

Ashish Kothapalli
Data Engineer
Mail: in.data89@gmail.com
Phone: (971)570-7413

PROFESSIONAL SUMMARY:

- 4+ years of overall IT experience working with storage and data analysis tools making use of cloud technologies.
- Proficient in Azure data services including Azure Data Factory, Azure Databricks, Azure SQL Database, Azure Cosmos DB, Azure Blob Storage, and Azure Synapse Analytics
- Experience working with **Databricks** (notebooks, Delta Lake, pipelines).
- Hands on experience in developing SPARK applications using Spark tools like RDD transformations, Spark core, Spark Streaming and Spark SQL.
- Experience working with python framework (Pandas, NumPy)
- Working knowledge of **Amazon's Elastic Cloud Compute** (EC2) infrastructure for computational tasks and Simple Storage Service (S3) as Storage mechanism.
- Setting up and integrating **Hadoop ecosystem tools** – HBase, Hive, Sqoop etc.
- Hands on experience in writing **Spark SQL scripting**.
- Experience with creating pipelines, SQL querying and Data ingestion on **Databricks**.

TOOLS AND TECHNOLOGIES:

Databases: SQL Server, MySQL, Oracle SQL Server, MongoDB, Cassandra.

Data Visualization Tools: Tableau

Analytics Tools: AWS Glue, AWS Athena, IAM, S3, EC2, GCP BigQuery, Cloud Storage, Databricks

Programming Languages: Python, NoSQL, Scala

Integration Tools: Git

Big Data Ecosystem: HDFS, MapReduce, PySpark, Hive.

Operating System: Linux, Windows

PROFESSIONAL EXPERIENCE:

Client: TCS, Hyderabad
Role: Data Engineer

Feb '19 – April'22

Project: Create data pipeline for ingestion, process, storage and access to meaningful insurance and claims data for various needs of business, from basic reporting to complex analytical problems.

Responsibilities:

- Developed **ETL** data pipelines using **Spark**.
- Responsible to ingest data from file streams and databases. Process the data with SQL databases, PySpark.
- Create **Pyspark** data frames for analysis.
- Worked on RDD's and related functions to work on partitioned data.
- Worked on reading and writing multiple data formats like **JSON, ORC, Parquet** on **HDFS** using **PySpark**.
- To implement and maintain the **ETL** (Extraction, Transformation, used to develop the jobs/Cleansing) Process using **Data stage**.
- Worked on building pipelines on **Databricks** to capture streaming data and load it into data warehouse.
- Worked on NoSQL databases like **Cassandra**.

- Performed data analytics on DataLake using Pyspark on databricks platform.
- Monitor data flow and maintain the data pipeline on **Databricks**.
- Perform Analytics on data extracted from different sources in Databricks.

Environment: Spark, Python, NumPy, Databricks, SQL, Cassandra, Pyspark, Databricks

Alp Consulting, Hyderabad

Oct '18 – Jun ' 20

Role: Jr. Data Analyst

Project :- Data extraction from SQL database and cleaning the extracted data using Python library, Pandas and loading the cleaned data to MySQL database.

Responsibilities:

- Used **Sqoop** to ingest data from relational databases.
- Provided technical assistance and data support for new incidents - both technical capability and business issues.
- Performed basic exploratory data analysis with **Pandas** and analysis using seaborn.
- Performed data extraction using SQL queries to extract useful data.
- Performed statistical calculations using Python **NumPy** library.
- Knowledge of **cloud data storage** (Azure Data Factory, Azure Blob Storage, Synapse Analytics, Azure AD, Azure SQL, Databricks).
- Used Jenkins as a CI/CD tool to deploy.
- Participate in the Agile meetings to discuss the daily progress.
- Intermediate working knowledge on MS Excel for data analysis.
- Provided data analysis to either support or oppose proposed system.
- Responded to queries either in person or over the phone.

Environment: Sqoop, Jenkins, Python, NumPy, MS Excel, SQL, Pandas, Azure cloud technologies
--

Client: IBM, Bangalore

Feb` 17 – Sep' 18

Role: TECHNICAL ANALYST

Responsibilities:

- Engage in technical analysis of software problems using problem determination / problem source identification skills
- Focus on customer satisfaction by managing client relationships and effectively communicating status and action plans.
- Collaborate across teams. Utilize technical and negotiation skills to work with team members in development, operations, and support to jointly diagnose problems.
- Worked on Tableau -Business Intelligence and Analytics tools for connecting to data, creating and managing data sources, Filtering, Actions, and types, formatting data, building Views, Maps, Dashboards publishing data etc.

- Take care of queries from the internal employees as the internal IT and resolve them via calls & emails. Have cross meetings with other teams to discuss and put forth possible solutions for recurring issues.
- Use Service-Now as a CRM ticketing tool to manage, create and retrieve user records.
- Create an agile environment and techniques to ease the ITIL process.
- Monitor incidents and problems raised via the support desk software; assess for service level agreement adherence and intervene with appropriate action.
- Work with support teams onshore and offshore.
- Personally, manage major and significant incidents and act as an escalation point to ensure timely resolution of incidents within expectations and SLAs.
- provide request fulfillment and/or allocate incidents to other members of Technical Run Services via a structured escalation path.
- Take care of escalations directly from the client and give a timely resolution.
- Handle a team of six and manage day to day tasks.
- Assign tasks to the team members and ensure on time resolution.
- Report to the sector lead on the team`s performance and deliver the daily report.

Education:

- Bachelor`s - Sri Bhagwan Mahaveer Jain College – Biotechnology - 2010
- Master`s – St. Aloysius College of Management & Technology – Bioinformatics – 2013
- Master`s – Wright State University, OH – Industrial Engineering - 2016