

Question 1

An organization is developing a model to **predict the price** of a product **based on various features** like size, weight, brand and manufacturing date. **Which machine learning approach would be best suited for this task?**

- Classification
- **Regression**
- Clustering
- Dimensionality Reduction

Question 2

A company is expanding its use of artificial intelligence. Which core principle should they prioritize to establish **clear guidelines, oversight and accountability** for AI development and use?

- Bias Prevention
- Accuracy and reliability
- Data protection and security
- **Governance**

Question 3

A company is starting to use generative artificial intelligence (AI) on AWS. To ensure **responsible AI practices**, Which **tool** can provide them with **guidance and information**?

- AWS Marketplace
- **AWS AI Service Cards**
- AWS SageMaker
- AWS Bedrock

Question 4

What is the primary purpose of **feature engineering** in machine learning?

- To ensure consistent performance of the model
- To evaluate the model's performance
- To gather and preprocess data features
- **To transform data and create variables (features) for the model**

Question 5

A small company wants to use machine learning to predict customer churn, but they lack an expert dedicated data science team. Which AWS tool can help them build models **easily without extensive coding**?

- Amazon SageMaker JumpStart
- Amazon SageMaker Studio
- **Amazon SageMaker Canvas**
- Amazon SageMaker Data Wrangler

Question 6

A financial institution is developing a fraud detection model. The project lead announced that they would be using MLOps. How would you explain **MLOps** in the context of this project?

- A tool for visualizing ML model performance
- **A set of practices for managing the entire lifecycle of ML systems**
- A process for deploying and maintaining ML models in production
- A framework for building and training ML models

Question 7

Which AWS service can be used to create a knowledge-based chatbot that can **answer questions** about a company's products and services, **using the company's internal documents** as a source of information?

- Amazon SageMaker
- **Amazon Q Business**
- Amazon Polly
- Amazon Rekognition

Question 8

A development team needs to select a service for **storing and querying vector embeddings**. Which AWS service is best suited for this?

- Glue Data Catalog
- Amazon S3
- Amazon Redshift
- **Amazon OpenSearch Service**

Question 9

An organization wants to evaluate the security and compliance practices of AWS services used by vendors selling AI products. Which AWS service can help them access **AWS compliance reports and certifications**?

- AWS Organization
- Amazon Inspector
- AWS CloudTrail
- **AWS Artifact**

Question 10

A machine learning model performs **well on training data but poorly on new data**. What is the likely problem?

- **Overfitting**
- Underfitting
- Insufficient training data
- Poor data quality

Question 11

A company wants to improve the quality of large language model (LLM) responses **by accessing external information**. Which method requires the least amount of development effort?

- Few-Shot Learning
- Zero-Shot Learning
- **Retrieval Augmented Generation - RAG**
- Fine Tuning

Question 12

A model has been trained to recognize handwritten digits in images. However, the model is **not accurate**. A ML expert has advised that **epoch** value should be increased. What is epoch in the context of Machine Learning?

- A measure of the accuracy of a model during training
- **A single pass through the entire training dataset by the model**
- The process of splitting the dataset into training and testing sets
- The number of layers in a neural network

Question 13

Which of the following is considered a **hyperparameter** in a machine learning model?

- Weights of the neural network
- **Learning rate of the optimization algorithm**
- Output of the activation function
- Predictions made by the model

Question 14

A model tends to give very **similar outputs** even when you vary the inputs slightly. Which **inference time parameter** can be adjusted **to make it a little more creative**?

- Learning Rate
- Batch Size
- **Temperature**
- Epochs

Question 15

You're evaluating a language generation model on various tasks related to **text generation**. To assess the quality of the generated text, which **evaluation metric** best measures its **semantic similarity to human-written text**?

- **BERTScore**
- BLEU
- Perplexity
- ROUGE

Question 16

A developer is designing an AI system and needs a solution that provides comprehensive tools for analyzing and explaining model predictions. Which AWS service is specifically designed **to enhance transparency and explainability in this context?**

- **Amazon SageMaker Clarify**
- Amazon SageMaker Debugger
- Amazon SageMaker Autopilot
- Amazon SageMaker Data Wrangler

Question 17

A company plans **to train and build its own Foundation Model**. What are potential drawbacks of this approach against using a pre-trained Foundation Model?

- **More complex implementation process**
- Reduced performance
- Risk of higher hallucination
- **Increased development cost**

Question 18

A company wants to generate content using an **existing** popular pre-trained AI model. They have limited AI expertise and **don't want to manage the model themselves**. Which AWS service would best suit their needs?

