ASTQB Certified Tester
Testing with Generative AI v1.0
Sample Exam Questions

American Software Testing Qualifications Board

46 Points Possible : Passing Grade = 30 Points
GenAI – Sample Exam

1. (1 pt) Which of the following uses deep learning techniques to create new content?

a. Symbolic AI
b. Classical machine learning
c. Deep learning
d. Generative AI

2. (1 pt) You have used the same prompt multiple times, but you have received different responses for the LLM. Why is this happening?

a. Because the prompt is incorrect
b. Because LLMs have non-deterministic behavior
c. Because the training data was faulty
d. Because LLMs are trained to always vary the response

3. (1 pt) How is tokenization related to embeddings?

a. Tokens are a numeric representation of how a given word is embedded into the model
b. Tokens represent characters, subwords, or words, and embeddings capture the semantic, syntactic, and contextual relationship of the tokens
c. Embeddings describe the text component in such a way that it can be tokenized for use within the model
d. Embeddings identify the non-deterministic behavior of a word component, which is then used to determine the positioning of its token using a vector to represent its position

4. (1 pt) You need an LLM that will help with decision-making based on a process similar to human logic. Which type of LLM should you use?

a. Foundation LLM
b. Instruction-tuned LLM
c. Reasoning LLM
d. Ad-hoc LLM

5. (1 pt) Which of the following is an advantage of using an LLM augmented with vision-language models for developing test cases?

a. There is no advantage because this type of model is used primarily for military applications
b. It can create test cases that can include text and visual cues
c. It can edit images much faster and more accurately than a human
d. It can generate executable code from wireframes and screenshots

6. (1 pt) You have created test automation that will be used to test driver's license applications. There is a huge amount of data available for this testing, but it will take too long to test with all of that data. How can an LLM help with this problem?

a. It can create new synthetic data that will protect private information
b. It can take the existing data and augment that with data that is missing according to the data rules
c. It can take the existing data and break it into smaller sets that will contain interesting variances
d. It can delete data that is duplicated and replace it with new data

7. (1 pt) Your organization is developing a test management tool that will use AI to generate, classify, and store test cases based on requirements, general testing knowledge, and previous issues found. Which of the following interaction models would be best to use?

a. Chatbots talking to the LLM
b. Chatbots talking to other chatbots that have been trained on your system
c. APIs of the LLM to allow the test task information to pass into and out of the LLM
d. APIs of the LLM to pass the test case information into the LLM to be stored there

8. (1 pt) Which of the following are the six components of an effective prompt?

a. Prompt Chain, Persona, Instruction, Assumptions, Constraints, Output Format
b. Role, Bias, Context, Instruction, Input Data, Constraints, Output Data
c. Persona, Background, Instruction, Input Data, Output Data, Assumptions
d. Role, Context, Instruction, Input Data, Constraints, Output Format

9. (1 pt) You have provided a structured prompt to a GenAI model regarding the creation of test cases. You have specified that you want to receive test cases in the following format:

- **Description of test**
- **Reference to source**
- **Steps to complete**
- **Expected outcome**

This is an example of what component of a structured prompt?

a. Role
b. Context
c. Instruction
d. Output format

10. (1 pt) If you have a test process with complex tasks and you will need to check intermediate results before proceeding, which prompting technique is the most appropriate?

a. Prompt chaining
b. Few-shot prompting
c. Zero-shot prompting
d. Meta prompting

11. (1 pt) Which of the following is true regarding a user prompt for an LLM?

a. It is set once at the start of the conversation
b. It is usually set by the developer
c. It usually changes with each interaction
d. It is user-independent, so the response will be consistent regardless of how the user words the prompt

12. (2 pts) You are working on a replacement of an e-commerce system. You have been given a large set of test conditions to test. There is not adequate time to test all the conditions. You have a testing-trained LLM that you can use. How can you best use it in this situation?

a. The LLM can create an estimate for how long testing will take based on similar test conditions. This schedule can then be incorporated into the overall project schedule to allow adequate time for testing.
b. The LLM can be fed information regarding previous defects, end-user requirements, and payment regulations to let it determine the relative risk of all the test conditions. This can then be used to prioritize the testing.
c. The LLM can generate and execute all the test cases based on the test conditions, which will save significant time.
d. The LLM can generate test data to cover every possible purchasing scenario. This will ensure testing will not miss any corner cases, and that will help the schedule by eliminating any testing escapes.

