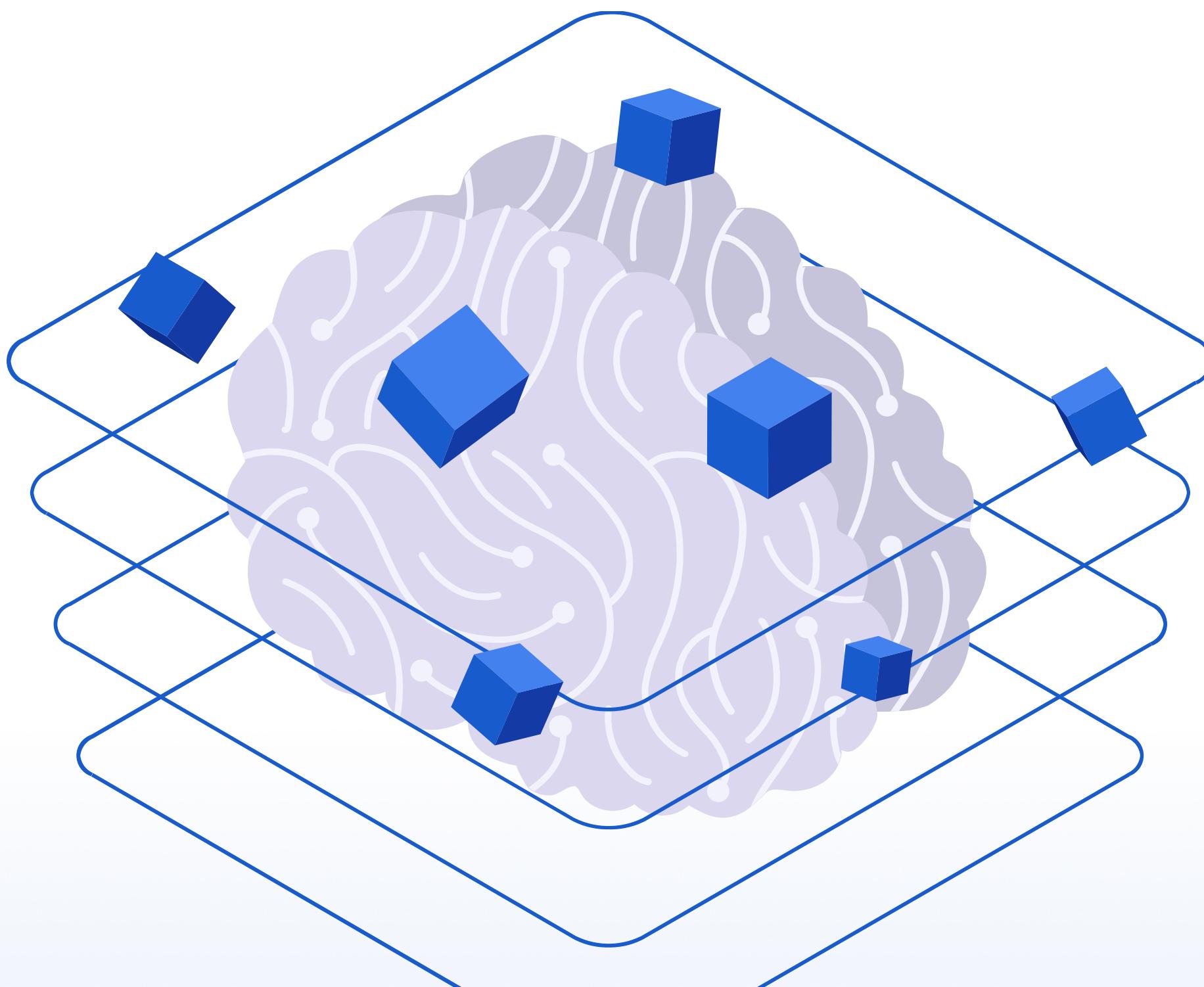


LEARN

# MACHINE LEARNING

through these

Most Asked Interview Questions

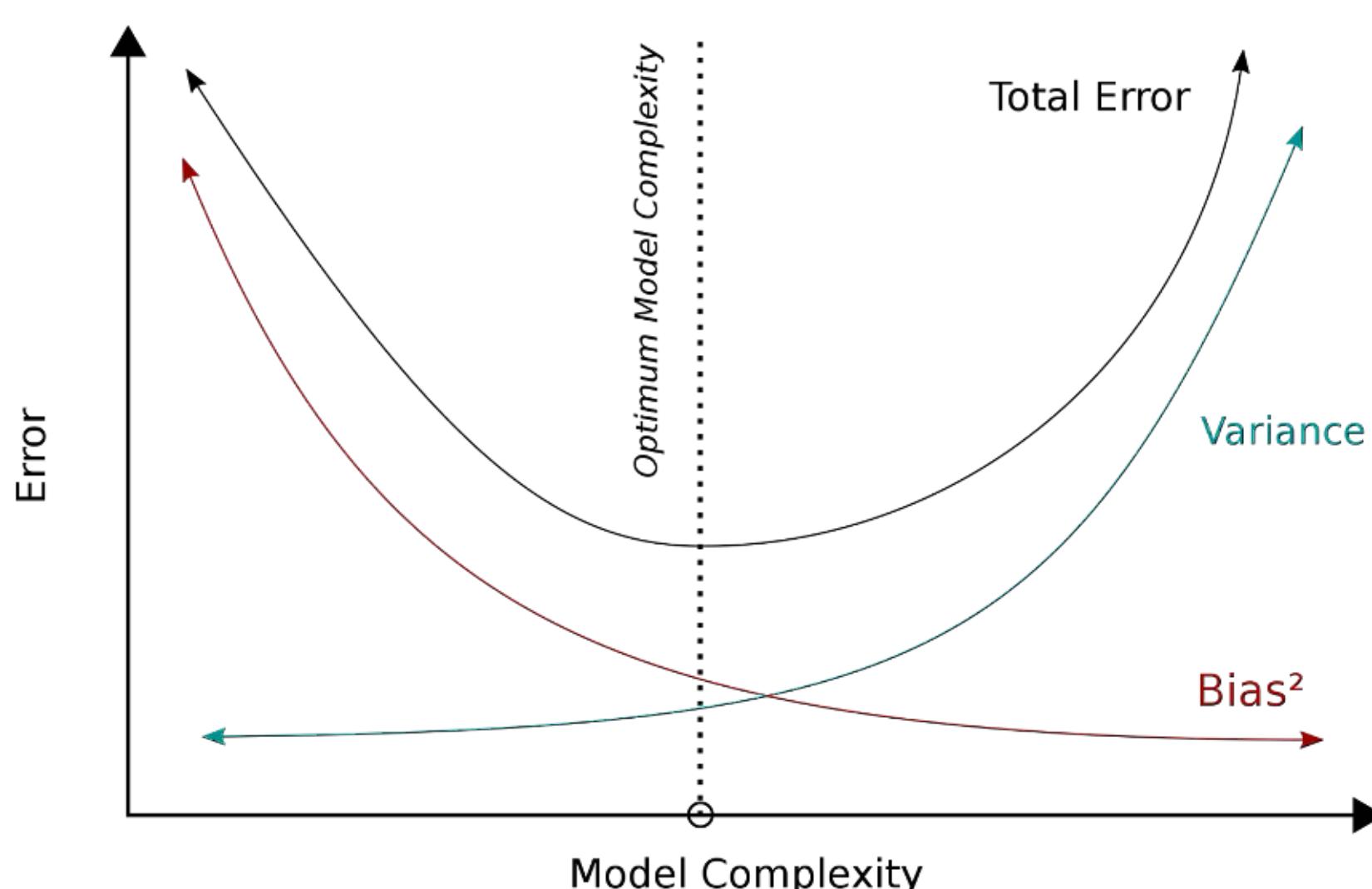


**Q. 1**

# Explain Bias-Variance Tradeoff.

**Ans. 1**

The bias-variance tradeoff represents the balance between the model's ability to generalize across different datasets (bias) and its sensitivity to small fluctuations in the training set (variance). A high-bias model is too simple and underfits the data, missing the underlying trend. A high-variance model is too complex, overfitting the data and capturing noise as if it were a real pattern. The goal is to find a sweet spot that minimizes the total error.

