Question 1 0 / 1 point Consider a dataset with p features and n observations. Which of the following statements is true for Principal Component Analysis? There are at most max(n + 1, p) principal components. There are at least min(n-1, p) principal components. There are at most max(n-1, p) principal components. There are at most min(n-1, p) principal components. Question 2 0 / 1 point Which of the following hyperparameters controls the size of the region of the input image that the convolutional filter processes at each step? Batch size Number of epochs Kernel size Learning rate **Question 3** 0 / 1 point Which of the following is a major limitation of the K-Means clustering algorithm? K-Means cannot handle non-numeric data. K-Means can only be used for clustering spherical clusters. K-Means requires the user to specify the number of clusters after the model is trained.

K-Means is sensitive to the initial placement of centroids.

Question 4 0 / 1 point

What is the primary goal of Principal Component Analysis (PCA)?

- To increase the number of dimensions in the dataset to improve performance.
- To reduce the dimensionality of the data while preserving as much variance as possible.
  - To remove outliers from the dataset.
  - To classify the data into different categories based on features.

Question 5 0 / 1 point

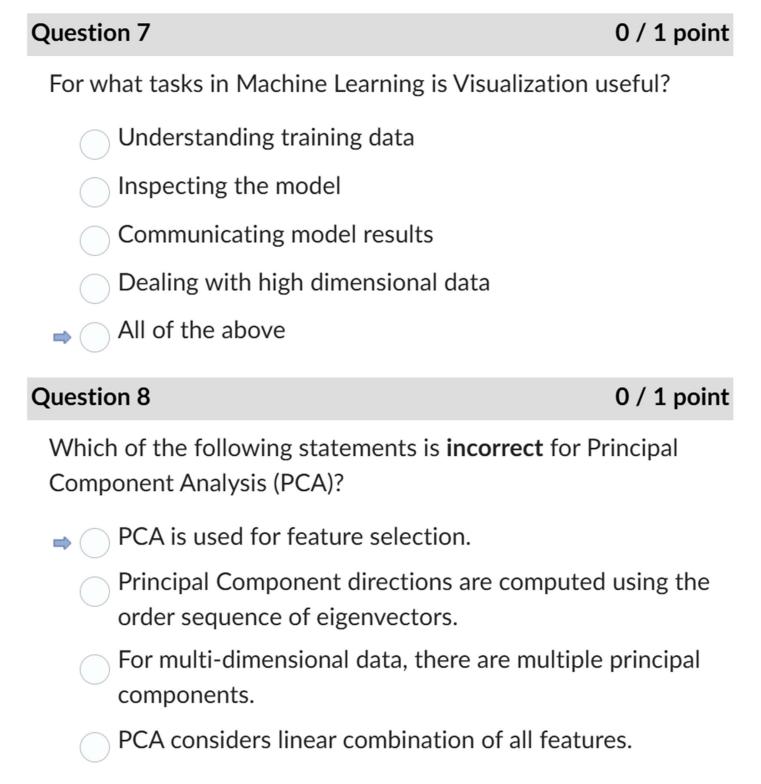
Which of the following types of problems is Naive Bayes best suited for?

- Problems where classes are not linearly separable.
- Problems where features are strongly correlated.
- Text classification tasks such as spam email detection.
  - Tasks involving a large number of continuous features with no categorical variables.

Question 6 0 / 1 point

Which of the following statements best describes the primary difference between **generative** and **discriminative** algorithms?

- Discriminative models can create new data instances, while generative models focus on modeling different kinds of instances.
- Discriminative models require fewer features than generative models to achieve the same level of accuracy.
- Discriminative models learn the probability of the classes given the data, while generative models learn the joint probability of the data and the class.
  - Generative models typically perform better on large datasets, while discriminative models are best for smaller datasets.



Question 9 0 / 1 point

Which of the following is **NOT** a core component of a Markov Decision Process (MDP)?

- States (S)
- → Loss Function (L)
  - Rewards (R)
  - Actions (A)

Question 10 0 / 1 point

What is the key feature of a Recurrent Neural Network (RNN) that distinguishes it from traditional feedforward neural networks?

- RNNs process data sequentially, maintaining a memory of previous inputs through feedback connections.
  - RNNs have no internal state and do not capture temporal dependencies.
  - RNNs use a fixed-size architecture with no connections between layers.
  - RNNs are only used for image classification tasks.

