Muhammad Haroon

mharoon@ucdavis.edu | (530) 760 9545 | www.muhammadharoon.xyz

 haroon96   𝕏 mharoon_   in haroon96

EDUCATION

Ph.D. in Computer Science   2021 - Present
University of California, Davis

B.S. in Computer Science   2014 - 2018
National University of Computer & Emerging Sciences, Pakistan

EXPERIENCE

University of California, Davis   2021 - Present
Graduate Student Researcher
Advisor: Prof. Magdalena Wojcieszak

Research on privacy and safety harms associated with online systems. I design and evaluate automated software techniques and machine learning models for auditing online systems.

Lahore University of Management Sciences   2019 - 2021
Research Associate
Advisor: Prof. Fareed Zaffar

Worked on privacy preservation and adversarial attacks using ensemble machine learning models and natural language processing for news media analytics. Additionally, I worked on automating workflows for archiving historic content.

Techlogix   2018 - 2019
Software Engineer

Worked as a full-stack developer on HealthCloud - a hospital management system. I improved the nearly hour-long deployment protocol for delivering critical updates to under 10 minutes.

PUBLICATIONS

Re-ranking Using Large Language Models for Mitigating Exposure to Harmful Content on Social Media Platforms   2025
Submission in Progress
Authors: R. Oak, **M. Haroon**, C. Jo, M. Wojcieszak, A. Chhabra
[pre-print]

Nudging the Recommendation Algorithm Increases News Consumption and Diversity on YouTube   2024
PNAS Nexus
Authors: X. Yu, **M. Haroon**, E. Menchen-Trevino, M. Wojcieszak
[paper] [website]

Auditing YouTube's Recommendation System for Ideologically Congenial, Extreme, and Problematic Recommendations   2023
Proceedings of the National Academy of Sciences (PNAS)
Authors: **M. Haroon**, M. Wojcieszak, A. Chhabra, X. Liu, P. Mohapatra, Z. Shafiq
[paper] [website] [code]

HARPO: Learning to Subvert Online Behavioral Advertising 2022
The Network and Distributed System Security Symposium (NDSS)
Authors: J. Zhang, K. Psounis, **M. Haroon**, Z. Shafiq
[paper] [code]

Avengers Ensemble! Improving Transferability of Authorship Obfuscation 2021
arXiv:2109.07028
Authors: **M. Haroon**, F. Zaffar, P. Srinivasan, Z. Shafiq
[pre-print] [code]

ONGOING RESEARCH

Estimating Political Ideology Online using LLMs

We study the extent to which LLMs understanding of politics allows them to estimate the political ideology of online content.

Towards High Fidelity Sock Puppets

Investigating approaches to augmenting sock puppets to make their online behaviors more realistic.

The Prosocial Ranking Challenge

Our personalized quality news up-ranking algorithm is one of ten finalists for the Prosocial Ranking Challenge.

RESEARCH TOOLS DEVELOPED

Personalized Quality News Up-ranker
One of ten finalist algorithms for the Prosocial Ranking Challenge: our algorithm boosts personalized news posts from credible and ideologically diverse news sources to make users more resilient to democratic threats.

YouTube Audit Platform
Identifies the daily top YouTube recommendations for politically left, right, and center users by automatically running several **Selenium**-based sock puppets in **Docker** containers. Web platform built using **Vue**. [website]

HARPO
Web extension that uses a A2C reinforcement learning model backend built using **PyTorch** to obfuscate the user's online profile. [code]

ResearchTube
Web extension for YouTube that subjects users to various algorithmic manipulations and observes changes to their content consumption patterns. [website]

CenterTube
A DQN reinforcement learning model implemented in **PyTorch** that monitors a user's YouTube homepage and systematically manipulates user recommendations towards moderate/neutral news content.

YENET
A Convolutional Neural Network for video frame interpolation. [code]

LUMS Digital Archive
Developed automated workflows for restoring and cataloging historic documents, and developed a platform in **AngularJS** to allow researchers easy access. [website]

YouTube/TikTok Driver
A series of libraries for programmatically interacting with YouTube, YouTube Shorts, and TikTok implemented using **Selenium** on web and **Android Debugging Bridge (ADB)** on mobile. [code]

OTHER PROJECTS *(can be found on my GitHub among other projects)*

ASL Handsigns Recognition
A Convolutional Neural Network implemented in **PyTorch** with a **Tkinter** frontend that recognizes ASL handsigns in real-time using the system webcam.

Scholar Lens
A VSCode Extension that auto-fills BibTeX entries using Semantic Scholar's API. Overleaf version in-progress.

Discord Stats Bot
Developed a Discord bot for video game statistics calculation that organically grew to over 200+ Discord servers.

INVITED TALKS + PODCASTS

Auditing Social Recommender Systems
- Politics and Computational Social Science Conference (PaCSS), Harvard University, 2022 [slides]
- Summer Institute in Computation Social Science (SICSS), University of Pennsylvania, 2022 [slides]
- The Backdrop - A Podcast by UC Davis [listen]
- UC Davis College of Engineering News [read]

Offensive Privacy to Counter AdTech Surveillance
- Nothing to Hide - A Data Privacy Podcast [listen]

SKILLS

Languages: *Python, C/C++, C#, R, Java, JavaScript, SQL, LaTeX*

Platforms/Frameworks: *Docker, Selenium, Git, Node.js, Vue, Android Debugging Bridge (ADB)*

Libraries: *PyTorch, SciPy, spaCy, BeautifulSoup, Pandas, Scikit-learn, Matplotlib, Seaborn, Plotly, Tkinter*

TEACHING EXPERIENCE

ECS 152A: Computer Networks
Teacher's assistant responsible for designing and conducting weekly discussion sessions, and lectures.

AWARDS & HONORS

Finalists ($6,000) — 2024
Prosocial Ranking Challenge

GGCS Summer Research Fellowship ($18,000) — 2024
University of California, Davis

GGCS Summer Research Fellowship ($18,000) — 2023
University of California, Davis

GGCS Spring Research Fellowship ($7,177) — 2023
University of California, Davis

Dean's Honor List — 2018
National University of Computer & Emerging Sciences, Pakistan

REFERENCES

Magdalena Wojcieszak
Professor of Communication
University of California, Davis
mwojcieszak@ucdavis.edu

Fareed Zaffar
Associate Professor of Computer Science
Lahore University of Management Sciences
fareed.zaffar@lums.edu.pk

Zubair Shafiq
Associate Professor of Computer Science
University of California, Davis
zubair@ucdavis.edu

Anshuman Chhabra
Assistant Professor of Computer Science
University of South Florida
anshumanc@usf.edu