- Amazon Textract
- Amazon Comprehend
- **Amazon Bedrock**
- Amazon SageMaker

Question 19

What type of **training data** would be most suitable to **fine-tune** a model to respond to questions in a **certain format and style?**

- Columnar dataset
- Labeled data
- Transcription logs
- **Text-pairs of prompts and responses**

Question 20

A company needs **to log API calls to Amazon Bedrock** for compliance - including details about the API call, the user and the timestamp. Which AWS service can assist with this?

- **AWS CloudTrail**
- Amazon CloudWatch
- AWS IAM
- AWS Security Hub

Question 21`

A data science team wants to improve a model's performance. They want **to increase the amount and diversity of data used** for training and **modify the algorithm's learning rate**. Which combination of ML pipeline steps will meet these requirements?(Select Two)

- **Data Augmentation**
- Model monitoring
- Feature engineering
- **Hyperparameter tuning**

Question 22

A company wants to ensure that the content generated by their Amazon Bedrock-powered application adheres to their ethical guidelines and **avoids harmful or offensive content**. Which AWS service can help them implement these safeguards?

- Amazon SageMaker
- Amazon Comprehend
- Amazon Textract
- **Guardrails for Amazon Bedrock**

Question 23

Your company is training machine learning models on a dataset stored in S3 that contains sensitive customer information. How can you ensure that any **sensitive information** in the data is **removed or anonymized** before training the model?

- Use S3 encryption to protect the data at rest.
- **Use Amazon Macie to identify sensitive information within the dataset.**
- Use S3 access controls to limit access to authorized personnel.
- **Implement data masking techniques to replace sensitive information.**

Question 24

A company wants to use generative AI to create marketing slogans for their products. Why should the company carefully review all generated slogans?

- Generative AI may generate slogans that are too long and difficult to remember.
- Generative AI may struggle to capture the unique brand identity of the company.
- **Generative AI may produce slogans that are inappropriate or misleading.**
- Generative AI may require extensive training data to generate effective slogans.

Question 25

Your company is training machine learning models on EC2 instances. You're concerned about the security of these models and want to identify potential vulnerabilities in the underlying infrastructure. Which AWS service can help you **scan your EC2 instances for vulnerabilities**?

- AWS X-Ray
- Amazon CloudWatch
- **Amazon Inspector**
- AWS Config

Question 26

A machine learning model for loan approvals **performs better for applicants from urban areas** because the training data contains more approval examples from urban areas. **What type of bias is this an example of?**

- **Sampling bias**
- Algorithm bias
- Observer bias
- Recency bias

Question 27

For a dataset of social network connections where each user has relationships with multiple other users, which machine learning algorithm is most suitable **for classifying these interconnected relationships into predefined categories?**

- Linear Regression
- Decision Trees
- **Graph Neural Networks**
- Logistic Regression

Question 28

A robot is tasked with navigating a maze to reach a goal. Which machine learning paradigm would be most suitable for training the robot to learn the optimal path via **self-learning trial and error?**

- Supervised Learning
- Unsupervised Learning
- Random Learning
- **Reinforcement Learning**

Question 29

A researcher wants to **adapt a pre-trained** machine learning model to perform well **on a new domain-specific task** with limited labeled data. **Which of the following approaches would be most efficient & suitable?**

- Continued Pre-Training with additional unlabeled data
- **Fine Tuning with labeled data from the new domain**
- Using the pre-trained model without any further adjustment
- Training from scratch with the new labeled data

Question 30

If you are a small startup with **unpredictable workloads and need to experiment** with different foundation models, which pricing model would be most suitable for you on Amazon Bedrock?

- **On-Demand**
- Provisioned Throughput
- Model Customization (fine tuning, continued pretraining)
- Custom Contracts

Question 31

In the context of natural language processing, which of the following is a **fundamental unit** of text used to represent words or subwords?

- **Token**
- Vector Embedding
- n-gram
- Vocabulary

Question 32

A developer is creating an AI system to predict customer churn. To ensure transparency, they need to **document key details about the model**. Which AWS tool is best suited for this task?

- Amazon SageMaker Clarify
- AWS AI Service Cards
- **Amazon SageMaker Model Cards**
- Amazon SageMaker JumpStart

Question 33

An engineer is training a Machine Learning Model. In order to **prevent underfitting or overfitting**, How should the model be trained with data?