13. (2 pts) You have been given requirements in the form of wireframes of the GUI and user stories. You want to use these to create the test conditions for a set of tests. How should you use prompts to evaluate the inputs and create the test conditions?

a. Apply multimodal prompts
b. Set well-defined system parameters to define the graphics to be used
c. Apply prompt chaining
d. Use meta prompting so the prompts can be continuously refined

14. (2 pts) You have been testing versions of the same software over several years. Rather than the quality improving, it seems to just stay the same, with the same high number of defects being caught during testing for each release. How could AI help with this problem?

a. It could analyze test failures and prioritize the tests that failed last time so they are executed first for each new release.
b. It could use meta prompting to review each test case to determine if it is an effective test. Ineffective tests could then be removed from the overall test suite.
c. It could create large sets of test data that are used across releases of the product to ensure that data variance is not causing the repeated failures.
d. It could conduct root cause analysis over all the defects found across releases to determine the common causes. This information could then be used to improve development and analysis processes.

15. (2 pts) You are working on a critical project that is under the scrutiny of senior management. As a result, simple and clear reporting of testing progress is paramount. You are using the test reporting capabilities of your test management system, but any analysis of the issues and trends has to be done manually by you. How could AI help with this situation?

a. AI could analyze all the open defects and verify that the prioritization is correct based on keywords in the defect report.
b. AI could track which test cases are finding the most defects, so you could highlight those on your reports.
c. AI could analyze the defect and test case execution information and provide trending information and root cause analysis mapped to the use cases for the software.
d. AI would automatically email the existing dashboards to the pertinent executives to ensure they always have the latest information available. For example, defect reports would go to the development management, defect reports to the test management, etc.

16. (2 pts) You are working on testing a new system. The requirements are high-level and lack detail. What prompting technique should you use to get started to get some ideas for how to progress with test case generation?

a. Prompt chaining
b. Few-shot prompting
c. Meta prompting
d. Zero-shot prompting

17. (1 pt) You are using GenAI to create test cases for you for a heavily rules-based application. You have requested that it use decision tables to build the test cases. Which metric would be the best one to use to determine how well it has done with this request?

a. Accuracy
b. Recall
c. Diversity
d. Time Efficiency

18. (1 pt) You have been iteratively modifying your prompt for GenAI test case creation. You are still getting the same types of issues with the output, even though you keep refining the prompts and adding more information. What technique can you employ to use those issues to better evaluate and refine the prompts?

a. Iterative prompt modification
b. A/B testing of prompts
c. Output analysis
d. Adjusting prompt specificity

19. (1 pt) Which of the following is a result of a reasoning error?

a. An output that favors certain types of information, approaches, or assumptions
b. An output that is based on misinterpretations of logical structures
c. An output that appears plausible but is factually incorrect or irrelevant
d. An output that is only partially complete

20. (2 pts) You are using GenAI to develop a test plan. You have fed it the requirements, high-level project plan, schedules, and tester profiles. Despite your instructions, when you ask it to generate the test schedule, it creates a new team member and puts them on the schedule to make the testing effort fit the project timeline. What is the problem, and what should you do about it?

a. This is likely a bias problem. You should give it a wider variety of requirements for similar systems so it generates a larger scope of tests.
b. This is likely a hallucination. You should instruct the AI model that this is incorrect so it learns.
c. This is likely a reasoning error. You should run the test cases against the test objects to verify if the expected results are valid.
d. This is likely a temperature problem. You should lower the temperature to allow a more specific focus.

21. (1 pt) In what way does comparing results across several different models help reduce hallucinations, reasoning errors, and biases?

a. You can combine the results into a superset of data, which will reduce the errors in the data
b. You can select the data that occurs with the highest frequency and know that it is the most accurate
c. You can compare different outputs from the same prompts to detect output errors and reliability
d. You can get the results faster by using multiple models concurrently

22. (1 pt) How does setting a random seed value help improve reproducibility in LLM outputs?

a. By setting a lower seed value, the temperature setting is automatically increased, which narrows the probability distribution
b. By setting a higher seed value, more seeds are allowed to be used, which will promote larger output sets from the LLM, which can then be used to extract consistent data, which improves reproducibility
c. By setting a variable seed value, the random sequence values are increased, which results in a larger set of data, which is easier to recreate
d. By setting a seed value, the random number generator has the same starting point, which will ensure the same pseudo-random sequence is used, which improves reproducibility

23. (1 pt) If you have a GenAI tool that is processing sensitive customer data, what must you ensure?

a. It cannot store the data in an unencrypted form because it will be readable by anyone
b. It cannot store the data where it might be accessed by an unauthorized user, including another GenAI tool
c. It cannot modify and then restore the data because it will be incorrect and might be misleading for other users
d. It cannot access any sensitive data for any use

