Sample Exam – Questions

Sample Exam set A
Version 1.0

ISTQB® Certified Tester
AI Testing Syllabus

Compatible with Syllabus version 1.0

International Software Testing Qualifications Board

Copyright Notice

Document Responsibility

The ISTQB® Examination Working Group is responsible for this document.

Acknowledgements

Revision History

Sample Exam – Questions Layout Template used:	Version 2.5	Date: May 21st, 2021	

Version	Date	Remarks
1.0	2021/10/01	Release for GA

Table of Contents

Introduction

Purpose of this document

The sample questions and answers and associated justifications in this sample exam set have been created by a team of Subject Matter Experts and experienced question writers with the aim of assisting ISTQB® Member Boards and Exam Boards in their question writing activities.

These questions cannot be used as-is in any official examination, but they should serve as guidance for question writers. Given the wide variety of formats and subjects, these sample questions should offer many ideas for the individual Member Boards on how to create good questions and appropriate answer sets for their examinations.

Instructions

In this document you may find:

- Questions[1], including for each question:
    - Any scenario needed by the question stem
    - Point value
    - Response (answer) option set
- Additional questions, including for each question [does not apply to all sample exams]:
    - Any scenario needed by the question stem
    - Point value
    - Response (answer) option set

- *Answers, including justification are contained in a separate document*

---

[1] In this sample exam the questions are sorted by the LO they target; this cannot be expected of a live exam.

Questions

Question #1 (1 Point)

Which of the following statements provides the **BEST** example of the 'AI Effect'?

- a) People lose their jobs as AI-based systems perform their roles cheaper and better
- b) Competitive computer games lose popularity as AI-based systems always win
- c) Rule-based expert systems for medical diagnosis are no longer considered to be AI
- d) People believe AI will take over the world, as shown in films

Select ONE option.

Question #2 (1 Point)

Which of the following options is **NOT** a technology used to implement AI?

- a) Support vector machine
- b) Decision tree
- c) Evolutionary reasoning
- d) Bayesian optimization

Select ONE option.

Question #3 (1 Point)

Which of the following statements about the hardware used to implement AI-based systems is **MOST** likely to be **CORRECT**?

- a) The processors used to train a mobile recommendation system must be the same as the processors on the mobile phone
- b) Graphical processing units (GPUs) are a reasonable choice to implement an AI-based computer vision system
- c) Deep learning systems need to be trained, evaluated, and tested using AI-specific chips
- d) It is always best to choose processors with more bits to achieve sufficient accuracy for AI-based systems

Select ONE option.

Question #4 (1 Point)

There are a number of good quality pre-trained models available in the market and you want to use one of them for an image-based classifier. You have decided to ask the provider of the model about the data used for training the model and its format.

Which of the following statements is the **BEST** example of a risk that you are trying to mitigate by asking these questions?

   a) Bad classification accuracy of the pre-trained models
   b) Differences in the data used to train the model and the operational data
   c) Performance efficiency issues of the pre-trained model
   d) Lack of explainability of the pre-trained model compared to that of a model trained by you

Select ONE option.

Question #5 (1 Point)

Which of the following statements is **MOST** likely to be specifying a requirement for autonomy in an AI-based system?

   a) The system shall maintain a safe distance to other vehicles until the brake or accelerator is pushed by the driver
   b) The system shall learn the preferred style of response to emails by remotely monitoring the email traffic
   c) The system shall compare its predictions of house prices with actual selling prices to determine if it needs to be retrained
   d) It shall be possible to modify the system's behavior to work with different types of users in less than a day

Select ONE option.

Question #6 (1 Point)

Which of the following statements about bias in AI-based systems is **NOT** correct?

a) Bias may be caused by users of a book recommendation system making choices that deliberately cause the system to make poor suggestions

b) Bias may be caused in the employee age of death prediction system by collecting the training data from a dataset of patients who are all retired

c) Bias may be caused in the creditworthiness system by using training data obtained from those who own and use a credit card

d) Bias may be caused in the navigation system by using a route planning algorithm that is too complex to be explained to typical users

Select ONE option.

