Module 1 Summary: Data Engineering and Generative AI

Congratulations! You have completed this module. At this point in the course, you know:

- Data engineers can integrate AI tools into their workflows and streamline their tasks with:
  - Synthetic data generation
  - Data augmentation
  - Data anonymization
  - Data cleaning and transformation
  - Data exploration and visualization
- Generative models focus on the distribution of data to generate new instances, while discriminative models separate data points into classes.
- GANs comprise two competing neural networks.
- VAEs are models that learn to encode data into a compressed representation and then decode it .
- Transformers' self-attention mechanisms allow for understanding and generating complex data patterns, revolutionizing creative and generative tasks across various domains.
- The core responsibilities of data engineers include maintenance of data infrastructure, development of ETL processes, and compliance with data quality and governance.
- DataRobot automates data cleaning and preparation, and GPT-3 generates predictive text and simulates scenarios.
- Transformative generative AI solutions include dynamic data storage and processing, advanced data integration, and efficient data management.
- Generative AI approaches schema design by learning from vast data patterns and existing database structures.

Mark as completed