Job Title   : Software Engineer (Databricks)
Skill   : Databricks, Python, Pyspark
Location   : Navi Mumbai
Type of Demand   : Full Time Employee
Work Details   : Office
Minimum Experience : 2+Years

Job Summary

We are seeking an experienced (2+Years) Software Engineer to join our team and help build and enhance our data engineering, analytics, and machine learning capabilities within the Databricks, Lakehouse Platform. In this role, you'll work with cutting-edge technologies and contribute to products that help organizations solve their most challenging data problems.

Key Responsibilities:

➢ Design, develop, and maintain high-performance Python applications that integrate with Apache Spark, Delta Lake, and other components of the Databricks platform
➢ Optimize and scale Python code for distributed computing environments
➢ Implement data processing workflows, ETL pipelines, and machine learning models
➢ Collaborate with cross-functional teams to integrate Python libraries and frameworks with the Databricks ecosystem
➢ Write clean, maintainable, and well-tested code following best practices
➢ Debug complex issues in distributed computing environments
➢ Provide technical guidance and mentorship to junior engineers

Qualifications:

➢ Bachelor's degree in Computer Science, Engineering, or related field (or equivalent practical experience)
➢ 2+ years of professional software development experience with Python/Databricks
➢ Strong knowledge of Python programming patterns, best practices, and the Python ecosystem
➢ Experience with data processing frameworks like Apache Spark, Pandas, or similar technologies
➢ Familiarity with SQL and database concepts
➢ Understanding of distributed computing principles
➢ Experience with version control systems (Git) and CI/CD pipelines
➢ Strong problem-solving skills and attention to detail

Preferred Skills:

➢ Experience with Databricks platform or Apache Spark
➢ Knowledge of Delta Lake, MLflow, or other Databricks open-source projects
➢ Experience with cloud platforms (AWS, Azure, GCP)
➢ Background in data engineering, analytics, or machine learning applications
➢ Contributions to open-source Python libraries
➢ Experience with performance tuning and optimization for big data applications
➢ Knowledge of software development methodologies (Agile, Scrum)