Deepak Kumar
Contact: +91 -9748217412; Email: deepak75719@gmail.com

CAREER OBJECTIVE

A highly skilled Azure Data Engineer with over 7 years of experience in developing data-intensive applications on cloud platforms. Possesses deep technical expertise in transforming heterogeneous data into actionable insights, effectively meeting the diverse needs and requirements of clients and stakeholders within the organization.

EXPERIENCE SUMMARY

- Possess 7.3 years of extensive experience within the IT industry, contributing to various high-impact projects and solutions.
- Demonstrate exceptional communication and interpersonal skills, with a proven ability to deliver compelling presentations and conduct effective product demonstrations.
- Accumulate over 6 years of specialized experience in Big Data technologies, showcasing expertise in handling vast datasets and implementing cutting-edge solutions.
- Hold more than 5 years of comprehensive experience in Microsoft Azure, with proficiency across a wide array of cloud-based tools and services.
- Proficient in leveraging Microsoft Azure services, including Azure Synapse Analytics, Azure Data Factory, Logic Apps, Azure Databricks, and Microsoft Fabric, to architect and implement advanced data solutions.
- Adept at understanding and thoroughly analyzing client requirements, ensuring the successful translation of business needs into technical solutions.
- Demonstrated ability to manage and oversee multiple concurrent projects, working collaboratively with large, geographically dispersed teams in fast-paced and highly dynamic environments.
- Extensive experience in designing and developing data pipelines using Azure Synapse and Azure Data Factory, alongside writing complex data transformations in Azure Databricks using PySpark for diverse use cases.
- Expert in implementing Spark with Python and Spark SQL, optimizing data processing workflows and accelerating testing cycles for improved performance and efficiency.

Core Competencies: *Azure Synapse~ Azure Data Factory~ Logic App ~ Azure Databricks~ Microsoft Fabric*

TECHNICAL SKILLS

Defect Tracking Tool:	JIRA
Development Tool:	Pycharm
Technology:	Azure Synapse Analytics, Spark, Azure Databricks, Azure Data Factory, Microsoft Fabric, Azure Logic App

EDUCATIONAL & PROFESSIONAL CREDENTIALS

Bachelor of Technology– 2017
Techno India Salt Lake, Kolkata
Intermediate in 2012 (CBSE)
High School in 2010 (CBSE)

PROFESSIONAL EXPERIENCE

ITC Infotech (Since December 2023)
Designation: Senior Data Engineer (Lead Consultant)

Tech Mahindra (October 2021 – December 2023)
Designation: Data Engineer (Senior Software Engineer)

Mindtree Limited (March 2020 – October 2021)
Designation: Data Engineer (Senior Software Engineer)

Tata Consultancy Services Limited (August 2017 – March 2020)
Designation: Systems Engineer

Trainings Attended:

Undergone Databricks and ADF trainings in Mindtree.
Undergone various Azure trainings in Tech Mahindra

Certifications:

Microsoft Azure Fundamentals (AZ-900)

Microsoft Azure Data Fundamentals (DP-900)

Azure Databricks Core Technical Training

Capstone

PROJECT EXPERIENCE

Customer Name: BAT

Duration: Dec 2023 – Dec 2024

Project	BAT Analytics Platform (BAP)
Role	As an ETL Developer, my primary role and responsibility involved modifying and optimizing existing pipelines to enhance data processing efficiency, ensure seamless integration, and meet evolving business requirements.
Tools and Technologies	Azure Data Factory, Microsoft Fabric, Synapse Notebook.
Operating System	Windows

Responsibilities	1. We acquire data from four distinct SAP systems, namely ECC, MDG, GOSS, and TMS. 2. A linked service was established to facilitate connectivity with the SAP source via SLT, utilizing the SAP CDC (Change Data Capture) Connector. 3. An ADF (Azure Data Factory) pipeline, along with generic data flows, was developed to ingest data from SAP (through SLT) into delta tables, successfully onboarding over 950 SAP tables. 4. Shortcuts were created within Microsoft Fabric, pointing directly to the delta tables for seamless access and integration. 5. The ADF pipeline was optimized to enhance the monitoring process, making it more streamlined and efficient by leveraging the annotation feature for better tracking and visibility.

