based model to operationalize a community centered definition of Taboo to analyze bias in hate speech detectors.
- **Gender Acquisition in Polish**: In this project we use the CHILDES Polish corpus to model noun-adjective disagreements in grammatical gender amongst Polish speaking children.
- **Investigating linguistic style in Online Communities**: The project uses a broad definition of community style to analyze the linguistic style of online communities on reddit, 4chan and voat.

Publications

Conferences

- Suum cuique: Studying bias in taboo detection with a community perspective.
  Osama Khalid, Jonathan Rusert, and Padmini Srinivasan.
  Findings of the Association for Computational Linguistics (ACL-IJCNLP) 2022.

- Smells like teen spirit: An exploration of sensorial style in literary genres.
  Osama Khalid and Padmini Srinivasan.
  Proceedings of the 29th International Conference on Computational Linguistics (COLING), 2022.

- Style matters! investigating linguistic style in online communities.
  Osama Khalid and Padmini Srinivasan.
  International AAAI Conference on Web and Social Media (ICWSM), 2020.

- No place to hide: Inadvertent location privacy leaks on twitter.
  Jonathan Rusert, **Osama Khalid**, Dat Hong, Zubair Shafiq, and Padmini Srinivasan.
  Privacy Enhancing Technologies (PETS),2019.