Aryan

**Professional
Summary**

- Data Analytics Professional with **'4+ years'** of experience in the Data Analytics domain.
- Conducting Process Study to enable Automation using Python for business Reports for daily, Monthly.
- Hands on Experience in **Python, PySpark**.
- Good Knowledge of ML Algorithms.
- Hands on Experience on **Linux** based **OS.**
- Working Knowledge of AWS EMR Service, S3, AWS EC2 Instance

Qualification:

- PG Diploma in **Big Data Analytics** (2016-2017) from **Centre for Development of Advanced Computing, ACTS (HQ) Bangalore** with **68% aggregate.**
- B.TECH in **Computer Science & Engineering** (2012-2016) from **Uttar Pradesh Technical University, Lucknow** with **65% aggregate. Skillset**

Scripting & Programming Languages	Python
RDBMS / NoSQL Databases	MySQL, MongoDB
Machine Learning Libraries(Python)	Pandas, Numpy, NLTK, Scikit-Learn, pandasql
Big Data Tools	Apache Spark, Apache Kafka, Hadoop
Python Libraries	pymongo, pymysql, pandasql, urllib, selenium, os, numpy, pyodbc
OS/IDE/Visualization Tool	Ubuntu, RHEL, Jupyter Notebook, DBeaver, Apache superset

Employment History

Employer	Role	Duration
Kestone IMS Pvt Ltd (A CL Group Company)	Data Analytics Engineer	Aug 2019 - Till Date
NMSWORKS Software Pvt Ltd	Data Analyst	July 2017 - Aug 2019

Projects:

Project Executed at Kestone IMS Pvt Ltd (A CL Group Company)	
Project Name	**Deeper Insights of Virtual Event Platform**
Summary	The goal of this project is to find the behaviour pattern of delegates(attendees) during the virtual event using Stats concept.
Tools & Environment	Python, MSSQL, PySpark, Jupyter Notebook, Flask, AWS
Responsibilities	I have involved as Data Engineer • Performed Data cleaning and transformation processes on the event log data. • Storing the cleaned data into MSSQL, • Automated different kind of depth reports using Python, MSSQL, • Wrote SQL Query for the generation of different Reports, • Deployed different Script on AWS using nginx , gunicorn, Flask.

Project Name	**Summary of all the Virtual Event**
Summary	This Project was the R&D on the Virtual Event Platform Data. The goal of this project is to find the behaviour pattern of delegates(different Level) in Virtual Platform based on different levels of Attendees for example: finding the resource consumption pattern in different levels of groups.
Tools & Environment	Python, MSSQL, PySpark, Jupyter Notebook

Responsibilities	I have involved as Data Engineer
	• Combined more than 240 Event data in one dataframe,
	• Used Python for visualization of end results ,
	• Performed Data cleaning and transformation processes.

Project Name	Job Level Classification Model
Summary	The goal of this project was to **predict** the **Job Level(**CXO, CXO-1, CXO -2, Managerial, Below Manager) based on their current designation. For Event Insights, we need this Tag for summarization
Tools & Environment	Python, MSSQL, Jupyter Notebook, Machine Learning
Responsibilities	• Cleaned the Text data
	• Applied Classification Algorithm to predict the Job Level.
	• Deployed Flask API on AWS Server.

Project Name	Web Scraping
Summary	Scraping website data from different websites stored in a database.
Tools & Environment	Python, MySQL, Jupyter Notebook, BeautifulSoup, Selenium
Responsibilities	• Scrap more than 20 websites like Fundoodata, Linkedin, Ackama, CollegeDunia, careers360, Acma.

Project Name	CRM Data
Summary	The goal of this project was to combine all the Event data(stored in Excel file) of 2017,2018,2019 and to clean and store in Structure format in DB so that it can be used for predictive analytics.
Tools & Environment	Python, MySQL, Jupyter Notebook, pandas
Responsibilities	• Performed Data cleaning and transformation.

Project executed at NMSWORKS Software Pvt Ltd.

Project Name	BSNL (BSNL BroadBand Network Management)
Summary	Mange all the device from bng to dslam using company tool "Cygnet NG" and collecting and analyzing performance and reliability of (BSNL Broad-Band) network using SSH, Telnet, SNMP. My roles and responsibility included to automate the collection of health data for all the devices deployed in BSNL network and generate reports. I used Python scripts and SQL queries and

	networking concepts to achieve this.
Tools & Environment	Python, MySQL, Jupyter Notebook, MongoDB, PySpark, Pandas

Responsibilities	- Automate the collection of health data for all devices deployed in BSNL Network, - validated the Data from MongoDB and MySQL DB using Python, - Collected the data from the devices and cleaned the data in required format for validation of Discovery using Python, - Automated the login validation using python, - Generated different kind of reports using python, - Automated the Audit of Network devices installed in BSNL Enterprise for this project like HighEnd Server, Switch etc, - Wrote different kinds of Reports by summarizing the Quarterly data, monthly and Yearly data using Python, pandas.

Project Name	Log Collector
Summary	The goal of this project is to collect all the server logs. It helps the team to debug failure.
Tools & Environment	Python, Kafka, PySpark
Responsibilities	I have done this project end to end, Collected Logs using kafka and parse to Spark for further cleaning and storing into MySQL.