Deepika Dodapanni

Bangalore, India

+91 9876543210

deepika.dodapanni@example.com

LinkedIn: deepikadodapanni

SKILLS

- **Data Engineering:** ETL Pipelines, Data Warehousing, Data Modeling
- **Programming:** Python, SQL, Scala
- **Big Data Technologies:** Apache Spark, Hadoop, Hive, Kafka
- **Databases:** PostgreSQL, MySQL, MongoDB, Snowflake
- **Cloud Platforms:** AWS (S3, Redshift, Lambda), GCP (BigQuery, Dataflow)
- **Workflow Orchestration:** Apache Airflow, Prefect
- **Data Visualization:** Tableau, Power BI
- **Soft Skills:** Problem Solving, Communication, Team Collaboration

EDUCATION

Indian Institute of Technology (IIT), Bombay

Master of Technology in Data Science & Engineering (2018 – 2020)

CGPA: 8.8

ABC High School, Bangalore

Intermediate in Science (2016 – 2018)

Percentage: 89.5%

XYZ School, Bangalore

SSC (2015 – 2016)

Percentage: 93.0%

EXPERIENCE

Data Engineer | Infosys Ltd. | July 2020 – Present

- Designed and implemented scalable ETL pipelines for large-scale data processing.
- Built real-time streaming data solutions using Apache Kafka and Spark Streaming.
- Optimized SQL queries and database performance for data warehouse solutions.
- Automated workflows using Apache Airflow to improve efficiency.
- Worked on cloud data lakes using AWS S3 and Redshift.

Data Engineering Intern | TCS Digital | Jan 2019 – June 2020

- Developed batch processing pipelines using Hadoop and Hive.
- Assisted in building a data ingestion framework for structured and semi-structured data.
- Conducted data cleaning and transformation using Python and SQL.
- Created dashboards for business insights using Power BI.

PROJECTS

Customer Churn Prediction System | Apache Spark, AWS

- Developed a predictive model to identify customers at risk of churn using ML algorithms.
- Processed large-scale customer transaction data using Spark and AWS Redshift.
- Implemented data pipelines to ingest and transform real-time customer data.

Data Lake for E-commerce Analytics | Hadoop, Snowflake

- Designed a scalable data lake architecture for an e-commerce platform.
- Used Hadoop and Snowflake for efficient data storage and analytics.
- Integrated various data sources and performed transformation using Spark.

Real-time Fraud Detection System | Kafka, MongoDB, Airflow

- Built a fraud detection system for financial transactions using real-time data streaming.
- Implemented Apache Kafka for real-time ingestion and MongoDB for NoSQL storage.
- Automated anomaly detection workflows using Apache Airflow.

CERTIFICATIONS & ACHIEVEMENTS

- Certified Data Engineer – Google Cloud
- AWS Certified Solutions Architect – Associate
- Winner of National Data Science Challenge
- Published research paper on Big Data Processing at IEEE Conference

EXTRACURRICULAR ACTIVITIES

- Member of Data Science Club at IIT Bombay
- Volunteer at Data for Good initiative, helping NGOs leverage data insights
- Finalist in National Hackathon on Data Analytics