Question #7 (1 Point)

Which of the following is **MOST** likely to be an example of reward hacking?

a) The programmer's assistant tool optimizes the code to provide reduced response times, while still ensuring that functional requirements are met

b) An anesthetic supply device with a goal of keeping patients stable during surgery supplies too many doses and patients do not wake up as quickly as expected

c) The third-party development organization paid their AI programmers based on the number of lines of code they write

d) A type of AI used to play competitive computer games against humans that is focused on getting the highest score

Select ONE option.

Question #8 (1 Point)

Given the following attributes for an AI-based system (I-V):

  I.   Probabilistic

  II.   Explicable

  III. Unfair

  IV. Non-deterministic

  V.   Deterministic

Which list of attributes below is likely to cause the **MOST** difficulties if the system is to be used as part of a safety-related system?

  a) I, IV
  b) II, IV
  c) II, III, V
  d) I, III, V

Select ONE option.

Question #9 (1 Point)

Which of the following statements BEST describes classification and regression as part of supervised learning?

  a) Regression is checking that the ML model test results do not change when the same test data is executed
  b) Classification is the grouping of unlabeled data into separate classes
  c) Classification is the labelling of the data for training the ML model
  d) Regression is predicting the number of classes that are output by the ML model

Select ONE option.

Question #10 (1 Point)

Which of the following options **BEST** describes an example of reinforcement learning?

a) The mobile game app updates its feedback, response timing and the number of user options it provides based on how much the players spend
b) The language translation app searches the internet to find text provided in multiple languages to improve its translation function
c) The factory quality control system uses video cameras and audio analysis to identify manufactured items that are faulty based on monitoring a human quality control operative
d) The software component test prediction system uses a range of quality measures to identify which components are likely to contain the most defects

Select ONE option.

Question #11 (2 Points)

You have been asked for your opinion on the ML approach to be used for a new system that is part of the traffic management for a SMART city. The idea is that the new system will control the traffic lights in the city to ensure traffic flows easily through and around the city.

Which of the following approaches do you expect **MOST** likely to succeed?

a) Unsupervised learning that is based on identifying clusters around the city where the traffic density is higher than average
b) A supervised learning regression solution based on thousands of journeys labelled with both journey length and duration
c) Reinforcement learning that is based on a reward function that penalizes solutions that result in higher levels of traffic congestion
d) A supervised learning classification solution that is based on drivers and passengers submitting their favorite routes for traversing the city

Select ONE option.

Question #12 (1 Point)

When performing testing of a trained model, an ML engineer found that the model was highly accurate when evaluated with validation data but that it performed poorly with independent test data.

Which of the following options is **MOST** likely to cause this situation?

    a) Underfitting
    b) Concept drift
    c) Overfitting
    d) Poor acceptance criteria

Select ONE option.

Question #13 (1 Point)

Which of the following is an example of a challenge that is likely to be encountered in the course of developing and testing an ML solution?

    a) Data anonymization operations typically require knowledge of various ML algorithms
    b) The data used might be unstructured data
    c) A large percentage of the budget gets spent just in data preparation
    d) The data pipeline scalability is a challenge when training the model

Select ONE option.

Question #14 (1 Point)

The data scientist has complained that the model cannot be trained with one particular algorithm, although other algorithms work with the same training data.

Which of the following options is the **MOST** likely reason for this?

    a) Wrong data
    b) Missing data
    c) Badly labelled data
    d) Insufficient data

Select ONE option.

Question #15 (1 Point)

DataSure is a start-up with a product that promises to improve the quality of ML models. DataSure claim that this improvement comes from checking if the data has been labeled correctly.

Which of the following defects is **MOST** likely to have been prevented by using this product?

   a) The model will have security vulnerabilities
   b) The model will have poor accuracy
   c) The model will not fulfill its intended function
   d) The model will produce biased outputs

Select ONE option.

