

# Shahrukh Hussain Shaikh

## Software Engineer

An Optimist, Hardworking and Keen Learner always looking forward to challenge myself. My Vision is to solve real world problem using my knowledge and skills in Data Science



✉ shahrukhcoder357@gmail.com

☎ 9820598286

📍 MUMBAI

🌐 linkedin.com/in/shahrukh-hussain-shaikh

## WORK EXPERIENCE

### Dataholic Analytics Solutions Software Engineer

03/2022 - Present

Mumbai

#### Achievements/Tasks

- Built pipeline that processes millions of documents on a monthly basis.
- Model integration into API.

### Cactus Communication NLP/Software Engineer

06/2020 - 02/2022

Mumbai

#### Achievements/Tasks

- **Entity detection** and **automated editing** in Document using **Java**
- Developed Graphene for hypothesis testing using graph based linguistic feature retrieval system
- Automating Inferencing pipeline on AWS using State machine
- Improved Sentences alignment using hunalign
- Develop **rule based system** that enhance the model output linguistically

### Quantiphi Analytics Solution Private Limited Data Scientist

04/2019 - 06/2020

Mumbai

#### Achievements/Tasks

- Developed projects on Image Analytics using computer vision & Deep learning.
- Worked on projects based on Python and SQL.
- Developed various Deep learning & machine learning models in tensorflow, numpy and scikit learn while working in R&D.
- Business Insight from viewership Data provided by BARC.
- Video Shot Extraction using Deep learning.
- Deployed DL Model on Docker and through Jenkins.

## EDUCATION

### Grey Atom -School of Data Science Master In Data science

02/2019

First class

### Bachelors of Engineering, Mechanical Mumbai University, GATE Percentile:-95%

## SKILLS

Python

Computer Vision

Machine Learning

Statistics

Deep Learning

NLP

Flask

Docker

SQL

Java

Maven

Git

Hadoop

Keras

Selenium

Spark

TensorFlow

AWS

Pandas

Tableau

openCV

Numpy

## WORK PROJECTS

### Document Entity Detection

- Built an Entity Detection System using **Java** for DOCX.
- Entities like **Introduction**, **Inline citation**, **Math-Type equation**, **Abbreviations**, **Reference**, **Super and Subscript**, **Bold**, **Underlined**, **Italic**, and similar other entities inside a document.
- Built an automated document manipulation system that edit the document rendering the **Track changes**, **Comment** and **Tip messages** predicted by ML model.
- Prettifying the track changes on document level.
- This system was entirely built on **Python and Java** language.

### Document In-Consistency

- Built a **suggestive system** that not only **detects** the **inconsistency** but also flags it across a **document**.
- It is a **rule based system** using **regex**.
- Inconsistency such as **Hypenation**, **Punctuation**, **Abbreviation**, **Capitalization**, and usage of **UK-US** spelling.

### AWS:Step Function

- Built a pipeline that runs on serverless ecosystem of AWS Step Function.
- Generating Precision, Recall and F1score among Source, Edited and Target Document.
- This automated pipeline even runs on weekly basis processing thousand of documents within hour. Achieved reduction of 90% in time for Analysis

### Business Logic Transfer (BLT)

- Extraction of the Characters, Camera, Location, Color palette related entity into feasible Format from a video at frame and roi level.
- Intershoot pooling of Character in the Mirror Shot to increase its screen time.
- Mapped the whole data with its impression with BARC Data.

### Video Camera Cut

- Production house need to manually cut video before broadcasting it in a country according to its censorship norm and other pertinent criteria.
- Designed an automated way using Deep learning model that cut video at shot level reducing the video editing time by 80%.