

Rajnish Bhatia

Big Data Engineer

To enhance the learnt skills through exercising them on ultimate platforms. Always passionate with strong growing and development skills to work with new technologies

✉ rbhatia.bs15seecs@seecs.edu.pk

📍 Umerkot, Pakistan

☎ +92 322-3941535 & +92 331 3476339

🌐 [linkedin.com/in/rajnish-bhatia-824b59126](https://www.linkedin.com/in/rajnish-bhatia-824b59126)

WORK EXPERIENCE

Data Engineer Turing.com

06/2022 - Present

Remote

Achievements/Tasks

- Working on improvements of products and processes along with Optimizations

Consultant Data Engineer - (Offshore Consultant - Regeneron Pharmaceuticals USA)

Visionet Systems Inc (Systems Limited)

06/2019 - Present

Lahore

Achievements/Tasks

- Worked on the GCP (GCS, Big query, Bigtable, Cloud functions, etc). Worked as combo of Devops Engineer including Operations automation.

Intern (Data Science team)

MTBC (Medical Transcription and Billing Company)

08/2018 - 09/2018

Islamabad

Achievements/Tasks

- Implemented a research paper algorithm for the process use. Worked on JavaScript modules for Image processing.

PROJECTS

Final Year Project

- Verification of In Air Handwritten Sequences (Sigm8) : Model that verify that signature is genuine or forged based on the sequential data that is result of user arm movement in the air wearing Myo arm band. The project is research work

GCP Project (Mattress firm)

- An Application that notifies Google pub sub on getting an email with csv attachment on authorized account and pub sub triggers an event based function that takes data from email and upload it on Google Cloud Storage, on uploading a file another function is triggered that takes file from GCP storage and uploads into a Big query table in particular dataset

Regeneron Pharmaceuticals

- Working on data Ingestion in different formats (excel, csv, txt, parquet, .gz, etc) from different sources (SFTP, DBMS, Cloud storages) into data lake through Apache Nifi. Cloud services used is AWS mainly (Athena, S3, Redshift, EC2).
- Working on Post Ingestion which includes ETL (data hardening, etc) and places it onto given storages, creates some external tables in Hive or other Databases and creates reports on the top of processed data. This task is achieved by PySpark scripts. These Spark jobs are scheduled, monitored and orchestrated on Apache Airflow. Some data is also loaded into Amazon Redshift.
- Working on Dynamic Data dictionary approach for Post Ingestion scripts that requires minimal code changes resulting in quick deployments. This includes automation of the processes. Focusing on Spark code and processing optimizations and automations to minimize the usage of cluster resources resulting in parallelism work approach.

EDUCATION

Bachelors of Computer Science NUST, Islamabad

09/2015 - 05/2019

CGPA : 3.73

SKILLS

Google Cloud Platform (Dataflow, Bigquery, GCS, Bigtable, Cloud Functions)

Apache Airflow

Machine Learning skills

AWS (Athena, RedShift, S3, EC2)

Big Data Tools

Hadoop

PySpark

Hive

Python Problem Solving

Strong Leadership skills

JIRA

Release Management

Data Science

Python Modules

Computer Vision

Keras

Tensorflow

Data Structures

Database

Algorithm Optimization

Apache Nifi

DQ modules

OOP

Sharp Analyzer

Strong Leadership skills

ACHIEVEMENTS

1. Microsoft Virtual Internship (07/2020 - 08/2018)
worked on the tasks given with some modules, created to check the intellect.

3. Problem Solving Silver Badge
Solved many of the different algorithms with useful techniques in limited lines of code

4. Winner of AI Boot camp (SEECS, NUST)
(08/2018 - 09/2018)
worked on the FYP with some techniques`

CERTIFICATES

Certified Resource Person and Trainer at OCEPS
(05/2013 - Present)
It is organization that feeds the creativity of students

Position in Writing Competition (SUPARCO)
(11/2014 - 11/2014)

Coursera: Serverless Data Analysis with Google BigQuery and Cloud Dataflow

Cognitive Class AI: Spark Fundamentals I
Cognitive class. ai course that goes through all the fundamentals concept of spark and makes user practice well on its components