Question #16 (1 Point)

An ML engineer, upon finding insufficient training data, is rotating labeled images to create additional training data.

Which of the following approaches to labeling is being applied in this above example?

   a) Crowdsourcing
   b) Augmentation
   c) AI-based labeling
   d) Outsourcing

Select ONE option.

Question #17 (2 Points)

The confusion matrix for an image classifier is shown below.

Confusion Matrix	Actual Positive	Actual Negative
Predicted Positive	78	22
Predicted Negative	6	14

Which of the following options represents the precision of the classifier?

a) 20/120 *100
b) 78/120 *100
c) 78/100 *100
d) 22/100 *100

Select ONE option.

Question #18 (1 Point)

ThermalSpace is a solution provider that helps thermal power plants to optimize their power output. Their solution is based on an ML model created using past data with clearly marked output. The model helps determine the amount of electricity to be generated at a given time of the day.

To determine the quality of the model using ML functional performance metrics, which of the following metrics is **MOST** likely to be used?

a) R-squared
b) Precision
c) Recall
d) False Positives

Select ONE option.

Question #19 (2 Points)

KnowYourPet is an app utilizing ML to determine whether a pet is hungry or not. It is understood that a dog is likely to be not hungry most of the time, as reflected in the training data. If the dog is mis-diagnosed as hungry then it may lead to overfeeding of the dog and this could lead to serious health issues.

Which of the following metrics would you choose for determining the suitability of the model under test?

   a) Accuracy
   b) Precision
   c) Recall
   d) F1-score

Select ONE option.

Question #20 (1 Point)

Which of the following options **BEST** describes a deep neural net?

   a) It is comprised of a hierarchical structure of neurons with the lowest (deepest) neurons making most of the decisions
   b) It is comprised of connected neurons where each neuron has an associated bias and each connection has an associated weight
   c) It is made up several layers with each layer (except input and output layers) connected to each other layer and errors are propagated backwards through the network
   d) It is made up of layers of neurons, each of which generates an activation value based on the other neurons in the same layer

Select ONE option.

Question #21 (1 Point)

Which of the following statements **CORRECTLY** describes a test coverage measure for neural networks?

    a) Value change coverage is based on individual neurons being seen to affect the overall output of the neural network

    b) Threshold coverage is based on neurons outputting an activation value greater than a pre-set value between zero and one

    c) Neuron coverage is a measure of the proportion of neurons that are activated at any time during the testing

    d) Sign change coverage measures the coverage of neurons that output both positive, negative and zero activation values

Select ONE option.

Question #22 (1 Point)

Which of the following requirements for an AI-based system is **MOST** likely to cause a significant challenge in testing?

    a) The system shall be more accurate than the system it is replacing

    b) The AI component in the system shall have 100% accuracy

    c) A human operator should be able to override the system in 1 second

    d) The system shall mimic the human emotions of a typical game player

Select ONE option.

Question #23 (1 Point)

Which of the following is a factor associated with the test data that can make the testing of AI-based systems difficult?

    a) Sourcing big data with high velocity

    b) Sourcing data from a single source

    c) Sourcing data separately from the data scientists

    d) Sourcing data from public websites

Select ONE option.

Question #24 (1 Point)

Why would the accuracy of human decisions be considered in testing as well as the accuracy of AI-based systems?

- a) Intuitive human decisions can be made faster than a corresponding AI-based system in some situations
- b) Unethical decisions can be made by humans as well as AI-based systems
- c) The accuracy of human decisions is not relevant to testing AI-based systems
- d) Human decisions may be of lower quality when they have been recommended by an AI-based system

Select ONE option.

Question #25 (2 Points)

An ML-based toll charging solution determines the type of incoming vehicles from the images captured by a camera. There are different types of cameras available and the solution provider claims to be able to use cameras of different resolutions. The images need to be in jpeg format with a size of 320X480 pixels for the purpose of training the model as well as for predicting the outcome. The model should be able to classify the vehicle types with certain desired high level of accuracy and should be tested against vulnerabilities. Each toll plaza will have its own complete system unconnected to any other system.