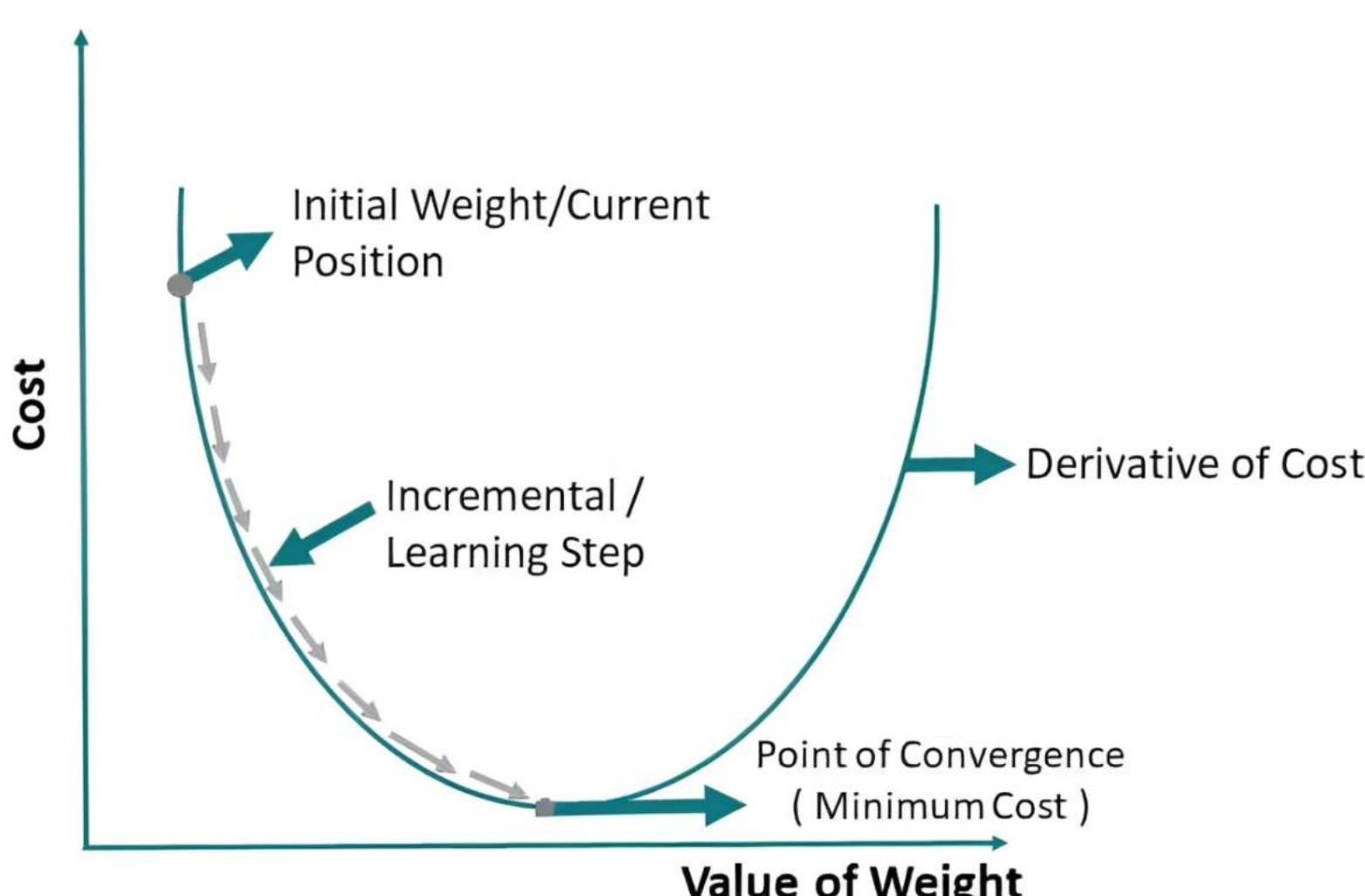


**Q. 2**

# How does Gradient Descent Work?

**Ans. 2**

Gradient Descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of the steepest descent as defined by the negative of the gradient. In machine learning, it's used to find the parameters of a model that minimize the cost function. The learning rate determines the size of the steps taken to reach the minimum.

