

✔ Congratulations! You passed!

Grade
received 100%

Latest Submission
Grade 100%

To pass 80% or
higher

Go to next item

1. What is an API-based application in the context of Generative AI?

1 / 1 point

- ☐ An application that retrieves relevant information from a database or external source to generate responses
- ☒ An application that communicates with a large language model over a network through HTTP requests
- ☐ An application that uses multiple specialized models for different tasks
- ☐ An application that integrates a large language model directly into its codebase

✔ Correct

Correct, this option describes an API-based application. It interacts with a large language model using HTTP requests to send information and receive responses. The model does not reside within the application itself.

2. What is an embedded model in Generative AI?

1 / 1 point

- ☐ An API-based application that communicates with a large language model over a network through HTTP requests
- ☒ A large language model integrated directly into an application's codebase
- ☐ A single, specialized model used for different tasks in a multi-model application
- ☐ A technique where an application retrieves relevant information from a database or external source to generate responses

✔ Correct

Correct, this option describes an embedded model. It integrates the large language model directly into the application's codebase for better performance but also increases the application's size and complexity.

3. What is a multi-model application in Generative AI?

1 / 1 point

- ☐ A specialized model used for different tasks
- ☐ A technique where an application retrieves relevant information from a database or external source to generate responses
- ☒ An application that uses multiple specialized models for different tasks
- ☐ An API-based application that communicates with a large language model over a network through HTTP requests

✔ Correct

Correct, this option describes a multi-model application. It uses multiple specialized models to handle different types of requests, improving efficiency and accuracy.

4. What is Retrieval Augmented Generation (RAG) in Generative AI?

1 / 1 point

- ☐ An API-based application that communicates with a large language model over a network through HTTP requests
- ☐ A single, specialized model used for different tasks in a multi-model application
- ☒ A technique where an application retrieves relevant information from a database or external source to generate responses
- ☐ An embedded model that integrates the large language model directly into the application's codebase

✔ Correct

Correct, this option describes Retrieval Augmented Generation (RAG). It is a technique where an application retrieves relevant information from a database or external source to generate responses.

5. What is one primary challenge in building Generative AI applications?

1 / 1 point

- ☒ Ensuring privacy and security of user data
- ☐ Asynchronous operations to and from the models
- ☐ Maintaining accuracy and format of generated content
- ☐ Context limitation

✔ Correct

Correct, this option describes one challenge in building Generative AI applications. Privacy and security are crucial when handling sensitive user information.

6. What is one benefit of using an API-based application in Generative AI?

1 / 1 point

- ☐ Improved accuracy in generated content through direct access to the large language model's internal mechanisms
- ☐ Improved performance due to direct integration with the large language model
- ☐ Simplified deployment due to integration with the application codebase
- ☒ Better separation of concerns between the application and the large language model

✔ Correct

Correct, this option describes one benefit of using an API-based application in Generative AI. It enables better separation of concerns between the application and the large language model, allowing for easier maintenance and updates to either component.

7. What are some potential drawbacks of using an embedded model in Generative AI?

1 / 1 point

- ☐ Simplified deployment due to integration with the application codebase
- ☒ Increased application size and complexity
- ☐ Better separation of concerns between the application and the large language model
- ☐ Reduced performance due to direct integration with the large language model

✔ Correct

Correct, this option describes one potential drawback of using an embedded model. Integrating the large language model directly into the application's codebase can increase its size and complexity.

8. What are some advantages of using Retrieval Augmented Generation (RAG) in Generative AI?

1 / 1 point

- ☐ Simplified deployment due to integration with the application codebase
- ☒ Improved accuracy and relevance of generated content through external information retrieval
- ☐ Better separation of concerns between the application and the large language model

☐ Improved performance through direct access to large language models' internal mechanisms

☒ **Correct**

Correct, this option describes one advantage of RAG. By retrieving relevant information from a database or external source, the application can generate more accurate and relevant responses.

9. What is one important benefit of using GitHub Actions for deployment?

1 / 1 point

☐ Reduce resource constraints and deployment requirements

☐ Increased speed for deployment

☐ Better management of data security requirements

☒ Multiple steps can be reproduced with accuracy reducing errors

☒ **Correct**

Correct. By automating the many steps required for deployment, you are ensuring reduced errors and high accuracy.

10. What is one use of transformer models in generative AI systems?

1 / 1 point

☐ Structuring datasets

☐ Processing sequential data

☐ Labeling data

☒ Generating language or images

☒ **Correct**

Correct! Transformer models are key to generating new language or image content in generative AI systems.