24. (1 pt) You've been working on an application that processes employee leave requests. The data you are using contains the birth date of the employee, but you have replaced all the birth dates with a default date of 1/1/89. You are using GenAI within your company's databases to build new test data. You have now discovered that the test data contains birthdates again, and they are the real dates. This is an example of what type of data privacy issue?

a. Lack of control over data usage
b. Vulnerability to security attacks
c. Evidence of a malicious introduction of false data
d. Unintentional data exposure

25. (1 pt) What type of attack vector is being used when the results of AI-generated data are rated incorrectly, intentionally?

a. Data exfiltration
b. Request manipulation
c. Data poisoning
d. Malicious code generation

26. (1 pt) What is the purpose of data anonymization?

a. To only use and process the data that is legally permitted and use only the smallest amount needed
b. To mask or replace any private information with non-identifiable data
c. To implement strong encryption and strictly control data access
d. To ensure everyone is trained in the proper handling of data and is minimizing access

27. (1 pt) Which of the following GenAI activities takes about as much power as charging a smartphone?

a. Generating a letter to your bank
b. Creating “business” penguin pictures for a presentation
c. Creating only one penguin image
d. Generating a financial year summary report

28. (1 pt) Which of the following establishes a legal framework addressing AI risks and classifying applications by risk level?

a. ISO/IEC 42001:2023
b. ISO/IEC 23053:2022
c. EU AI Act
d. NIST AI Risk Management Framework

29. (1 pt) Which part of the architecture in an LLM-powered test infrastructure is responsible for integrating multiple data sources?

a. The front-end
b. The LLM itself
c. The back-end
d. The data controller

30. (1 pt) What are the two steps in RAG processing?

a. Input / Retrieval
b. Processing / Output
c. Access / Processing
d. Retrieval / Generation

31. (1 pt) What does "orchestration" mean regarding LLM-powered agents?

a. It is when several agents work together in a coordinated manner to more efficiently solve complex problems
b. It is when more than three LLM-powered agents work independently to generate outputs that are then reviewed by another agent for accuracy and post-processing
c. It is when RAG is an integral part of an LLM-powered test infrastructure
d. It is a term for background processing that uses a combination of autonomous and semi-autonomous agents

32. (1 pt) How can overfitting compromise your model during fine-tuning?

a. It can make the LLM too generalized to be applicable and efficient
b. It can introduce hallucinations and biases
c. It can make the model unable or inaccurate when processing data that is new to it
d. It can create a lack of transparency regarding LLM decision-making

33. (1 pt) Which of the following LLM deployment types is most concerned with data privacy?

a. Chatbots
b. Commercial testing tools with AI components
c. In-house developed commercial test tools with AI components
d. Data privacy is a critical concern for all of these deployment types

34. (1 pt) How is the risk of using copyrighted data without proper authorization categorized when using shadow AI?

a. Information security weakness
b. Data privacy weakness
c. Compliance issues
d. Vague intellectual property

35. (1 pt) Why is high-quality input data important when implementing an LLM-powered testing approach?

a. It will be faster with good data that produces fewer errors
b. It will produce reliable results
c. It will allow the largest possible data set to be built
d. It will ensure access is protected and limited to only the correct users

36. (1 pt) You are working to convince your manager that a GenAI tool would be helpful in generating test reports. One of his issues is that he feels he could hire interns to write the test reports, and that would be faster than incorporating a new tool. You are sure that it will be able to write the reports faster and more accurately, freeing critical testing time.

What is an important capability of the LLM that you should evaluate to be sure it will prove you right?

a. Model performance
b. Fine-tuning potential
c. Direct cost
d. Community and support

37. (1 pt) What is the first key phase when adopting GenAI into the test organization?

a. Discovery
b. Initiation and Training
c. Usage Definition
d. Exploitation and Iteration

38. (1 pt) Why does a tester need to understand "prompt engineering" in order to work effectively with GenAI?

a. Because they will have a better grasp of the internal layers of the neural network
b. Because they need to understand what user prompts will need to be tested and the relative risk of each
c. Because they will be asked to generate new user prompts based on the AI-designed UI
d. Because they will need to design and refine the input prompts to guide the LLM

39. (1 pt) When conducting GenAI training meetings with the testing team, what is an important learning method?

a. Discussing challenges and their resolutions
b. Highlighting erroneous prompts used within the team
c. Processing AI outputs to better analyze manual testing gaps
d. Promoting the use of AI methods and nominating an "AI champion of the week"

40. (1 pt) Who is usually responsible for ensuring that test teams maintain both traditional testing competencies and AI literacy?

a. Each tester
b. The senior testers
c. The test manager
d. The AI master