Pranith Kumar Dongari

pranithkumardongari3@gmail.com | linkedin.com/in/pranithkumar | +1 (561) 757-8003 | Boca Raton, Florida, USA

Data Engineer with hands-on experience designing and building scalable ETL pipelines using Azure Data Factory, Databricks, and PySpark. Proficient in SQL, Python, and Big Data technologies, with a strong focus on data quality, performance optimization, and workflow automation. Skilled in transforming complex datasets into actionable insights to drive data-informed business decisions.

EDUCATION

Florida Atlantic University **10/2023 – 05/2025**
Candidate for Master of Science in Information Technology Management
GPA: 4.0/4.0 | Dean's List

TECHNICAL SKILLS

Big Data & ETL Tools: Databricks, Apache Spark (RDD, DataFrame, Spark SQL), Apache Hadoop (HDFS, Spark), Snowflake
Cloud Platforms & Solutions: Azure Data Factory, Azure Synapse, Databricks
Programming & Scripting: Python (Pandas, NumPy, PySpark), SQL (Advanced)
Data Modeling & Warehousing: Dimensional Modeling, Star Schema, Snowflake Schema,
CI/CD & Version Control: Git, GitHub, GitHub Actions (Basic), Azure DevOps (Basic), SDLC Understanding
Databases: MySQL**,** SQL Server
Data Visualization: Power BI (Basic), Looker (Familiar), MS Excel (Advanced)
Web Development (Basic): HTML, CSS, JavaScript

EXPERIENCE

9thPixel Technologies India Pvt.Ltd **03/2021- 06/2023**
Data Engineer *Hyderabad, India*

- Designed and implemented scalable ETL pipelines in Azure Data Factory (ADF), automating data integration across multiple sources and improving data consistency, reducing manual data handling by 20%.
- Optimized ADF activities (copy, lookup, metadata) and integrated stored procedures, boosting pipeline execution efficiency by 30% and enhancing data flow reliability.
- Developed reusable, parameterized ADF pipelines for scalable workflows, cutting future development time by 40%.
- Built data transformations in PySpark on Azure Databricks, streamlining the processing of large datasets and improving the speed of business intelligence reporting.
- Collaborated with stakeholders to deliver real-time analytics solutions, driving a 5% increase in revenue by supporting data-driven decision-making.

PROJECTS

Lending Club Loan Data Pipeline| SQL, PySpark, Databricks Community Edition | Personal Project **Git Hub**
Duration: 02/2025 – 04/2025

- Implemented a PySpark-based data cleaning and transformation module within the Databricks ETL pipeline, improving data quality scores by 40% and enabling more accurate risk assessments.
- Utilized PySpark for data cleaning, transformation, and optimization, reducing processing time by 30%.
- Conducted automated data processing workflows, ensuring efficient handling of large-scale datasets.
- Established a standardized data validation process, reducing data-related errors by 15% and improving overall data integrity.

CERTIFICATIONS

MS Excel - Advanced, Cengage **04/2024**
Big Data Analytics, Florida Atlantic University **05/2025**