

Name

Personal Details:

Location: Bangalore
Cell: +91-1212121212

Email: xyz@gmail.com

LinkedIn: <https://www.linkedin.com/in/xyz/>

Experience Summary:

Highly motivated seasoned data engineer serve as a key technologist with strong technology skills, commitment to quality, positive work ethic and a "can do" attitude, responsible for scripting using Python, Spark, Bash and migrating the data from unstructured to semi/structured data& Quality of the products, code, and making sure the system is built accordingly. Strong end-to-end feature development background, Delivery, Standards and Industry Alliances.

Demonstrated Technical Capabilities

- Very Strong Data Engineering skills ability to craft clean interfaces and operate at the right levels of abstraction
- Solid coding skills with ability to drive teams through massive refactoring exercise & improve coding standards across large code bases
- Deep knowledge, understanding & experience of working with a large variety of multi-tier architectures. Awareness of pitfalls & use cases for a large variety of solutions
- Deep understanding & experience of high-performance scale & real-time response systems
- Experience & expertise in a variety of large-scale persistent systems including large databases
- Scalable product development cycles with large data handling
- Demonstrated expertise in problem solving and technical innovation

Demonstrated Management Capabilities

- Promotes effective communication, teamwork and mentor
- Liaison for communications/issue resolution with internal and external teams
- Develops comprehensive plans for methodology and deliverables

Summary of qualifications

- Over 3+ years of Comprehensive experience working as Data Engineer
- Proven record of deliverables through Agile& Kanban
- Worked with scrum master on defining the requirement and priority
- Delivered large and complex projects with business standards effectively
- Owned and manage commitments to the business
- Passion to learn new technologies

Education

Bachelor of Technology (Information Technology) 2011– 2015
Saveetha Engineering College, Chennai – TN

Technical Skills

Big Data Technologies: HDFS, MapReduce, HiveQL, Sqoop, PySpark, Spark SQL, Kafka, MongoDB,
Air Flow

Platform : Windows and UNIX

RDBMS : Oracle 10G/11G & MySQL

ETLTools : Informatica Power center 8x/9x.

Programming Language: Python 3X

Awards:

- Received **Gold Excellence Award** for developing engine agnostic data science life cycle to predict sulphidation impact of Trent engines from Rolls Royce.

Certifications:

- ITIL Foundation Certificate in IT Service Management.
- IBM Hadoop Certification.

Professional Experience

Rolls-Royce through LTTS-Bangalore Data Engineer

01/2019-10/2019

Handled the responsibilities for advanced data analytics, industrial Artificial Intelligence and machine learning techniques, R2 Data Labs develops data applications that unlock design, manufacturing and operational efficiencies within Rolls-Royce, and creates new service propositions for customers

- Working with on a live 10 nodes hadoop cluster running on HDP and applying Data Engineering in Aerospace Domain.
- Built Spark scripts by utilizing pyspark depending on the requirement.
- Developed python scripts for Sortt and Maximo Data sources.
- Experienced in creating features engineering for Trent_700 and Trent_1000 engines.
- Multi pipeline architecture design for ingestion, ETL and analysis of more than 10 different impression data sources.
- Work collaboratively with Data Science team members to develop algorithms to automate data consumption processes

DXC .Technology, Bangalore Big Data Developer

04/2016-01/2019

Integrated vast amounts of structured and unstructured data across research and development, supply chain, customer-facing operations, and customer interactions, both from traditional data sources and new sources of online data. It supports the company ability to complete analysis on multiple data sets quickly.

- Installing, Configuring Hadoop Eco System and Cloudera Manager using CDH Distribution.
- Worked on a live 30 nodes Hadoop cluster running CDH5.
- Worked with highly structured and semi structured data of 90 TB in size (270 TB with replication factor of 3).
- Built Spark Scripts by utilizing Pyspark depending on the requirement.
- Experienced with Spark Context, Spark SQL and Data Frame.
- Extracted the data from Oracle into HDFS using Sqoop.
- Created and worked Sqoop (version 1.4.3) jobs with incremental load to populate Hive External tables.
- Developed Hive (version 1.2) scripts for end user / analyst requirements to perform ad hoc analysis.
- Very good understanding of Partitions, Bucketing concepts in Hive and designed both Managed and External tables in Hive to optimize performance.
- Designed, developed and did maintenance of data integration programs in a Hadoop and RDBMS environment with both traditional and non-traditional source systems as well as RDBMS and NoSQL data stores for data access and analysis.
- Experienced with performing CRUD operations in HBase.

Technical Support Engineer

10/ 2015 – 03/ 2016

- Oversee range of company's application, providing design, technical support and problem analysis and resolution. Involved Manual Testing in SIT and UAT.
- Use system metrics to compile data for analysis and identification of system issues.
- Ensure optimal performance of applications, investigating and resolving operational issues via range of media channels, including phone, chat and meeting.
- Work with R&D Team to ensure compliance of application use and to address issues.