Which of the following types of testing are the **MOST** appropriate options for the tests you would choose for system testing?

- a) Testing for concept drift
- b) Adversarial testing
- c) Scalability testing
- d) Fairness testing
- e) Data pipeline testing

Select TWO options.

Question #26 (1 Point)

Which of the following statements **BEST** describes a testing challenge that specifically applies to a self-learning system?

- a) The system requires regular retraining and therefore requires regular testing
- b) The system is regularly released which means regression testing is required
- c) The system changes in such a way that tests that previously passed can fail
- d) The system requires a human operator, who is also required for testing

Select ONE option.

Question #27 (1 Point)

Which of the following is **NOT** likely to be required to test a system for bias?

- a) Involving selected users that are known to be biased
- b) Measuring how changes in test inputs change test outputs
- c) Observing how production outputs correlate to production inputs
- d) Obtaining additional data from other sources

Select ONE option.

Question #28 (1 Point)

Which of the following statements **BEST** describes how system complexity can create challenges when testing an AI-based system?

- a) Testing for bias may require data that the team does not have
- b) Manual generation of white-box tests can be difficult
- c) Determining whether a system is ethical can be subjective
- d) It can be difficult to find representative data to train a model

Select ONE option.

Question #29 (1 Point)

An AI-based system is being used by the health ministry to identify vulnerable groups of patients, who will be provided with support and advice to help prevent them suffering future illnesses to which they may be susceptible. The results will also be shared with other government agencies and medical insurance companies. The system is initially being trained on a large set of data collected by the health ministry from two surveys of 5,000 men over 50 years of age and 25,000 women over 30 years of age. The system will continue to identify vulnerable patients by gathering information from publicly available social media.

Which of the following attributes should be **MOST** carefully considered when specifying the objectives and acceptance criteria for the system?

   a) Adaptability
   b) Bias
   c) Explainability
   d) Flexibility
   e) Autonomy

Select TWO options.

Question #30 (1 Point)

An ML engineer is trying to find exploitable inputs and then use these inputs to retrain the models to make them immune to these inputs.

Which of the following options **BEST** describes the approach being used by the ML engineer?

   a) Validation
   b) Adversarial testing
   c) Data pipeline testing
   d) Scalability testing

Select ONE option.

Question #31 (1 Point)

A test manager has to select test techniques to be used for testing autonomous vehicle software. There are a large number of environmental conditions (>50) that need to be considered for seven vehicle functions.

Which of the following test techniques is MOST likely to be used when testing the variety of vehicle functions (VF) in different environmental conditions (EC)?

   a) A/B testing based on the VF and EC parameters
   b) Combination testing of all the parameters of VF and EC
   c) Pairwise testing of the relevant values of VF and EC
   d) Back-to-back testing of relevant VF and EC values

Select ONE option.

Question #32 (1 Point)

A test manager decides to have a non-AI system with similar functionality to the AI based system under test (SUT) built to support system testing.

Which of the following statements is most likely to be **CORRECT**?

   a) The test manager has chosen back-to-back testing because it helps solve the test oracle problem by using a pseudo-oracle
   b) The test manager has chosen A/B testing because it helps solve the test oracle problem by using a pseudo-oracle
   c) The test manager has chosen back-to-back testing because the non-functional requirements of the SUT can be verified against the pseudo-oracle
   d) The test manager has chosen A/B testing because the non-functional requirements of the SUT can be verified against the pseudo-oracle

Select ONE option.

Question #33 (2 Points)

An AI-based mobile phone search system provides a list of phones that it believes are most suitable for the user based on its knowledge of the user's previous mobile phone usage and their specified preferences.

