



Data Engineer | Dehradun

About TorcAI

TorcAI, a data science company and global provider of Audience Infrastructure and Programmatic solutions to advertisers, publishers, agencies and innovative technology companies. We provide highly customizable audience platforms, real-time trading systems with machine learning capabilities. TorcAI has provided technology solutions to enterprises in the AdTech and MarTech space enabling them to monetize data assets and bring in competitive technological advantage.

Our platform's strength lies in our ability to re-engineer complex ad and audience tech work flows and bring in efficiencies with an objective to meet business goals. Know more at www.torcai.com.

About Profile:

As a Data Engineer, you will be responsible for designing, implementing, and optimizing data pipelines and systems that enable the organization to make data-driven decisions. You will collaborate with cross-functional teams to ensure data quality, accessibility, and security. Additionally, you will support the development of KPI measurements and deliver actionable insights for the organization.

Roles and Responsibilities:

- Design, develop, and maintain scalable, reliable, and efficient data pipelines to process large volumes of structured and unstructured data using Python and PySpark.
- Collaborate with data scientists, analysts, and other stakeholders to understand data requirements and transform them into technical solutions.
- Ensure data quality, consistency, and accuracy by implementing data validation, cleansing, and monitoring processes.

tomorrow's tech, today.



- Optimize and enhance existing data infrastructure to improve performance, reliability, and scalability.
- Develop and maintain data ingestion processes from various sources, such as APIs, databases, and third-party services.
- Implement data warehousing solutions and maintain data storage systems, ensuring data security and compliance with relevant regulations.
- Collaborate with the data team to build and maintain data models, metadata, and data dictionaries to support reporting and analytics.
- Contribute to the evaluation and selection of data-related tools and technologies.
- Assist in the development and tracking of KPI measurements, providing actionable insights to drive business decisions and improvements.
- Provide technical guidance and support to team members and stakeholders, while continuously learning and staying up-to-date with industry trends.

Skill (Tech & Functional)

1. Strong programming skills in Python are a must-have.
2. Hands-on experience with PySpark and processing large datasets in a streaming environment.
3. Proficient in SQL and experience with relational databases (e.g., MySQL, PostgreSQL, Oracle) and NoSQL databases (e.g., MongoDB, Cassandra, DynamoDB).
4. Experience with data warehousing solutions and technologies, such as Snowflake, Redshift, or BigQuery.
5. Familiarity with data integration and ETL tools, such as Apache NiFi, Talend, or Informatica.
6. Strong understanding of data modeling, schema design, and data management principles.
7. Excellent problem-solving and analytical skills, with the ability to work independently and as part of a team.
8. Strong communication skills, with the ability to articulate complex concepts to both technical and non-technical stakeholders.

tomorrow's tech, today.



Nice to have:

1. Experience with Apache Druid.
2. Experience with cloud platforms such as AWS, GCP, or Azure.
3. Knowledge of dashboard APIs and data visualization tools, such as Tableau, Power BI, or Looker.
4. Familiarity with machine learning concepts and tools.

Qualification & Experience

- Bachelor's degree in Computer Science, Engineering, or a related field.
- 2-4 years of experience working as a Data Engineer or in a similar role.

Please mail your resume at: hire@torcai.com

tomorrow's tech, today.