

Debanjan Sarkar – About me

I'm a skilled Data Engineer with a wealth of experience and a strong educational background in Electronics and Communication Engineering. Currently serving as a Junior Associate in Big Data and EDW at Celebal Technologies, I have showcased exceptional proficiency in a multitude of technologies and tools.

In my current role, I have been at the forefront of implementing robust data solutions. Utilizing Azure Data Factory, I adeptly applied Data Flows transformations, focusing on hash column application, deduplication of client's product sales raw data, and the conversion of file formats.

Notably, my expertise lies in the usage of PySpark on Databricks, where I processed incoming client data using the medallion architecture. My role involved cleaning and validating data, and building Statistical Tables over silver data. I also played a pivotal role in designing Data Models for Data Warehousing, demonstrating my ability to normalize tables, identify Dimensions and facts, and implement SCD Type 1 and Type 2 on related dimensions. My contribution extends to implementing logic for lineage table population, tracking the metrics of ETL full load and incremental load, showcasing my holistic understanding of the data engineering landscape.

My educational journey at Jalpaiguri Government Engineering College culminated in a B. Tech degree in Electronics and Communication Engineering, where I achieved an impressive 8.84 DGPA. This academic excellence provides a solid foundation for my practical skills in the field.

Certified in AZ900, DP900 and Databricks Lakehouse Fundamentals V2, I showcased a commitment to staying abreast of industry trends. Beyond certifications, I was engaged in real-world projects, including the development of a full-stack web application, a CRUD API using FastAPI, and a web application for Book Record Management using Django, all deployed on various platforms such as railway.app and HEROKU.

So, as some of them are already mentioned, my technical skills encompass Databricks, Apache Spark (PySpark, Spark SQL), SQL Server, Azure Data Factory, LAMP stack, Python, Django, FastAPI, BeautifulSoup, and proficiency in version control with Git & GitHub. Additionally, Debanjan I competency in C/C++ and JAVA programming languages.

Beyond my technical prowess, I bring a keen problem-solving acumen to the table. My ability to create versatile and scalable data-driven applications underscores my commitment to delivering solutions that align with business objectives.

In conclusion, I am a dynamic and results-driven Data Engineer, equipped with a potent combination of technical skills, practical experience, and a strong academic foundation. My contributions, coupled with a passion for continuous learning and staying at the forefront of technology, places me as a standout professional in the field of data engineering. As I seek new opportunities in the fields of Data Science and Generative AI, I am poised to make substantial contribution to any domain that values innovation, efficiency, and excellence in data-driven decision-making.