Given that metamorphic testing is being used with the following source test case:

Inputs		**Outputs**
Selected price range:	$200-$300	Recommended Phones:
3D camera:	Don't care	SnapHappy_X1
Screen size:	mid to large	SnapHappy_M2
OS:	Android or iOS	SnapHappy_M3
Battery Life:	Don't care	ClickNow_1000x
		ClickNow_1000xs

And this test data for two corresponding follow-up test cases:

Input T1	
Selected price range:	$200-$300
3D camera:	yes
Screen size:	mid to large
OS:	Android or iOS
Battery Life:	Don't care
Input T2	
Selected price range:	$200-$300
3D camera:	no
Screen size:	mid to large
OS:	Android or iOS
Battery Life:	Don't care

Which of the following options is **MOST** likely to be a valid list of recommended phones for the follow-up test cases?

    a)   T1: SnapHappy_X1, SnapHappy_M2
           T2: ClickNow_1000x, ClickNow_1000xs

    b)   T1: SnapHappy_M2, SnapHappy_M3, ClickNow_1000xs
           T2: SnapHappy_X1, ClickNow_1000x

    c)   T1: SnapHappy_X1, SnapHappy_M2, SnapHappy_M3, ClickNow_1000x, ClickNow_1000xs
           T2: SnapHappy_X1, SnapHappy_M2, SnapHappy_M3

    d)   T1: SnapHappy_X1, SnapHappy_M2, SnapHappy_M3, ClickNow_1000x, ClickNow_1000xs
           T2: SnapHappy_X1, SnapHappy_M2, SnapHappy_M3, ClickNow_1000x, ClickNow_1000xs

Select ONE option.

Question #34 (1 Point)

System testing of an AI-based system is being planned. It has been suggested that exploratory testing is used in addition to scripted test techniques.

Which of the following scenarios is **MOST** likely to be an example of exploratory testing being performed?

    a)   Training data is visualized using tools to look at various aspects of the data
    b)   Tests written using equivalence partitioning during the previous test cycle are being run
    c)   The Google 'ML test checklist' is being used
    d)   ML functional performance metrics are being calculated

Select ONE option.

Question #35 (2 Point)

LAIgal systems has an AI-based product for extracting relevant favorable judgements similar to a given legal case. This product is used by judges in the courts. Details of the current case are provided, and the system produces relevant judgements. The system needs to be safe from malicious inputs. A similar open-source product exists and is available. Not having a suitable test oracle is a challenge when testing.

Which of the following test techniques should be selected to test the new version during system testing?

   a) A/B testing
   b) Back-to-back testing
   c) Adversarial testing
   d) State transition testing
   e) ML functional performance metrics calculation

Select TWO options.

Question #36 (1 Point)

Which one of the following statements is an example of a difference between a test environment for AI-based systems and a test environment for conventional systems?

   a) Test environments for AI-based systems may require some mechanism to determine how a particular decision is made
   b) Test environments for AI-based systems need simulators and virtual environments whereas conventional systems do not need these
   c) Test environments for AI-based systems need large amounts of data, whereas conventional systems do not need large amount of data
   d) GPUs are required for test environments for AI-based systems whereas conventional systems do not need these

Select ONE option.

Question #37 (1 Point)

In which of the following situations would AI be MOST useful when categorizing new defects?

    a) A small number of defects requires categorization on a new application
    b) A large number of defects is reported on a small application
    c) Minimal data is provided in typical defect reports
    d) A new development team needs to know the most appropriate developer to fix a defect

Select ONE option.

Question #38 (1 Point)

Which of the following is an AI tool **MOST** likely to use as the basis for generating functional test cases?

    a) A test charter
    b) A picture of the system as a flow chart
    c) Web server logs
    d) Crash reports

Select ONE option.

Question #39 (1 Point)

Which of the following options **CORRECTLY** states how an AI-based tool can perform optimization of regression test suites?

    a) By analyzing false positive test results
    b) By analyzing information from previous testing activities
    c) By using genetic algorithms to create new test cases
    d) By updating the expected results to counter concept drift

Select ONE option.

