Senior/Lead Data Software Engineer

Senior/ Lead Data Software Engineer Description

EPAM is a leading global provider of digital platform engineering and development services. We are committed to having a positive impact on our customers, our employees, and our communities. We embrace a dynamic and inclusive culture. Here you will collaborate with multi-national teams, contribute to a myriad of innovative projects that deliver the most creative and cutting-edge solutions, and have an opportunity to continuously learn and grow. No matter where you are located, you will join a dedicated, creative, and diverse community that will help you discover your fullest potential.

No less important is the safety, well-being, and experience of our applicants. Therefore until further notice, all EPAM employment interviews will be conducted remotely. Our recruitment professionals and hiring managers are standing by to ensure a robust and engaging virtual candidate experience. We look forward to speaking with you!

Responsibilities

- Design, construct, install, test and maintain highly scalable and optimized data pipelines with state-of-the-art monitoring and logging practices
- Bring together large, complex, and sparse data sets to meet functional and non-functional business requirements and use a variety of languages, tools and frameworks to marry data
- Design and implement data tools for analytics and data scientist team members to help them in building, optimizing, and tuning of use cases
- Tackle challenging and varied problems related to gaining further maturity in Data Platform tools, processes, and engineering capabilities
- Focus on engineering excellence to deliver high-quality solutions that provide leverage for the company's objectives

Requirements

- 3+ years of hands-on experience in data processing focused projects
- Proficiency with Java or Python or Scala and SQL
- Knowledge of Apache Spark
- Experience with one of the major cloud providers AWS, Azure or GCP
- Hands-on experience with a few selected data processing technologies, e.g., Hadoop, MongoDB, Cassandra, Kafka, Elasticsearch, Python libraries (Pandas/NumPy/…) data processing tools of cloud providers (EMR/Glue/Data Factory/Big Table/…)
- Relevant experience with version control and code review
- Knowledge of Agile methodologies
- Linux and Bash scripting basics

Nice to Have

- Good hands-on experience with Databricks and Delta-Lake
- Ability to build Apache Airflow pipelines
- Experience with the Snowflake platform

We offer

- Dynamic, entrepreneurial, high speed, high growth corporate environment
- Diverse multicultural, multi-functional, and multilingual work environment
- Opportunities for personal and career growth in a progressive industry
- Global scope, international projects
- Widespread training and development opportunities
- Unlimited access to LinkedIn learning solutions
- Competitive salary and various benefits
- Sport and social teams support, recreation area, advanced CSR programs