Customer Name: Communisis
Duration: Aug 2022 – Dec 2023

Project	Modern Data Platform (MDP)
Role	As an ETL Developer, my role and responsibility involved architecting and implementing comprehensive end-to-end data ingestion pipelines utilizing Azure Synapse Analytics, ensuring the efficient extraction, transformation, and seamless loading of data for advanced analytics and reporting.
Tools and Technologies	Azure Synapse Analytics, Azure dedicated SQL pool, Azure logic App and Azure Key Vault.
Operating System	Windows
Responsibilities	We receive a multitude of files on a daily basis from seven distinct Towers, each in varying formats such as JSON and CSV. Initially, these files are categorized and organized into their respective directories based on their filenames. Subsequently, we transform these files into a structured format and load them into the RAW layer tables. After populating the RAW layer, we proceed to ingest the data into the STG layer, adhering to an incremental load strategy and implementing the Slowly Changing Dimension (SCD) Type-2 methodology. Additionally, we have developed a robust logging mechanism that captures logs for every event throughout the process. In the event of an error, an automated email notification is triggered via Azure Logic Apps, containing the pertinent details, which is sent to the designated group of users.

Customer Name: Anglo American
Duration: Oct 2021 – Aug 2022

Project	Anglo American
Role	As an ETL Developer, my role and responsibility involved designing and implementing end-to-end data ingestion pipelines, ensuring the seamless extraction, transformation, and loading of data to support robust data processing and analytics workflows.

Tools and Technologies	Spark, Azure Data Factory, Azure Databricks, cosmos DB
Operating System	Windows
Responsibilities	1. Developed a robust ingestion pipeline capable of transferring data from on-premise systems to the cloud, triggered by predefined events within the pipeline. 2. Actively engaged in data modeling, cleansing, ensuring data quality, and preparing input datasets for the development of predictive models. 3. Designed and implemented Azure Data Factory pipelines to orchestrate seamless data integration and transformation processes, facilitating efficient workflows into Azure Databricks.

Customer Name: 7-11
Duration: July 2020 – Oct 2021

Project	7-11
Role	As an ETL Developer, my role and responsibility encompassed the design and development of end-to-end data ingestion pipelines, ensuring the efficient extraction, transformation, and loading of data to facilitate seamless data flow and support analytical processes.
Tools and Technologies	Spark, Azure Data Factory and Azure Databricks
Operating System	Windows
Responsibilities	1. Engineered a dynamic ingestion pipeline to seamlessly transfer data from on-premise systems to the cloud, driven by event-based triggers within the pipeline. 2. Led efforts in data modeling, cleansing, data quality validation, and preparing input datasets essential for the development of sophisticated models. 3. Developed and optimized workflows using Azure Logic Apps to automate and streamline business processes. 4. Designed and implemented data pipelines in Azure Databricks, utilizing PySpark to write transformations for various complex use cases. 5. Constructed and managed Azure Data Factory pipelines to efficiently orchestrate data integration and transformation workflows into Azure Databricks.

Customer Name: Unilever
Duration: June 2020– July 2020

Project	Unilever
Role	As an ETL Developer, my role and responsibility involved designing and building end-to-end data ingestion pipelines, ensuring the smooth and efficient flow of data from source to destination while adhering to best practices in extraction, transformation, and loading (ETL) processes.

Tools and Technologies	Spark, Azure Data Factory and Azure Databricks
Operating System	Windows
Responsibilities	1. Developed Pipeline 2. Did transformation, analysis and unit testing using Azure Databricks.

Customer Name: Aldi-Data Bureau

Duration: Apr 2019 – Mar 2020

Project	Aldi-Data Bureau
Role	As an ETL Developer, my core responsibility was to modify and optimize existing pipelines, driving improvements in data processing efficiency, ensuring seamless integration across systems, and adapting to the dynamic needs of the business.
Tools and Technologies	Spark, Azure Data Factory and Azure Databricks
Operating System	Windows
Responsibilities	1.Implemented retry logic in many jobs to reduce the number of pipeline failure. 2. Worked on making existing pipelines parameterized to make them more generic. 3. Provided on call support in case of pipeline failure.

Customer Name: Tata Sky, India

Duration: Jan 2018– Feb 2019

Project	Tata Sky, India
Role	Worked as one of the Hadoop Developer in the team
Tools and Technologies	Hadoop, HDFS, Hive, Pig, Sqoop
Operating System	Windows
Responsibilities	1. Contributed to the development of Pig scripts to process and analyze large-scale data stored in HDFS. 2. Contributed to the development of Pig scripts to process and analyze large-scale data stored in HDFS. 3. Designed and implemented partitions for Hive tables, enhancing data organization and query performance. 4. Created and managed Hive tables to store processed results in an optimized, tabular format. 5. Authored and executed Pig Latin scripts to load and transform data into Hive tables for further analysis and reporting.