Question 11 0 / 1 point

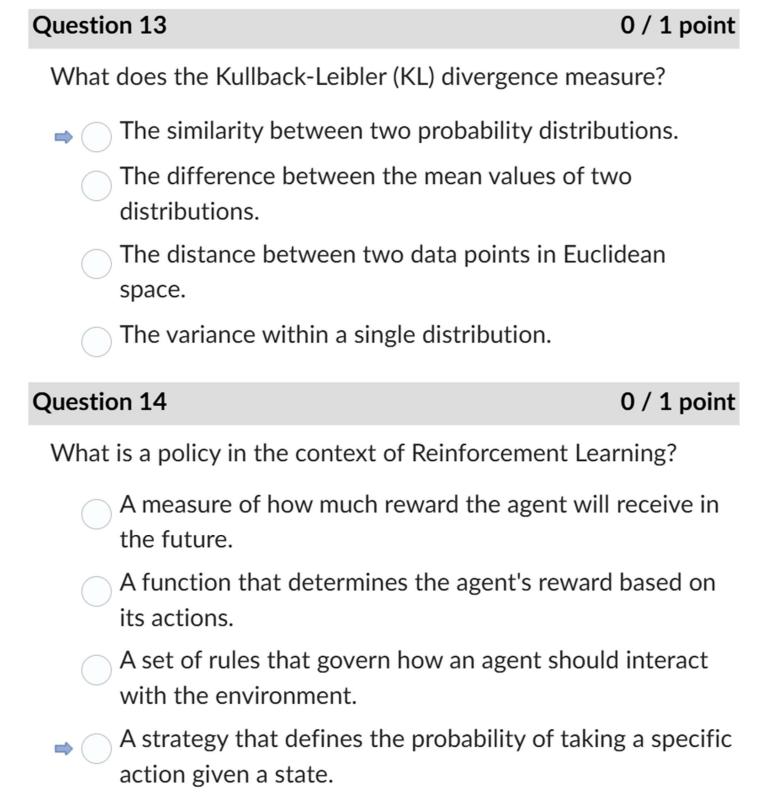
Which of the following is a key assumption made by the Naive Bayes algorithm?

- Features are conditionally independent given the class label.
  - Features are highly correlated with each other.
  - The class distribution follows a Gaussian distribution.
  - The dataset contains only binary features.

Question 12 0 / 1 point

Which of the following is a variant of Recurrent Neural Networks (RNNs) designed to address the vanishing gradient problem?

- → Long Short-Term Memory (LSTM)
  - Convolutional Neural Network (CNN)
  - Generative Adversarial Network (GAN)
  - Autoencoder



Question 15 0 / 1 point

What does the Bellman Equation represent in the context of Reinforcement Learning?

- A recursive relationship that helps calculate the optimal state-value functions for the agent.
  - A method to calculate the transition probabilities between states.
  - A function used to update the agent's position in the environment.
  - A formula that calculates the reward for each action taken by an agent.

Question 16 0 / 1 point

Consider a 5x5 input image with the following values:

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Consider you are applying a 3x3 filter (kernel) with the following values:

1	0	-1
1	0	-1
1	0	-1

Perform a **valid** convolution operation, which means the filter is applied only where it fits entirely inside the input image. What is the resulting value at the **top-left** corner of the output feature map?

- -12
- <u>6</u>
- **⇒** -6
  - 12

Question 17 0 / 1 point

Consider the below equation for reinforcement learning, where R is the reward at state t (T is the terminal state) and there are k states to sweep over. With G being the overall reward during learning, what is gamma (  $\gamma$ )?

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots + \gamma^{k-1} R_{T+k+1}$$

- Number of Actions
- → Discount Rate
  - Learning Rate
  - Iteration Number

Question 18	0 / 1 point
Which of the following is a key characteristic of hierarchical c	lustering?
It uses a predefined number of clusters to begin with a to optimize them.	nd iterates
It requires the user to specify the number of clusters be starting the algorithm.	efore
It starts with all data points as individual clusters and m them iteratively.	nerges
It starts by treating all data points as one single cluster them iteratively.	and splits
Question 19	0 / 1 point
Which of the following layers is typically used in Convolutional Networks to reduce the spatial dimensions of feature maps, where we will be a spatial dimension of feature maps, where we will be a spatial dimension of feature maps.	
Convolutional Layer	
Fully Connected Layer	
ReLU Layer	
→ Pooling Layer	
Question 20	0 / 1 point
In Reinforcement Learning, what does an agent aim to maxim	ize?
⇒ The cumulative reward over time.	
The total number of actions taken.	
The total number of states available.	
The number of states visited	