

Shaoor Munir

☎ +1 (530) 953-3757 | @ shaoormunir@gmail.com | in [shaoormunir](#) | [shaoormunir](#) | techjuice.pk/author/shaoor

EDUCATION

University of California, Davis

Doctor of Philosophy in Computer Science

Davis, US

September 2021 – 2026 (Prospective)

National University of Computer and Emerging Sciences

Bachelors in Computer Science, 3.73/4.0

Lahore, PK

August 2014 – May 2018

Govt. College of Science

F.Sc Pre-engineering, 1012/1100

Lahore, PK

August 2011 – August 2013

STANDARDIZED TEST SCORES

GRE

321/340 (Quantitative 162, Verbal 159, AWA 5.0)

May 2019

TOEFL

116/120 (Writing 30, Reading 30, Speaking 26, Listening 30)

December 2019

EXPERIENCE

Lahore University of Management Sciences

Research Assistant

November 2019 – Present

- Working in collaboration with academics from The University of Iowa and UC Davis to explore state of the art in Natural Language Generation and Natural Language Understanding
- Working on exploring automated methods to attribute a synthetic text to its source language model
- Collaborating with external and internal resources to work on projects related to news event detection and Urdu language text corpus

Northbay Solutions

Software Engineer

Apr. 2019 – Nov. 2019

- Worked on an internal Big Data Research Project making use of Kylo, Nifi, Apache Spark, and Amazon Web Services. The end product was used as a complete end to end Data Orchestration and Transformation tool for organizations dealing with large amount of unstructured data
- Worked as part of team designing core framework with Java API endpoints for Data Lake management in Amazon Web Services

Systems Limited

Associate Software Engineer

July 2018 – September 2018

- Worked as a trainee in their Machine learning and Data Science team
- Worked on a R&D project for an American bank client, making use of Random Forests, SVM, Word2Vec, Doc2Vec to automatically categorize banking forms and extract useful data

TechJuice

Editor

October 2018 – April 2019

- Managed a writer panel consisting of 15 writers, spread across Lahore, Islamabad and Karachi
- Worked towards improving the editorial quality, hired and trained new writers to expand the existing team

Content Writer

November 2016 – April 2018

- Wrote and published more than 450 articles with one of the largest technology media organizations in Pakistan
- Organized, managed, and carried out interviews for TechJuice 25 under 25 2017 and 2018 editions
- Worked closely with major companies in Pakistan like PTCL, Jazz, Telenor and many more for brand promotions and strategic content production

PROJECTS AND RESEARCH EXPERIENCE

- Protecting Personal Data Flow on the Internet** (Ongoing Research)
- Ongoing multi-disciplinary project which aims to improve transparency and control of how personal data of users flows on the internet through browsing history, data collected by websites, advertisement, and a number of other ways where user data might be stored by first or third parties.
 - Working on developing tools using open-source technologies for measurement studies and automatic detection of privacy policies to ensure privacy of online user data.
- CookieWatcher: Privacy-Invasive Cookie Exchanges on the Web** (Ongoing Research)
- Ongoing research on how third party scripts run in a first party context can misuse this privilege and, along with bounce tracking, can use this privilege to track users across different sites in presence of restriction implemented by modern web browsers on third party cookies
 - Working on analysis of common cookie behavior on popular websites and how it enables cross-site user tracking and multi-device user profiling.
- Through the Looking Glass: Learning to Attribute Synthetic Text** (*EACL 2021*) (Published Research)
- Analysis of different classification techniques to learn how to attribute text generated by a language model
 - Analysis of how different parameters and fine-tuning can effect performance of attributors
- Makhzan** (Urdu Text Corpus)
- Built open-source standardized Urdu Text corpus to help advance Natural Language Processing in Urdu language.
- Newsreader** | *Python, MySQL* (Mainstream Media Research Project)
- Differential analysis of online news in Pakistan encompassing sentiment, topics, and subjects
 - Economic Policy Uncertainty analysis based on frequency of articles across various national newspapers
- K2** | *Java, AWS* (Data Lake Solution)
- Big Data Project to provide end to end Data Lake solution with a plug and play Java API for end users
- Intelligent Mock Interviewer** | *Python, Django, OpenCV* (Final Year Project)
- An intelligent interviewer which used OpenCV, SVMs and Neural Networks to analyze audio and video of an interview to judge the personality of interviewee. The final assessment of the interview was based on the big five personality scale - one of the most commonly used scales in professional environments
- OneBlood** | *Java, Glide, Volley, Firebase* (Undergraduate Group Project)
- A community sourced blood donation mobile application

TECHNICAL SKILLS

- **Languages:** Python, C/C++, Java
- Experience working with Machine Learning models, AI algorithms and Data Science libraries
- Experience working with Amazon Web Services and working knowledge of other cloud services
- Understanding and familiarity with common programming practices and design patterns

REFERENCES

- | | |
|--|---|
| Zubair Shafiq
<i>Associate Professor of Computer Science</i> | zshafiq@ucdavis.edu
<i>University of California, Davis</i> |
| Fareed Zaffar
<i>Associate Professor of Computer Science</i> | fareed.zaffar@lums.edu.pk
<i>Lahore University of Management Sciences</i> |
| Padmini Srinivasan
<i>Professor of Computer Science</i> | padmini-srinivasan@uiowa.edu
<i>University of Iowa</i> |