# Mirza Elaaf Shuja

elaaf.shuja@gmail.com | linkedin.com/in/elaaf | github.com/elaaf | +92 317 4246496 | Lahore, Pakistan

## TECHNICAL SKILLS

| | |
|---|---|
| **Languages:** | Python, C/C++, JavaScript, SQL, Shell Scripting |
| **Platforms:** | Linux, Google Cloud Platform, Microsoft Azure |
| **Tools:** | Git, Jira, GitHub, Docker, Hadoop, Spark, Talend, Sqoop, HIVE, Vertica |
| **Philosophies:** | Agile Development (Scrum), Test Driven Development |

## EXPERIENCE

**ADDO AI** - addo.ai                                                                            Lahore Pakistan
*Software Engineer - Big Data and Machine Learning*                                         *Apr 2019 – Present*
+ Developed POCs for domain specific Machine Learning use-cases (Fraud Detection, Customer Segmentation)
+ Built scalable Machine Learning model training pipelines using Tensorflow and PyTorch
+ Deployed Machine Learning models to production (GCP, Flask)
+ Developed Spark ETL jobs for Data Warehouse Migration from Teradata to Vertica
+ Developed Sqoop jobs for source data loading into Data Lake on HIVE
+ Optimized Spark jobs to reduce job runtime and cluster resources
+ Unit Testing, Live Scheduling and User Acceptance Testing of Spark ETL jobs

**Information Technology University (ITU)** - itu.edu.pk                                        Lahore Pakistan
*Machine Learning Researcher*                                                               *Sep 2018 – Feb 2020*
+ **Unsupervised Image to Image Translation using Attention guided Cycle GANs**        *Computer Vision*   Github
  Built an Unsupervised Image Gender Swapping system using Attention guided Cycle GANs trained on the CelebA dataset. Architecture
  based on a NIPS 2018 paper tweaked to fit our use case. Achieved better FID score than baseline CycleGAN.
+ **Mis(Informations) Dissemination in semi-closed Social Networks**     *Natural Language Processing*
  Analyzed prevalent MisInformation on WhatsApp and Twitter regarding COVID-19. Statistically differentiated Informative and
  MisInformative textual content and compared message lifetime across both platforms.

**BlueEast** - blueeast.com                                                                     Lahore Pakistan
*Software Engineer - Firmware*                                                              *Apr 2018 – Sep 2018*
+ Development and Testing of Firmware for IOT enabled home appliances
+ Implemented a new firmware extension to store 14 days of IOT sensor data locally using SPIFFS storage
  on the onboard ESP8266 memory (2-4Mb) via Google's protobuf

**Side Projects** - github.com/elaaf
+ **RedScrap:** A wrapper written in Python around the PushShift Reddit API to retrieve Reddit Submissions. Repo
+ **StanceDetect:** A python implementation of a research paper for Unsupervised Stance Detection of Twitter users. Repo

## EDUCATION

**Information Technology University (ITU)**                                                      Lahore Pakistan
*Masters in Data Science*                                                                          *2018 – 2020*
+ Machine Learning Deep Learning Big Data Analytics Computer Vision Information Retrieval and Text Mining Pattern Recognition
+ Offered Graduate Fellowship based on academic performance to top five students (2019-2020)

**National University of Sciences and Technology (NUST)**                                     Islamabad Pakistan
*Bachelors in Electrical Engineering*                                                              *2013 – 2017*
+ Object Oriented Programming Data Structures and Algorithms Embedded System Design Industrial Process Control

## RESEARCH PUBLICATIONS

**A First Look at COVID-19 Messages on WhatsApp in Pakistan -** arxiv - *Natural Language Processing (NLP), Computational Social Science*
A novel study of temporal and statistical properties of prevalent textual and image based information/misinformation in semi-closed social networks
(WhatsApp) in Pakistan with collaborators from **MIT**, **QMUL** and **ITU**.
*Accepted as a Full Paper at the IEEE/ACM International Conference on Social Networks Analysis and Mining (**ASONAM 2020**) December 2020.*