Question #40 (1 Point)

Which of the following options **CORRECTLY** states how an AI-based tool can perform defect prediction?

  a) Using natural language to ask developers where they predict defects will occur
  b) By analyzing the causes of defects raised on a similar code base
  c) By analyzing false positive defects
  d) Scanning code to identify defects using rules.

Select ONE option.

Appendix: Additional Questions

Question #1 (1 Point)

Which of the following statements about AI is **MOST** likely to be **CORRECT**?

a) An autonomous robot that can act as a worker in a house, shop or office is an example of general AI
b) A robot exhibiting similar skill levels as a human is considered to have achieved singularity
c) AI-based systems that support a range of test management functions are considered to possess general AI
d) An AI-based system that cannot access the internet is said to exhibit narrow AI

Select ONE option.

Question #2 (1 Point)

Which of the following statements is MOST likely to be describing a conventional system (as opposed to an AI-based system)?

a) This system assigns customers into groups, based on their historical buying patterns
b) This system controls the braking of the car dependent on its speed
c) This system taught itself to recognize different words by listening to recordings
d) This system detects anomalies from its experience of seeing anomalies in many X-rays

Select ONE option.

Question #3 (1 Point)

Which of the following options is **NOT** a framework used to develop AI-based software?

a) scikit-learn
b) CNTK
c) MxNet
d) EZPy-AI

Select ONE option.

Question #4 (1 Point)

Which of the following statements is **MOST** likely to be describing a system that includes the use of AI as a Service (AIaaS)?

a) The image classifier identifies defects in the gyroscope casings produced by the company and was built using a transfer learning approach so that it is of high accuracy
b) The underwater AI-based vehicle steering system uses a third-party obstacle avoidance component based on decision trees and Bayesian optimization
c) The contract checker uses an exclusive algorithm for determining levels of legal liability, but the pricing part of the contract is separately checked by a generic contract pricing AI component
d) The car rental pricing system is built using AI to support a demand-based algorithm and is hosted in the cloud and made available to all of the company's car rental offices

Select ONE option.

Question #5 (1 Point)

Which of the following options is the **MOST** likely use of a regulatory standard/regulation for an AI-based system?

a) Use of ISO/PAS 21448 (SOTIF) for an unmanned autonomous submarine
b) Use of GDPR for a bank loan decision-making system
c) Use of ISO 26262 for a fully self-driving car
d) Use of GDPR for a drone collision-avoidance system

Select ONE option.

Question #6 (1 Point)

Which of the following statements about flexibility and adaptability is **MOST** likely to be **CORRECT**?

a) Adaptability is important in unsupervised learning as it allows the ML model to learn from data without labels
b) Flexibility is important in supervised learning as it allows the ML model to recognize meaning even when data is poorly labelled
c) Adaptability is important in reinforcement learning systems as such systems must adapt themselves to optimize their reward function
d) Flexibility is important in self-learning systems as it allows them to adapt themselves to unexpected changes in their environment

Select ONE option.

Question #7 (1 Point)

Which of the following statements about the evolution of AI-based systems is **CORRECT**?

 a) Self-learning AI-based systems that continue to work in the same operational environment are not expected to change their behavior
 b) Side effects are not a concern for AI-based systems that change themselves to cope with changes in their environment
 c) AI-based systems must change themselves to cope with changes in system requirements during development
 d) Self-learning systems that physically interact with people, need to be managed to ensure system changes are not dangerous

Select ONE option.

Question #8 (1 Point)

Which of the following examples of an AI-based system is LEAST likely to require special attention with regard to ethical issues?

 a) A computer game that teaches children the benefits of democracy by challenging them to become president
 b) An application that uses data available on social media to provide a trustworthiness rating for job applicants
 c) An autonomous self-powered underwater marine mapping system to create a map of the seabed in international waters
 d) A mobile app that monitors each employee's daily exercise and rewards them with health insurance benefits

Select ONE option.

