



Name: Vaibhav Kunwar
Designation: Data Engineer, Accenture ATCI, Bangalore
Mobile: +91-8009220801
Email: vaibhavkunwer@gmail.com



ACADEMIC DETAILS

Year	Degree / Exam	Institute	GPA/Marks (%)
2018	MTech in Opto-Electronics & Optical Communication	Indian Institute of Technology Delhi	75.29/100
2016	B. Tech Engineering Physics	Delhi Technological University	71.09
2012	CBSE	St Anthony School	88.6/100
2010	CBSE	St Anthony School	9.4/10

ABOUT

Experienced Data Engineer with 2+ years of working in the Information Technology and Services Industry. Skilled in Python, PySpark, SQL, CI/CD, ETL, Data Transformation and Data Collector Pipeline. Data Lake Architecture and Amazon Web Services (AWS). Proficient in Big Data Tools and Technology with industry level projects.

WORK EXPERIENCE & PROJECTS (Sept 2018-Present)

1. Data on Cloud

- Experience using infrastructure as code (IAC) tools service like **AWS CloudFormation**.
- Configured and deployed cloud services leveraging an **infrastructure as code (IaC) approach using AWS CloudFormation Templates**.
- Deployed AWS cloud services like **EMR, Redshift, Glue, S3, Kinesis Streams** using CloudFormation Template
- Knowledge on **CI/CD tools like AWS Code Deploy, Code commit and Codebuild**
- Configured build and released CI/CD pipelines for Code pipeline merge automation using CloudFormation Template for AWS services like **S3 EMR, Redshift, Glue, Athena, Streams, Firehose, Kinesis Analytics and DynamoDB**
- Worked on Command line integration of AWS with **CloudFormation Template**.
- Experience with Serverless platforms like **AWS Lambda** for making merge automation functions for use in **Code pipeline Automation and creating IAM roles for AWS services in CloudFormation Template**.
- Data lake Architecture

2. Splice Machine

- Data Visualizations on **Tableau and MicroStrategy** for importing databases on Tableau dashboard.
- Developed **Python Application for ODBC Connection** for Database performance Testing for Splice Database.
- Developed **Java Application for JDBC Connection** for Database performance Testing for Splice Database.
- Setup a **Cloudera Hadoop ecosystem** to store and manage network Traces with analytics using Tableau.

3. British Telecom

AWS

- Data migration from Collector Sources to AWS Redshift Datawarehouse.
- **CI/CD** for services like Redshift, Glue Job, Crawlers for Merging across different Environment.
- **Data Pipeline** for **ETL** in AWS Glue Jobs.
- **Lambda function**-Python for automating Glue Job status report generation and Logging Framework.
- Automation using Lambda Function for Sensitive Data Tagging of S3 data object.
- Design and Built ETL pipelines to automate ingestion of structured data.
- Streamsets Transformer for Data Transformation.
- Streamsets Orchestration Pipeline for Data Transformation

GCP

- Data Pipeline development on GCP with Airflow/Composer
- ETL
- Big Query

CERTIFICATIONS

- **AWS Developer Associate**, April 2019
- **Machine Learning with Python** on Coursera, IBM

TECHNICAL SKILLS

- **Programming tools:** Python, JavaScript, SQL ,Java ,Shell
- **Databases:** IBM DB2, Splice dB, NOSQL, MongoDB ,Redshift
- **Machine Learning:** Exposure to ML algorithms for classification like KNN, Logistic Regression, Decision Trees. Clustering algorithms -K means. Multiple Linear Regression.
- **ML Python Libraries:** Matplotlib, Tensorflow, Numpy, Pandas
- **Platforms:** AWS