AMRITANSH BISHEN
VPO-Chhotka Rajpur
Distt: BUXAR, BIHAR
Contact No: +91-9873214524
E-mail: abamritanshbishen@gmail.com

Summary:
* 8+ years of working experience and 5+ years in the field of Bigdata and **AWS Certified Developer Associate.**
* Well-versed with Big Data Hadoop, Hive, Spark, Python, MapReduce, Oozie, Talend, AWS Services-EMR, EC2,S3, Elasticsearch, Logstash, Kibana, Glue, Athena, Airflow, Autosys, Shell scripting.
* Good Hands-on experience with the Hadoop eco-system framework (HDFS, Hive, Spark, Kafka, Sqoop, Python, SQL, NoSQL db's)
* Develop, validate and maintain HiveQL queries, Pyspark queries, Glue Jobs, Shell Scripting, Airflow Dag creation, Autosys jil creation.
* Designed DDL and DML in Hive/Imapala, Athena.
* Worked on Talend open source edition, Informatica MDM, Delphix MDM (Masking), Azure Synapse, Azure Data Explorer.
* Capable of processing large sets of structured, semi-structured and unstructured data and supporting systems, Streaming data in spark cluster.
* Working experience of Agile Methodology.
* Good communication and documentation skills, having ability to learn the fundamental concepts effectively and efficiently
* Strong knowledge on SAS / R / PYHTON programming.
* Clustering R & SAS (Iris Data), Multiple Linear Regression R & SAS (Real Estate), Logistic Regression R & SAS (Health Care), Decision Tree Model R & SAS (Sales prediction for Car Seats in a continent) and Market Basket Analysis R.
* Projects done for R, SAS & Python – Telecom, Banking Sector, Internet, Aviation Industry.

Career Objective:
To obtain position of responsibilities that utilizes my skills and experience and keen to work in an environment where I can enrich my knowledge. I wish to succeed in an environment of growth and excellence and earn a job which provides me job satisfaction and self-development and help me achieve personal as well as organizational goals.

Professional Experience:

1. **Sr. Software Engineer- Iris Software**
   May'21 - till date

Project Name – CitiBank MarketRisk
Development Tools: Hive, Spark, SQL, Python, Pyspark, Shell, Autosys

Description:
Working under Data Engineerinrg team of Citi Bank, primary task is to move data from Oracle side to Hive for Other teams so that there analysts can utilize the data from Impala. Write pyspark jobs and sql views for ETL and data availability and data science team.

Responsibilities:
- Write pyspark code for moving tables from oracle to Hive and vice-versa in periodic approach.(Batch processing)
- Optimize the jobs for better performances and debug issues faced.
- Prepare Autosys jils for each tables, individual or box level.
- Prepare templates for new tables for Talend team for moving those tables.
- Debug the daily base issues happening in the pipeline and ensure smooth running of all jobs, peform Data Quality checks.
- Coordinate with all teams and understand there requirements and prepare the ETL for them and onboard there tables into the cluster.
- Time to time work on POC for tech stack selection.

2. Elastic Search Consultant–Vcare Infotech
   Mar'20 – May'21
Project Name: Airtel Payment Bank Limited
Development Tools: JOCATA(AML/FRMS), SQL, Kafka, Spark, ElasticSearch, Grafana, Shell scripting, Airflow

Description:
APBL is using third party tools for compliance and fraud management for submitting its report to FIU and RBI. Primary Job is to make sure all the services are up and running and assist business in generating required reports every month, as per requirement. This data is being consumed by Data Science team for further development and being part of Data Science team assist them in creating useful reports.

Responsibilities:
- All the services of 3rd party tool should be up and working in 24/7 environment.
- Anti-Money Laundering and Fraud Risk Management system rule creation and validation.
- 100+ reports are running through airflow, Create Airflow DAG, monitor and maintain it.
- FRMS Spark cluster and all micro services should be up and maintain them.

3. Sr. Big Data Developer - Shree Infotech Pvt Ltd
   Oct'19– Mar'20
Project Name: Crawford& Co. (EDL)
Development Tools: Spark, Python, Pyspark, AWS Glue, AWS Athena, Redshift and Git.

Description:
Creating EDL for Crawford and Co. Under EDL, we have created end-to-end pipeline, which provides Data management, data ingestion and data processing. In addition, we have integrated Data Engineering with Data science in this project.