Question #9 (1 Point)

Which of the following statements about the transparency, interpretability and explainability for AI-based systems is **MOST** likely to be **CORRECT**?

 a) The search engine algorithm used for training students on search engine technology was selected as it was considered to be the most explainable
 b) The loan system was considered transparent as for each loan application it was clear to users how it decided whether to approve agree to the loan or not
 c) The doctors were happy with the level of interpretability of the rule-based oncology system as they could understand how the given rules were implemented in the system
 d) The drone operators were happy with the transparency of the control system as they felt that the system responded correctly to their instructions

Select ONE option.

Question #10 (1 Point)

Which of the following **BEST** describes the unsupervised approach to machine learning?

  a) Data and labels are analyzed to group them into clusters
  b) A system automatically learns by satisfying a fitness function
  c) A system teaches itself to meet objectives based on rewards
  d) Data are analyzed to identify patterns in the data

Select ONE option.

Question #11 (1 Point)

Given the following descriptions:

   I.   Model performance is checked using validation data

   II.   The origin of test data used to test the model is identified

   III. The tuned model is made ready for its target hardware

   IV.   Test data are used to ensure the agreed ML functional performance criteria are met

   V.   The model is created from source code

   VI.   The critical data features are identified

Which of the following options **BEST** matches the descriptions with the activities in the ML workflow?

  a) III - Build & Compile Model / II - Prepare the Data / IV - Evaluate the Model
  b) V - Build & Compile Model / II - Prepare the Data / I - Evaluate the Model
  c) V - Build & Compile Model / VI - Prepare the Data / IV - Evaluate the Model
  d) III - Build & Compile Model / VI - Prepare the Data / I - Evaluate the Model

Select ONE option.

Question #12 (1 Point)

Which of the following statements is **LEAST** likely to be used as a rationale for selecting an ML algorithm?

  a) The amount of memory available for training the translation system in the mobile device
  b) The maximum time allowed for retraining the embedded health monitoring system
  c) The number of measured characteristics used as the basis for a sports prediction system
  d) The number of expected clusters of customer types for a retail marketing system

Select ONE option.

Question #13 (1 Point)

Which of the following statements about the test dataset is **CORRECT**?

    a) The test dataset comes from a source totally different from the validation dataset
    b) The format of the test dataset is different from the format of the validation dataset
    c) The test dataset can be used as the validation dataset but not as the training dataset
    d) The test dataset should not be exposed to the model during the training process

Select ONE option.

Question #14 (1 Point)

Which of the following options would **MOST** likely be a reason for poor labeling of data?

    a) Insufficient data
    b) Synthetic data
    c) Translation errors
    d) Algorithm chosen for the ML model

Select ONE option.

Question #15 (1 Point)

An ML team asserts that the ML functional performance metrics based on validation data collected as part of training an ML model are sufficient for determining the quality of the system.

Which of the following statements is a valid reason to show that this may be **INCORRECT**?

    a) The ML functional performance metrics may not work well if the ground truth is not correct
    b) The ML functional performance metrics cannot be used for measuring quality as these are tool dependent
    c) Validation data is biased resulting in skewed functional performance measurements
    d) Data may need to be transformed prior to training the model, so the functional performance measurements do not reflect the quality of the model

Select ONE option.

Question #16 (1 Point)

Which of the following options regarding benchmark suites, **BEST** completes the following statement?

a) ML benchmark suites help choose a particular model by indicating the time it takes to train
b) ML benchmark suites help choose a particular model by indicating the time it takes to test
c) ML benchmark suites help choose a particular model by indicating the time it takes to validate
d) ML benchmark suites help choose a particular model by indicating the time it takes to deploy

Select ONE option.

Question #17 (1 Point)

Which of the following test levels provides the **BEST** choice for performing bias-related testing?

a) Component testing
b) Input data testing
c) System testing
d) Model testing

Select ONE option.