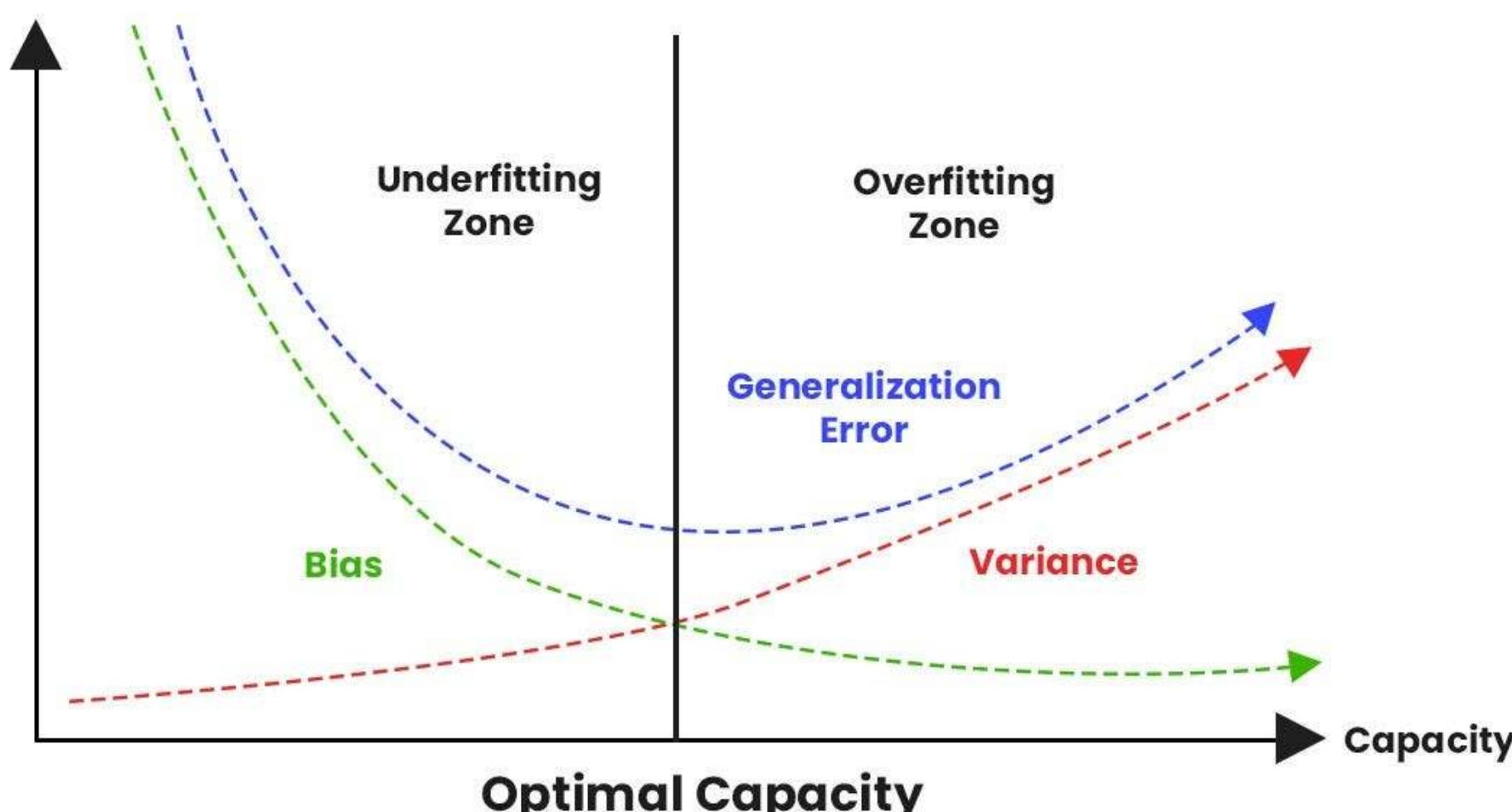


**Q. 3**

# What is Regularization? Give Examples.

**Ans. 3**

Regularization is a technique used to prevent overfitting by adding a penalty on the size of the coefficients. The penalty term discourages complex models and thus reduces variance without substantially increasing bias. Examples include L1 regularization (Lasso), which adds the absolute value of the magnitude of coefficients as penalty, and L2 regularization (Ridge), which adds the square of the magnitude of coefficients.