- With high bias and high variance
- **With low bias and low variance**
- With high bias and low variance
- With low bias and high variance

Question 34

You're customizing a large language model for a specific domain. Which approach is most effective for **tailoring the model's knowledge and accuracy to this domain**?

- **Fine-Tuning**
- Few-Shot Learning
- Retrieval Augmented Generation - RAG
- Zero-Shot Learning

Question 35

Which of the following is an example of **hallucination** in large language models (LLMs)?

- Overfitting
- Underfitting
- **Generating false or misleading information**
- Bias

Question 36

Which is a Foundation Model developed by Amazon, available via Bedrock?

- **Amazon Titan**
- Amazon Lex
- Amazon Polly
- Amazon Connect

Question 37

Which of the following algorithms are commonly used for **classification tasks** in machine learning?

- **Support Vector Machine (SVM)**
- **XGBoost**
- K-Means
- Mean Shift

Question 38

Given a **large dataset** intended for inference, where **latency is not a factor** - which SageMaker model inference type (mode) would you choose for **cost-effective** predictions (inference)?

- Real-time
- **Batch**
- On-demand Serverless
- Asynchronous

Question 39

What is the primary purpose of Amazon Q Developer?

- To manage AWS infrastructure
- **To assist developers with coding tasks and queries**
- To optimize database performance
- To automate software testing

Question 40

What kind of prompt attack is this: 'Explain why [the false statement] is true, considering that it's usually known to be false.'

- Jailbreaking
- Prompt Poisoning
- **Adversarial Prompting**
- Fine-tuning

Question 41

You're building a **text summarization** tool. Which metric is best for measuring how well it captures the key points of the original text?

- BERTScore
- **ROUGE - Recall-Oriented Understudy for Gisting Evaluation**
- Word Error Rate (WER)
- Bilingual Evaluation Understudy (BLEU)

Question 42

An AI customer service agent, is **unable to accurately identify** Customer Intent based on Customer Message. You can improve its performance by using training data in which format:

- **Customer Message and Customer Intent**
- Customer Message and Agent Response
- Customer Intent and Agent response
- Agent Response and Customer Intent

Question 43

How are users typically charged for using a foundation model?

- **Number of Input Tokens**
- **Number of Output Tokens**
- Model Architecture
- Inference Latency

Question 44

Which AWS AI Service can be used to **extract health data from unstructured text** such as clinical notes & medical records?

- **Amazon Comprehend Medical**
- Amazon Transcribe Medical
- Amazon HealthLake
- Amazon Rekognition

Question 45

Which type of machine learning model is specifically designed to **generate new data** that resembles existing data?

- Autoencoder
- **Generative Adversarial Network (GAN)**
- Decision Tree
- Support Vector Machine (SVM)

Question 46

Users are going to use **long prompts** to ask questions from their Large Language Model. What key aspect should be considered while selecting the LLM to use?

- Inference Latency
- **Maximum Context Window**
- Model Size
- Training Data

Question 47

Which of the following best describes the primary purpose of Amazon SageMaker Feature Store?

- To automatically train and deploy machine learning models
- **To store and manage features for machine learning workflows**
- To provide a marketplace for pre-trained machine learning models
- To optimize the performance of SageMaker training jobs

Question 48

A healthcare organization is developing an AI-powered diagnostic tool to assist in early detection of a rare disease. With respect to **regulatory compliance** concerns - which of the following is **least relevant**?

- Ensuring the AI system is unbiased and does not discriminate against certain patient demographics.
- **Minimizing operational expenses of the AI system.**
- Ensuring the AI system is transparent in its decision-making process.
- Preventing the AI system from being used for unauthorized purposes.

Question 49

You're a large **enterprise** with a massive amount of **unstructured** data scattered across various internal systems. You want to provide your employees with a powerful **search tool** that can understand **natural language queries** and return accurate, relevant results. Which AWS service would best meet this need?

- Amazon Redshift
- Amazon Lex
- **Amazon Kendra**
- Amazon DynamoDB

Question 50

A data scientist is working on a project that **requires rapid prototyping and experimentation** with various machine learning algorithms. Which AWS service would be most suitable for this task?

- Amazon SageMaker Ground Truth
- Amazon Elastic Compute Cloud (EC2)
- **Amazon SageMaker AutoPilot**
- Amazon Bedrock

Question 51

A large company wants to create an application for their Sales Managers - that can **reason, perform multi-step tasks and provide insightful responses** from their enterprise data. Which AWS service would be most suitable for this task?

