

Basar Sohail Chowdhury

✉ sohail.chd0202@gmail.com | ☎ +1-857-930-1243 | Irvine, CA | **Github:** sohailchd | **LinkedIn:** sohailchd

Backend Software engineer with experience in design and development of backend systems and microservices for large scale applications in gaming, finance and data backup & storage domain. Created data pipelines for processing high velocity data in finance, custom tools for application monitoring and automated testing frameworks. During coursework I focused on large scale and distributed systems using map reduce and GPU programming, machine learning projects on CNN for pneumonia detection using chest x-rays images, genomic data processing, NGS.

Skills and Technologies

Languages/Scripting: Python, Java, Golang, Bash, C++, JavaScript

Frameworks: Flask, Anaconda, TensorFlow, Java spring boot, ReactJS, golang web frameworks

Tools/DB: Docker, Redis, Spark, Hadoop, Elasticsearch, AWS stack, GIT, MySQL, MongoDB, Kafka

Skills: Design and development of data processing systems, microservices and internal frameworks and Services, Unit/Automated Testing, Map Reduce, REST APIs, GPU CUDA, Linux

Course work: Statistics, Genome sequencing, Advance Algorithms, Machine Learning, Distributed computing

Academic projects

Hadoop MapReduce and spark:

- Implemented PageRank algorithm on 40GB of data using hadoop map reduce paradigm in AWS EMR clusters
- Goal was to understand distributed computation using map reduce paradigm and its limitations
- Created a data processing pipeline using spark to understand the basics of distributed realtime computation using spark on big data (10GB of followers data) from Twitter
- Goal was to understand the real time processing and understanding of spark computations and RDD concepts for in memory computation

Histogram Equalization using GPU CUDA:

- implemented Histogram Equalization using optimized GPU code in CUDA that provides the same functionality of the histogram equalization from OpenCV but can perform the algorithm faster

NGS sequencing of genomics data:

- Created a NGS genomics pipeline for sequencing of genomic data and annotation using BioPython.
- Used bioinformatics tools for linux and python as scripting to create the pipeline and R markdown for reporting

Pneumonia Detection using radiograph images:

<https://github.com/sohailchd/PneumoniaDetection>

- Goal was to detect pneumonia infection in patients using convolutional neural network (CNN) of radiograph images and create a prediction model to correctly identify the infection in patients
- Achieved about 96% accuracy on augmented datasets of 20GB, used google colab and python for implementation, training and predictions

Work Experience

Verys (SDE)

California, USA June 2019 to Dec 2019

- As a backend engineer created core system APIs for flight reservation system to be integrated with cruise
- Frontend UI development and integration with the backend using ReactJS, flight search using Solr

- Middleware components for processing of flight booking objects and creating adapters for cruise system

Technologies: *Java, Spring, AWS, MYSQL, ReactJS, Solr search*

Druva (SDE / Data backup and discovery / Unicorn startup)

India, Jan-2017 to Aug-2018

- Developed APIs for endpoint backup and storage related services (AWS EC2)
- Implemented cloudapps integration (Exchange mail/GDrive/Box) backup and restore APIs
- Designed and developed microservices on AWS for parsing data after backup for compliance (HIPAA/PHI/Legal Hold)
- Created Asyn APIs for uploading data from desktop clients to block storage service

Achievements: Optimization of custom parser, reducing the data parsing from 4 hrs to 30mins

Technologies: *Python, Golang, AWS, Terraform, Elasticsearch, DynamoDB, Flask*

GSLabs (SDE / Chicago mercantile exchange)

India, July-2015 to Jan-2017

- Design and implement large scale microservice based architecture of processing finance transaction
- Cloud migration of the legacy applications to spark for real time data processing(40k messages/sec)
- Talking to clients and stakeholder, managing a team of 5 people and growing the project
- Created a container based distributed scheduler for automated testing of UI and APIs using selenium. Used round robin strategy to schedule jobs and tests using a shared master file.

Technologies: *Java, Golang, Spark, Hadoop, Redis, MongoDB, Linux, AWS EMR, EC2, S3, GIT*

Ubisoft (Test Engineer / Gaming and Entertainment)

India, July-2013 to July-2015

- Feature improvement/testing of payment gateway integration API for in game purchases
- Developed Internal tool for logging crash logs from PS4 dev kits to Jira using Sony SDK using python
- Debugging and unit test cases for PS4 games ('The Crew' and 'Assassins creed black flag') in C++ and system tools in linux

Technologies: *C++, Python, Network packet analysis, CI/CD automation, Linux*

Education

1. Master of Science, *Bioinformatics*, (Northeastern University, Boston) *May, 2020*
2. Bachelor Of Engineering, *Information Technology*, (University Of Pune, Pune) *May, 2013*