**Q. 4**

# Explain the Difference between Bagging and Boosting.

**Ans. 4**

Both Bagging and Boosting are ensemble techniques to improve model predictions, but they work differently.

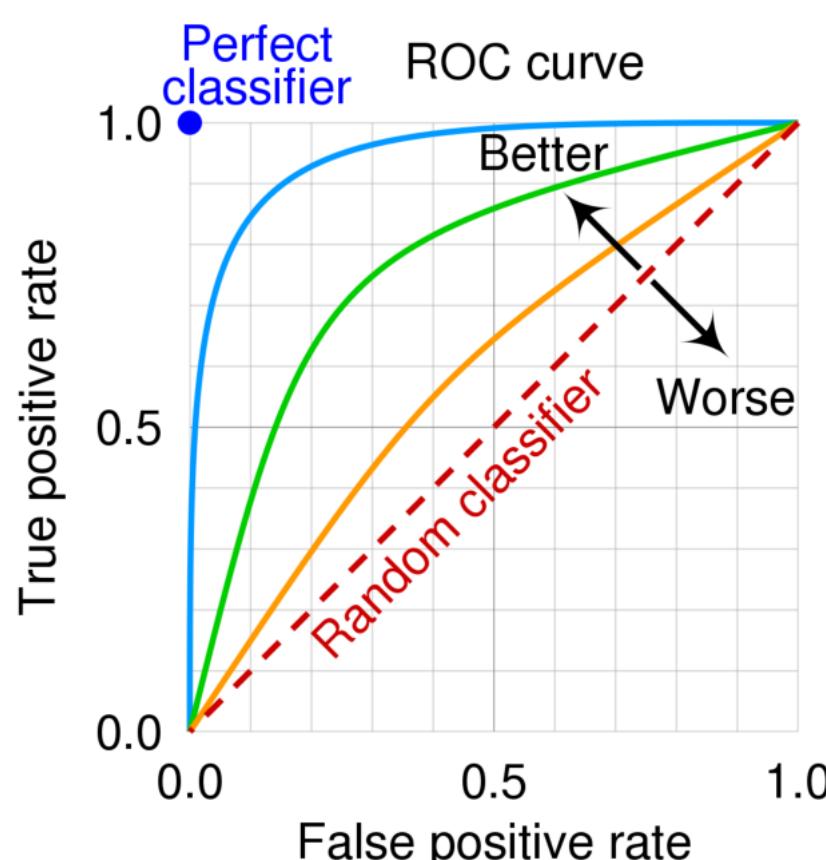
| S.NO | <b>Bagging</b>  | <b>Boosting</b>  |
|------|---|--|
| 1    | The simplest way of combining predictions that belong to the same type.                               | A way of combining predictions that belong to the different types.                 |
| 2    | Aim to decrease variance, not bias.   | Aim to decrease bias, not variance.  |
| 3    | Each model receives equal weight.   | Models are weighted according to their performance.                                |
| 4    | Each model is built independently.  | New models are influenced by the performance of previously built models.           |
| 5    | Different training data subsets are randomly drawn with replacement from the entire training dataset. | Every new subset contains the elements that were misclassified by previous models. |
| 6    | Bagging tries to solve the over-fitting problem.  | Boosting tries to reduce bias.   |
| 7    | If the classifier is unstable (high variance), then apply bagging.                                    | If the classifier is stable and simple (high bias) the apply boosting.             |
| 8    | Example: The Random Forest model uses Bagging.  | Example: The AdaBoost uses Boosting techniques                                     |

**Q. 5**

# Describe the ROC Curve and AUC.

**Ans. 5**

The ROC Curve (Receiver Operating Characteristic Curve) is a graph showing the performance of a classification model at all classification thresholds. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR). AUC (Area Under the ROC Curve) measures the entire two-dimensional area underneath the entire ROC curve and provides an aggregate measure of performance across all possible classification thresholds. An AUC of 1 represents a perfect model; an AUC of 0.5 represents a worthless model.