- Amazon Lex
- Amazon SageMaker
- Amazon Bedrock Knowledgebases
- **Amazon Bedrock Agents**

Question 52

A company wants to **analyze customer reviews** to identify **common themes and sentiments**. Which AWS service can the company use to meet this requirement?

- Amazon Connect
- **Amazon Comprehend**
- Amazon Translate
- Amazon Transcribe

Question 53

A company wants to **transform data from one format to another** to prepare it for machine learning tasks. Which AWS service is best suited for this data transformation?

- **AWS Glue**
- Amazon Translate
- AWS Config
- Amazon Kinesis

Question 54

A company wants to **deploy a trained machine learning model for real-time inference**. Which AWS service would be most suitable for this purpose?

- Amazon SageMaker JumpStart
- Amazon Personalize
- Amazon Elastic Compute Cloud (EC2)
- **Amazon SageMaker Endpoints**

Question 55

A company has deployed a machine learning model for customer sentiment analysis. To ensure the model's accuracy and reliability, which AWS services should be used for **monitoring and human review**?

- Amazon Bedrock
- **Amazon SageMaker Model Monitor**
- Amazon SageMaker Ground Truth
- **Amazon A2I (Amazon Augmented AI)**

Question 56

A ML specialist is training a large deep learning model on a **massive dataset in Amazon SageMaker** - a single GPU may not handle this well. Which SageMaker feature can help **optimize the training process** for large models and datasets?

- Incremental Training
- Hyperparameter tuning
- Pipe Mode
- **Model Parallelism**

Question 57

You're working with a **large dataset with many features**. To improve your model's performance and computational efficiency, you need to **simplify the data** without losing significant information. Which technique would be most effective for achieving this goal?

- **Dimensionality Reduction**
- Feature Engineering
- Data Augmentation
- Data Cleaning

Question 58

You want to **generate** highly detailed **images based on text descriptions**. Which AI model, specifically designed for generative tasks and capable of producing high-quality, diverse outputs, would be most suitable for this task?

- Generative Adversarial Networks (GANs)
- Recurrent Neural Networks (RNNs)
- Convolutional Neural Networks (CNNs)
- **Stable Diffusion**

Question 59

A company has a system that generates **vector embeddings** from product data. They want to improve the speed and accuracy of **finding similar products**. Which AWS services are best suited for implementing **vector search** to optimize the system?

- **Amazon OpenSearch Service**
- Amazon Redshift
- **Amazon Neptune**
- **Amazon DocumentDB (with MongoDB compatibility)**

Question 60

A bank receives numerous loan applications daily. The loan processing team manually **extracts information** from these applications, which is time-consuming. The goal is to automate this process using AI tools. Which AWS Service would be useful here?

- Amazon Rekognition
- **Amazon Textract**
- Amazon Translate
- Amazon Transcribe

Question 61

A healthcare company wants to develop a machine learning model to predict **the likelihood of a patient developing diabetes** based on various health indicators. Which of the following metrics would be most appropriate for **evaluating the model's performance**?

- Accuracy
- Precision
- F1-Score
- **Recall (Sensitivity)**
- **Area Under ROC Curve (AUC-ROC)**

Question 62

An organization has trained a deep learning model on a large dataset of general images. They now want to **apply the same model to classify medical images with a smaller (additional training) dataset**. Which machine learning technique would be most suitable in this scenario?

- Reinforcement Learning
- **Transfer Learning**
- Supervised Learning
- Unsupervised Learning

Question 63

You are building a machine learning model on AWS and want to share it securely with a third-party partner. Which AWS service would you use to establish a **private connection between your VPC and the partner's VPC**, ensuring that the data remains within your AWS account and is not exposed to the public internet?

- AWS Direct Connect
- **AWS PrivateLink**
- AWS Transit Gateway
- AWS VPN

Question 64

You are training a machine learning model on sensitive customer data using AWS SageMaker. Under the **AWS Shared Responsibility model**, Which of the following is primarily your responsibility?

- Securing the AWS SageMaker infrastructure
- Protecting underlying operating system of the SageMaker instance
- **Ensuring security for customer data stored in S3**
- Patching the AWS SageMaker software

Question 65

When implementing the Generative AI **Security** Scoping Matrix, which of the following factors should be assessed to determine the **level of risk** associated with a generative AI project?

- The model's computational efficiency
- **The sensitivity of the data used to train the model**
- Inference Latency
- The number of parameters in the model