Responsibilities:
- Import the historical as well as Incremental data from Oracle, Mysql DB's.
- Written Pyspark Job under Glue hood for the data movement and cost reduction.
- Stored the data and made Tables on top of them from Athena.
- Created view on Athena side for Data Science/Analytics team.
- Scheduled all the jobs on daily basis also data cleansing jobs on biweekly basis.

4. Sr. Hadoop Developer – TI Technologies Pvt Ltd(RCG INDIA) Oct'18 – Oct'19

Project Name: RCG Sentimental analysis and Real Time Visualization
DevelopmentTools: AWS, Hive, HDFS, Kafka, S3, Pyspark, Talend, ElasticSearch, Logstash, Kibana.

Description:
Perform analysis on near real time data, live streaming with the help of Spark and Python and then showing the response in Elasticsearch Kibana visualization with the help of Live Dashboards. Pull the data from Source with Kafka and store in EDL(S3) and for cleaning in HDFS then apply all the Algorithms, Web Scrapping, Text Blob, Ngrams, TF-IDF etc. and calculate sentimental polarity.Built End to end pipeline and deployed on AWS.

Responsibilities:
- Created Kafka Producers, consumers, brokers and incorporated.
- Created HIVE tables and stages for data cleaning and stored 1 copy of raw in S3.
- Written Pyspark code for Reviews sentiment analysis and sent the data into ELK for Visualization.
- Created the workflow in Talend for easy deployment in any services.
- Created Kibana Dashboard for real time visualization of the score generated on reviews.

Project Name   : NCR Data Ingestion
Development Tools:HDP, Hive, HDFS, Sqoop, JIRA, Intellij, Confluence

Description:
Finance Data of the organization was generated from multiple sources and data have to be collected from different data Sources i.e. NetSuite, ERP Clouds, Synergy, Manual and after that change the format in csv then store it in EDL.

While the data enter in EDL perform DML's on every staging as per Organization standard.built EDL and the layers so that Manipulation can be tracked and then in last store the required data in EDW.

Responsibilities:
- Created DDL for all layers of EDL and DML's for the same.
- Import data and change the file formats of received data and then push it into HDFS.
- Taking care of JIRA tickets raised for bug issue or project tracking.
- Created workflows for all the level DML's.
- Developed advance level hive queries to meet the requirements of different layers.
- As per business demand made data available in EDW/RDMS's for daily needs.

5. Hadoop Developer – Inferlytix Pvt Ltd   Feb'17 – Oct'18
Project Name:ETL and Data Analytics
Development Tools:SQOOP, PIG, HIVE, MySQL, Oozie

Description:
Developed complete ETL framework for migrating the data from different sources to Hadoop Ecosystem. Development of a neat data-driven framework, which can handle not just single migration of data from RDBMS to Hadoop Ecosystem, but can also handle recurring incremental dumps from the RDBMS. Automatic scheduling of the jobs to trigger this migration activity. Capture Audit Information.

Responsibilities:
- Writing the workflows/Oozie jobs
- Importing data into HDFS and Hive from RDBM's.
- Involved in creating Hive tables, loading with data and writing Hive queries which will run internally in MapReduce way.
- Perform ETL/Incremental append operations on Hive table data using HQL.
- Perform Analysis on data as per customer requirement.

Business Analyst – GoIbibo-MMT   Mar'15-Feb'17
- Optimization report with other OTA's (Reduced burn from -27% to -7%)
- Handling queries for Sister concern companies (RedBus, TBO, Seat Seller)
- Spoors Audit for BDM's(India) and Audit for Market Managers reports.
- Spoors Data structuring & mapping the market Visits of BDM's and MM's.
- Crawling data study on Disparity on market & Hotel level

Academic performance:
- BA from Manav Bharti University 2015.
- Intermediate from central board of secondary education – 2008.
- High school from central board of secondary education – 2006.

- Languages – English, Hindi.
- **Interest and Hobbies -** Listening Music and Enjoy travelling.

Description:
I consider myself familiar with Computer Science aspects. I am also confident of my abilities to work in a team.
I hereby declare that the information shared above is true to the best of my knowledge and belief.

14th Dec' 23   (**Amritansh Bishen**)