Prasad Ambati

Data Engineer

Email: ambati0809@gmail.com

Phone: +91-9392328984

Professional Summary

Technically sound and result oriented professional with 6.5 years of total experience and 3.5 years relevant experienceas as a Data Engineer in the field of on-premises and Excellent communication and interpersonal skills with a willingness to pick challenging roles.

Big data development had made me learn how to create information from data. You see numbers and letters; I see meanings and insights

- Expertise in working with large datasets and insights within structured and unstructured data of Hadoop File Systems (HDFS), Hive, Spark, Sqoop, Spark SQL to build robust Big Data solution and Programming languages like Scala and SQL.
- Leverage expertise and experience in data migration into the Hadoop ecosystem like Hive.
- Extensive experience in data ingestion technologies using Sqoop to load the data from relational database management into HDFS and vice versa.
- Experience in working with data load from HDFS and Linux file system to have external and managed tables.
- Experience in working with different file formats like TEXTFILE, CSV, JSON, ORC, XML, AVRO
- Worked on improvising the tuning options using Hive partitioning, Bucketing, Query Level Optimizations, Join Optimizations and Window functions
- Experience in query optimization to achieve low latency and faster processing.
- Experience in core Spark and Spark SQL using functional programming in Scala
- Experience in understanding and managing Hadoop Log Files.
- Knowledge in using Oozie for configuring job flows.
- Experience in designing both time driven and data driven automated workflows using Oozie.
- Implemented Spark using Scala and SQL and performed in-memory data processing utilizing data frames and dataset API
- Hands on experience in agile methodologies.
- Good experience in SQL, Microsoft SQL Server and PostgreSQL database
- Knowledge on AWS.
- Experience on GitHub and Bitbucket
- Attended Trainings on Snowflake.

Work Experience

Working as a Senior Software Developer at Excelra Knowledge Solutions Pvt. Ltd. Hyderabad from Feb 2022 - Till date.

Charter Global Technologies Pvt Ltd	:	Sep 2021 - Dec 2021
FactSet Systems India Pvt. Ltd. (DE)	:	Jun 2019 – Aug 2021
FactSet Systems India Pvt. Ltd. (DA)	:	Sep 2018 – May 2019
FactSet Systems India Pvt. Ltd. (RA)	:	Mar 2016 – Aug 2018

Tech Stack

- ➢ Skilled in Big Data Eco-Systems: Hadoop (HDFS), YARN, Hive, Hue, Sqoop, Spark, SparkSQL and Scala
- ➢ Databases & Data Storage: SQL Server, Oracle, PostgreSQL and MySQL
- ➢ Cloud Technologies: AWS – Amazon S3, Lambda, Glue, EC2, EMR and Athena
- ➢ OS (Operating System): Windows, UNIX
- ➢ Programming Languages: Scala, SQL
- ➢ Version Control System: GIT and Bitbucket
- ➢ Big data Platforms: Cloudera Data Platform

Professional experience:

- ➢ **Client Name : GSK, UK**
- ➢ **Project   : Genomics Data Ingestion Tracker**
- ➢ **Duration   : Feb -2022 - Till date**
- ➢ **Role   : Developer**

Roles & Responsibilities:

- ➢ Understanding Project related PPTs, Master data spreadsheets, and Data source spreadsheets in detail.
- ➢ During this project, I spend most of my time working with PostgreSQL database and Excel Spreadsheets.
- ➢ I have associated with businesspeople on daily basis and according to their suggestion I was worked on the database and data source spreadsheet.
- ➢ As per schema changes, I have reconciled whether all the fields are inline or not in the required sources – I.e., Master data spreadsheet, Source data spreadsheet, database and UI API fields - and found any missing fields.
- ➢ Initially, I manually inserted the data from spreadsheets into the Dev database, then I did a sanity check on the Dev UI application, and I prepared a PPT with my findings, then I met with businesspeople to discuss my findings and get their perspectives.
- ➢ I entered data into the test database after the Dev database information was confirmed by businesspeople
- ➢ As per businesspeople requirement I have associated dataset_id's with request_id's.
- ➢ Worked on Agile methodology and participated in daily Scrum Meetings and Demo Meetings.
- ➢ Working knowledge on Git.

- **Client Name** : **ADP & LLC, USA**
- **Project** : **T - Mobile Employees payroll Automation**
- **Duration** : **Sep-2021 To Dec-2021**
- **Role** : **Developer**

Roles & Responsibilities:

- Obtained IDmod's information from a variety of sources
- In accordance with the migration sheet, rebuilt the MySQL queries in PostgreSQL.
- The Data pipe line should be structured according to the USA government's statutory deductions for salaried employees. i.e., Pre-tax deductions, Mandatory deductions, post-tax deductions and Voluntary deductions.
- Responsible for maintaining source code repositories using Bitbucket and to track document and control changes to sourcecode
- Responsible for building the pipelines using datasets, linked services and activities like stored procedure, copy, lookup, For-Each etc.
- Worked on Agile methodology and participated in daily Scrum Meetings and Demo Meetings.
- Working knowledge on Bitbucket

- **Project** : **Merger and acquisition, corporate actions and Financial Data**
- **Duration** : **Jun - 2019 To Aug -2021**
- **Role** : **Associate Developer**

Roles & Responsibilities:

- Involved in the design and development of technical specifications using Hadoop Technology.
- Developed Hive queries to analyze large data sets.
- Create Hive Jobs to Parse the logs and structure them in tabular format to facilitate effective querying of the data.
- Performing unit tests for Hive queries and subsequent development integration test.
- Writing complex transformation logic and translate them into HQL queries
- Written the Apache Hive to process the HDFS data
- Created Hive tables to store the processed results in a tabular format
- Ability to achieve project goals within project constraints
- Involved in the design and development of technical specifications using Hadoop Technology
- Monitoring and Debugging the Hive Jobs

Academic Qualifications

- MBA from RVR &JC College of Engineering, affiliated to Acharya Nagarjuna University in the year 2015

Achievements

- Appreciation from GSK Client for the GDIT application support
- Star performer award for the year 2019.
- Participated in Road Safety & Traffic Management hackathon conducted by SCSC (Society for CyberabadSecurity Council)