SARIM ZAFAR

sarim.zafar71@gmail.com
https://github.com/sarim-zafar
+923060040469

PROFESSIONAL AFFILIATIONS

***Data Science Lead, Finja*, Lahore, Pakistan** Nov 2022 – Current

- Successfully implemented and managed the deployment of an Early Warning System utilizing online learning principles to predict loan repayment timeliness and enhance organizational efficiency.
- Facilitated substantial cost savings by designing and implementing an on-prem analytics solution, resulting in a monthly cost reduction of thousands of dollars.
- Empowered data-driven decision-making through the democratization of access to critical organizational metrics.

***Consultant Data Scientist,* Mowasalat, Doha, Qatar** Sept 2022 – Nov 2022

- Developed visibilities for tracking KPIs on PowerBI

***Data Scientist,* Airlift Technologies, Lahore** Nov 2021 – Jul 2022

- Developed a replenishment system from the ground up leveraging Deep Learning
- Improved Dark Ops algorithms to curb stockist activity using Pattern Recognition Algorithms
- Led a team to develop a Churn Prediction and Attribution model

***Senior Data Scientist,* ADDO.AI, Lahore** Dec 2020 – Nov 2021

- Developed a system to help cabbies estimate their revenue.
- Developed a state-of-the-art Singlish Speech to Text model for low-quality audio
- Demonstrated leadership abilities by leading a team to develop a QA automation framework for call centers

***Machine Learning Engineer,* ADDO.AI, Lahore** Jul 2017 – Dec 2018

- Developed multiple AI-powered systems ranging from risky driving behavior detection, and super-resolution of satellite imagery to signature verification using few-shot learning
- Demonstrated viability for customer propensity modeling using historical sales data
- Developed and deployed a Content retrieval system using visual similarity on the Google Cloud Platform
- Developed an email categorization system using Convolutional Neural Networks

***Software Developer,* Vivid Technologies, Lahore** Aug 2015 – Apr 2016

- Developed and maintained a Lead Management System in Laravel and MySQL
- Developed and maintained a Visual IVR system in Angular, Laravel, MySQL
- Developed fast response APIs with low system resource consumption as a constraint in the Go language
- Automated various small tasks using python scripts such as making system backups and cleanups

TOOLS/ SKILLS

Programming language and Tools: Python, C, C++, PHP, Java, Go, MySQL, Laravel, JavaScript, Matlab
Libraries: Scikit-Learn, Matplotlib, Seaborn, Plotly, PostgreSQL, Pandas, Tensorflow, Numpy, NLTK, PyTorch
Data Science and AI Techniques: Supervised/ Semi-supervised/ Unsupervised /Self-supervised learning, Classification, Clustering, Regression, Time-series, Machine Learning, Deep Learning, and Social Networks Mining.

EDUCATION

MSCS, North Dakota State University, Fargo Jan 2019 – Dec 2020

- Thesis: On the applicability of deep metric learning to address source code authorship attribution problem under simulated real-world constraints
- Advisor: Dr. Zubair Malik
- CGPA: 3.79

BSCS, Information Technology University, Lahore Sept 2013 – June 2018

- Thesis: Analysis of Pakistan's Political Landscape on Twitter
- Advisor: Dr. Junaid Qadir
- Selected Relevant Courses: Deep Learning, Artificial Intelligence (CS331), Topics in AI (CS333)

Publications

1. Zafar, S., Sarwar, U., Gilani, Z. and Qadir, J., 2016, November. Sentiment analysis of controversial topics on Pakistan's twitter user-base. In *Proceedings of the 7th Annual Symposium on Computing for Development* (p. 35). ACM.
2. Zafar, S., Malik, M.Z. and Walia, G.S., 2019, September. Towards Standardizing and Improving Classification of Bug-Fix Commits. In *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)* (pp. 1-6). IEEE.
3. Gronneberg, Bethlehem, Joshua DeNio, Anusha Inugurthi, Matti Kariluoma, Nafiz Rifat, Muhammad Sarwar, Sarim Zafar, Brian M. Slator (2020). Bridging the Computer Science K-12 Education Gap by Integrating Conventional Materials with an Immersive Virtual Environment. *Proceedings of the 53rd Annual Midwest Instruction and Computing Symposium*. Milwaukee, WI, April 3-4, 15 pg. online
4. Zafar, Sarim. On the Applicability of Deep Metric Learning to Address Source Code Authorship Attribution Problem Under Simulated Real-World Constraints. Diss. North Dakota State University, 2020.
5. Sarwar, Muhammad Usman, et al. "Multi-label Classification of Commit Messages using Transfer Learning." *2020 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. IEEE, 2020.
6. Zafar, Sarim, et al. "Language and Obfuscation Oblivious Source Code Authorship Attribution." *IEEE Access* 8 (2020): 197581-197596.

Research/Teaching Experience

***Graduate Research Assistant,* North Dakota State University, Fargo** Jan 2019 – Aug 2020

- Developed a Multi Class and Multi-Label Commit Classification through transfer learning
- Developed a Natural Language Interface to SQL databases using CNN
- Developing robust techniques for Code Authorship Re-Identification in large databases

***Research Assistant,* National University of Computer and Emerging Sciences, Lahore** Jun 2016 – Jul 2017

- Implemented different Machine learning models on provided genetic sequences in Affymetrix's microarrays for cancer classification. Developed ensemble models for better performance and used feature selection techniques to get robust and relevant features

***Graduate lecturer,* North Dakota State University, Fargo** Aug 2021 – Dec 2021

- Course: CSCI-122: Programming in Basic - Visual Basic .NET
- Taught the course and was responsible for conducting tutorials as well

***Graduate Teaching Assistant,* North Dakota State University, Fargo** Jan 2020 – May 2020

- Course: CSCI-160: Intro to Comp. sci.
- Conducted lab sessions and TA hours for the students

***Teaching Assistant,* Information Technology University, Lahore** Feb 2018 – Jun 2018

- Course: Deep Learning
- Assisted in the development of quizzes, exams, and homework
- Conducted tutorials and labs for the students to teach them tools such as Tensorflow

Soft Skills

- **Effective Communication**: Demonstrated effective communication skills by delivering effective data-driven insights to technical and non-technical audiences which results in optimized processes
- **Team Player**: Displayed effective team collaboration by working with diverse teams during the different industry as well as academic projects
- **Proactive Problem Solver**: Demonstrated practical problem-solving skills by identifying multiple business use cases and providing pragmatic solutions to get the right answers
- **Business Sense**: Demonstrated Business sense by understanding the organization's problems, providing solutions, and translating the insights into results that work for the organization

Honors And Awards

- PEEF Scholarship: Fully funded bachelor's degree by the ministry of education
- Demonstrated entrepreneurial skills by securing 3rd position in Startup Weekend Lahore '16 competition