Sumit Sahay
Data Engineer/ Scientist

A passionate, hardworking and aspiring Data Scientist who's seeking to leverage his analytical and programming skills to increase productivity and meet the needs of your company.

sahaysumit06@gmail.com ✉
+91-9829211249 📱
Gurugram 📍
linkedin.com/in/sumit-sahay 🔗
github.com/sumitsahay68 🔗

WORK EXPERIENCE

Programmer Analyst
Cognizant Technology Solutions

09/2019 - Present   Bangalore, India

Achievements/Tasks

- Worked on Network Traffic Monitoring & Analysis for EdgeCloudNF 5G Architecture R&D Project.
- Worked on the development of an STB Test Automation framework for running automated tests on STB/OTT Platforms.
- Built a Computer Vision-based STB Test Automation Framework using Python and OpenCV. Improved accuracy of the framework by 30%, by using Tesseract along with edge detection.
- The deployed version of the framework is able to perform automated tests on various STB/OTT Platforms with high reliability.
- Got trained and completed courses through Databricks Academy on PySpark, ETL and Data Engineering concepts.
- Utilized Spark DataFrames and Spark SQL API for faster big data processing and various ETL processes as training hands-on.
- Worked with dimensions table and performed QC on various datasets at various levels of granularity.
- Attended AWS Developer training and got certified as AWS Developer Associate.

Full Stack Developer Intern
Cognizant Technology Solutions

02/2019 - 05/2019   Bangalore, India

Achievements/Tasks

- Completed training on Java, SQL and Shell Scripting.
- Gained valuable training and experience in various Full Stack technologies which included Java, Angular, Oracle and Python.
- Collaborated with team to develop a Web Application based on Angular, Python and Oracle SQL for Database Audit as the internship project.

EDUCATION

Postgraduate Diploma
IIIT Bangalore

12/2020 - Present

Courses

- Machine Learning and Artificial Intelligence

B.Tech
Chandigarh Engineering College

06/2015 - 05/2019   8.0 CGPA

Courses

- Computer Science Engineering

SKILLS

Python · SPARK · SQL · Databricks · Regression Models · Classification Models · Ensembling Models · Model Selection · OpenCV · Image Processing · AWS · ETL · Computer Vision · Machine Learning · EDA · MS Excel · Java

ACHIEVEMENTS

CISCO NETRIDERS CCENT 2017
Ranked 229 in ASIA Pacific and Japan region in CISCO NETRIDERS CCENT 2017.

Global Innovation Challenge 2020 (Cognizant) (12/2020 - 01/2021)
Participated with the team and presented the idea to global level judges. Got selected for the prototyping phase. Developed and demonstrated the prototype and was selected among the top 5 ideas in BU.

Received Silver Award "Seeking data, building knowledge" (04/2021 - Present)
Awarded for the investment of time in learning initiatives and contributions to the BU from Global Delivery Head.

PROJECTS

Image Compression Using Machine Learning (06/2018 - 07/2018)
- Implementation of K-means clustering Machine Learning algorithm for the purpose of compressing an image by reducing the number of colors in the image to only those which are most common in it.

House Prices - Advanced Regression (Kaggle Competition) (05/2021 - 05/2021)
- Built regularized regression models to understand the most important variables to predict the house prices in Australia. Submission ranked in Top 30%.

Telecom Churn Case Study (05/2021 - 05/2021)
- Solved the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn. Tested multiple models to find the best predictive model. Also built additional models for good interpretability.

Gesture Recognition (Deep Learning) (07/2021 - 08/2021)
- Recognize hand gestures using Neural Networks. Experiment with models to find the model with high accuracy, low latency and low memory footprint. Experimented with Conv3D, Conv2D+LSTM/GRU, and Transfer Learning + LSTM/GRU model architecures.

CERTIFICATES

AWS Developer Associate (DVA-C01) (12/2020 - 12/2023)
AWS Developer - Associate certification demonstrates your technical expertise in developing and maintaining applications on the AWS platform.

Machine Learning (08/2018 - Present)
Authorized by Stanford University and offered through Coursera

Kaggle - Deep Learning (07/2020 - Present)

Apache Spark Programming with Databricks (10/2020 - Present)