Contents

[The Magic Behind Generative AI 2](#_Toc177378505)

[Quiz #3 – Gen AI 2](#_Toc177378506)

The Magic Behind Generative AI

Quiz #3 – Gen AI

|  |
| --- |
| **Question #1** |
| What is the main **building block** for processing and storing the knowledge of a Gen AI model? | |
| 1. Deep learning 2. GPUs and Tokens 3. Artificial Neural Networks 4. Supervised learning 5. Unsupervised learning | |

|  |
| --- |
| **Question #2** |
| What are the three main layers of an Artificial Neural Network? | |
| 1. Input, Output, and Processing 2. Data, Analysis, and Results 3. Input, Hidden, and Output 4. Start, Middle, and End | |

|  |
| --- |
| **Question #3** |
| What are the **key advantages** of the Transformer architecture over Recurrent Neural Networks (RNNs) in processing input data? | |
| 1. Sequential processing 2. Parallel processing 3. Attention mechanism 4. Unlimited context window 5. Answers b and c | |

|  |
| --- |
| **Question #4** |
| What is the purpose of using the **attention mechanism** in a Transformer architecture? | |
| 1. To speed up the training process. 2. Pay more attention to specific elements as part of the data stream. 3. To generate new data. 4. To classify data into different categories. | |

|  |
| --- |
| **Question #5** |
| What is a foundation model? | |
| 1. Small-scale model specialized in a specific domain. 2. Large-scale generic model for a wide range of tasks. 3. Small-scale model trained on massive amount of data. 4. Small-scale model trained on small amount of data. | |

|  |
| --- |
| **Question #6** |
| A large language model (LLM) is an example of a foundation model? | |
| 1. Yes 2. No | |

|  |
| --- |
| **Question #7** |
| What is the significance of large language models (LLMs) in the field of Generative AI? | |
| 1. Fast image generation. 2. Fast image and video generation. 3. Natural language processing. 4. None of the answers above. | |

|  |
| --- |
| **Question #8** |
| What is the primary factor used to estimate the power of an LLM model? | |
| 1. Model size – number of parameters. 2. Training data quantity. 3. Training duration and data quantity. 4. Model architecture. | |

|  |
| --- |
| **Question #9** |
| Which factors will be influenced by the **complexity** of training large scale model (LLM)? | |
| 1. Large datasets to train the model effectively. 2. Robust and scalable computing infrastructure. 3. Extended training time. 4. More specialized skills and expertise in machine learning. 5. All answers are correct. | |

|  |
| --- |
| **Question #10** |
| What is the primary function of a **prompt** in relation to an LLM? | |
| 1. That’s the output response of the model. 2. Control the output based on text input. 3. Store image data inside the LLM. 4. Divide the input into tokens. | |

|  |
| --- |
| **Question #11** |
| Why is the input prompt converted into tokens? | |
| 1. Improve the model's creativity. 2. Make the data more relevant and fix mistakes. 3. Simplify processing and storage by using a numerical format. 4. Reduce the number of unique words. 5. Encrypt the data for security. | |

|  |
| --- |
| **Question #12** |
| How are tokens used to measure usage in generative AI systems? | |
| 1. By calculating the number of words in the input and output. 2. By measuring the processing time required for each query. 3. By counting the total number of tokens in the input and output. 4. By monitoring the system's resource usage. | |

|  |
| --- |
| **Question #13** |
| What is the **context window** of a generative AI model? | |
| 1. The maximum number of tokens the model can process at once under the same context. 2. The minimum number of tokens required for a prompt. 3. The number of tokens generated by the model. 4. The maximum number of characters the model can handle. 5. The number of words in the model's vocabulary. | |

|  |
| --- |
| **Question #14** |
| Why are tokens an important issue to consider for applications APIs design and integration with generative AI models? | |
| 1. To improve the model's accuracy. 2. To reduce the cost of training the model. 3. To optimize APIs performance, usage and costs. 4. None of the above is correct. | |

|  |
| --- |
| **Question #15** |
| How does an LLM generate a complete text? | |
| 1. By understanding the meaning of words and sentences. 2. By using a sequential prediction process of tokens. 3. By following a predefined template. 4. By copying a text from a training data. 5. By hiring AI-powered librarian robots. | |

|  |
| --- |
| **Question #16** |
| What is the basis for the LLM's prediction of the next token in a sequence? | |
| 1. Predetermined rules. 2. Random selection. 3. Statistical analysis using probability distribution. 4. None of the answers are correct. | |

|  |
| --- |
| **Question #17** |
| Why does an LLM sometimes generate slightly different outputs for the same prompt? | |
| 1. Due to a model design error. 2. Improve accuracy and make it more deterministic. 3. Simulate creative thinking. 4. Reduce processing time. 5. Ensure consistency with the generated outputs. | |

|  |
| --- |
| **Question #18** |
| How does self-supervised learning work in the context of LLMs? | |
| 1. By applying supervised learning with training dataset of labeled data. 2. By using pre-defined rules to guide the learning process. 3. By training the model to predict missing parts of the input data. 4. By combining supervised and unsupervised learning techniques. | |

|  |
| --- |
| **Question #19** |
| What is the most common and simple practice to influence the output of a model? | |
| 1. Contextual prompting. 2. Fine-Tuning. 3. Retrieval Augmented Generation (RAG). 4. Training a foundation model from scratch . | |

|  |
| --- |
| **Question #20** |
| Is the primary goal of **contextual prompting** in prompt engineering to guide the LLM towards generating a specific response? | |
| 1. Yes 2. No | |

|  |
| --- |
| **Question #21** |
| What is the primary goal of **Retrieval Augmented Generation** (RAG)? | |
| 1. Replace pre-trained LLMs with external databases. 2. Limit the amount of data LLMs can process. 3. Enhance LLMs with external information during the query process. 4. Make LLMs more expensive to use. 5. Reduce the accuracy of LLM responses. | |

|  |
| --- |
| **Question #22** |
| How does **Retrieval Augmented Generation** address the issue of **private data** in public Gen AI systems? | |
| 1. By exposing private data to the public. 2. By training LLMs on private data. 3. By using public data instead of private data. 4. By incorporating private data into the LLM's training. 5. By leveraging private data as additional information inside the input prompt. | |

|  |
| --- |
| **Question #23** |
| What are the potential drawbacks of RAG? | |
| 1. Increased accuracy and relevance. 2. Limited context window and latency. 3. Lower cost and faster response times. 4. Improved domain-specific knowledge. 5. Enhanced privacy and security. | |

|  |
| --- |
| **Question #27** |
| What are the benefits of using fine-tuning compared to prompt engineering with context or RAG? | |
| 1. Increased latency and larger prompt size. 2. Less latency and smaller prompt size. 3. Reduced accuracy and relevance. 4. More complex implementation. 5. Less flexibility. | |

|  |
| --- |
| **Question #28** |
| Can fine-tuning be applied to both open-source and proprietary LLMs? | |
| 1. Only to open-source LLMs. 2. Only to proprietary LLMs. 3. To neither open source nor proprietary LLMs. 4. To both open-source and proprietary LLMs. | |