Question #18 (1 Point)

Which of the following statements about the documentation of AI components is **CORRECT**?

a) Because non-functional requirements are not a part of the documentation of an AI component, non-functional testing cannot be performed
b) White-box testing of the interaction of AI and non-AI components is not possible if the interfaces are a part of the documentation
c) Checking for bias in the data is made possible by including the source of the data in the documentation
d) Self-adapting AI systems require each change made by the system to be fully documented

Select ONE option.

Question #19 (1 Point)

An ecommerce application recommends products to the user on the basis of their purchase history and purchases being made by other people using the site, among other factors. As the tester, you have been asked to measure the current conversion rate of recommendations to compare with the original required conversion rate.

Which of the following is the **MOST** likely underlying reason for this request?

    a) AI effect
    b) Adversarial attacks
    c) Concept drift
    d) Lack of fairness

Select ONE option.

Question #20 (1 Point)

Which of the following options is **MOST** likely to be relevant when testing a system's autonomy?

    a) Testing over a sustained period of time
    b) Testing the accuracy of system predictions
    c) Testing how quickly the system can adapt
    d) Static analysis of training data

Select ONE option.

Question #21 (1 Point)

Which of the following statements demonstrates how non-deterministic systems can create challenges in testing?

    a) Non-deterministic systems produce a different result each time, normally preventing the generation of expected results
    b) Non-deterministic systems are difficult to test because they are not explainable, which hinders the investigation and fixing of defects
    c) A system that is given the same inputs and initial state may produce different outputs, so can require multiple test executions
    d) Non-deterministic systems are usually biased and require additional tests to allow this bias to be excluded from the results

Select ONE option.

Question #22 (1 Point)

When testing a deep neural network, which of the following characteristics can be evaluated without using dynamic testing?

- a) Explainability
- b) Transparency
- c) Automation bias
- d) Accuracy

Select ONE option.

Question #23 (1 Point)

Which of the following characteristics of an example AI-based system might cause a test oracle problem?

- a) It is not known where the training data was obtained
- b) The output of the system is a prediction for which the ground truth is unknown
- c) The system runs with no human intervention and is considered to be autonomous
- d) There is a lack of transparency into how the system was implemented

Select ONE option.

Question #24 (1 Point)

Which of the following statements is **CORRECT** with regards to A/B testing?

- a) A/B testing is also known as differential testing because two different programs are used for this type of testing
- b) A/B testing is mostly useful for testing simple ML models as it does not produce accurate results for complex models
- c) A/B testing requires multiple expected results from the same inputs to be compared to identify significant differences in tested models
- d) A/B testing is a good technique for writing test cases for various types of ML models, especially self-learning systems

Select ONE option.

Question #25 (1 Point)

Which of the following statements is **LEAST** likely to be an example of a benefit provided by virtual test environments in the testing of an AI-based system?

  a) A smart city application is tested in a virtual test environment to allow scenarios that very rarely happen (e.g., crowd control at a new year event) to be tested
  b) An AI-based money market trading system is tested in a virtual test environment as it would be impractical to test on real money markets due to the potential costs
  c) A smart fruit picker is tested in a virtual test environment to allow many picking scenarios to be run in an accelerated timeframe to check that fruit is not damaged
  d) An autonomous car is tested in a virtual test environment to allow potentially dangerous test scenarios to be run safely

Select ONE option.

Question #26 (1 Point)

It is possible to use an AI-assisted test automation tool to recognize objects through image processing, rather than using references to their location.

To which type of AI software engineering technology is this example **MOST** likely referring?

  a) Classification, Learning and Prediction
  b) Probabilistic Software Engineering
  c) Search-based Software Engineering
  d) Clustering

Select ONE option.

Question #27 (1 Point)

Which of the following statements is **CORRECT** with respect to visual testing?

  a) Visual testing makes use of object IDs to determine changes
  b) Visual testing uses images to do pixel-by-pixel comparisons
  c) Visual testing helps find overlapping user interface elements
  d) Visual testing fails when the screen layout changes

Select ONE option.