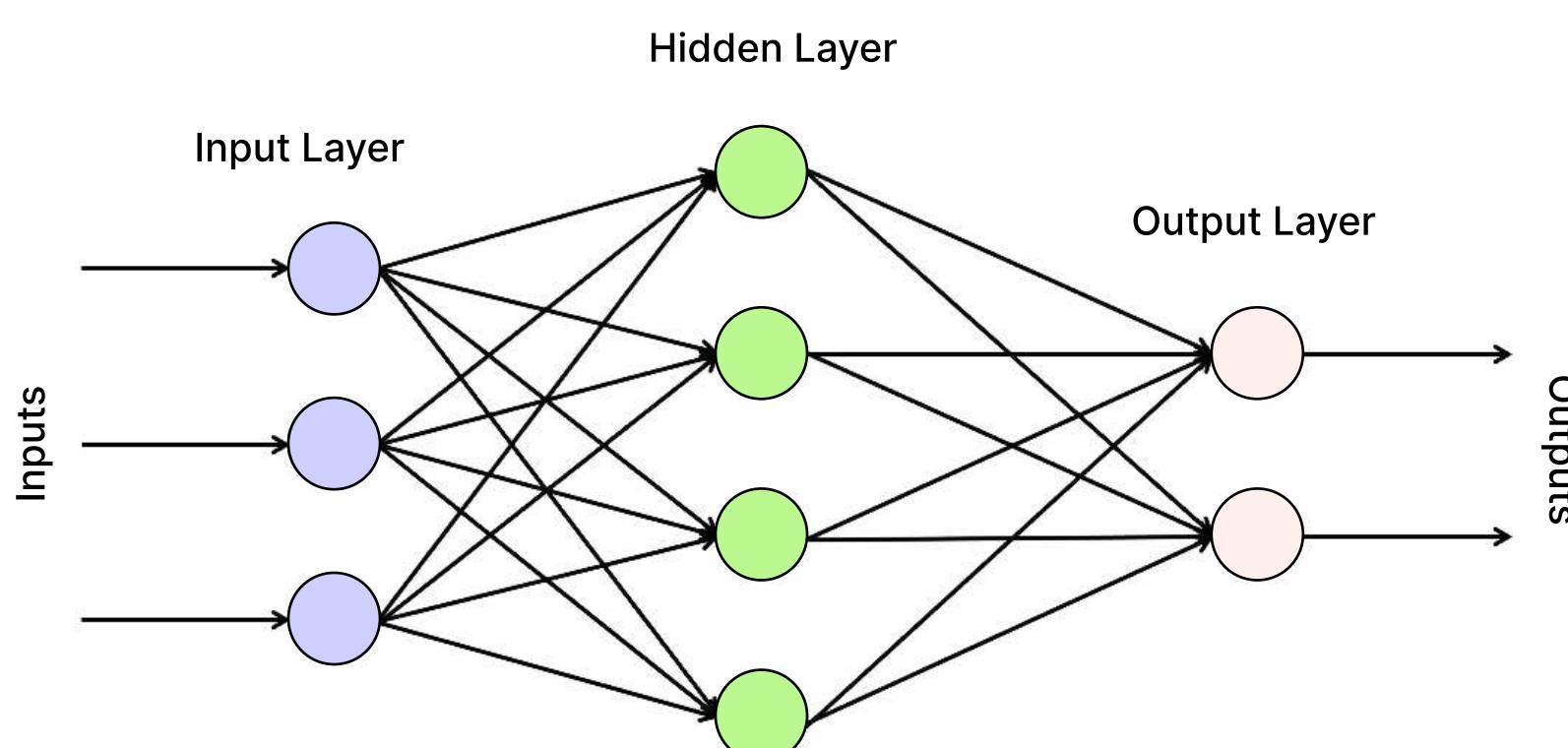


**Q. 6**

# What are Convolutional Neural Networks (CNNs) and where are they used?

**Ans. 6**

CNNs are a class of deep neural networks, most commonly applied to analyzing visual imagery. They use a mathematical operation called convolution in at least one of their layers. A key feature of CNNs is their ability to automatically and adaptively learn spatial hierarchies of features from images. CNNs are widely used in image and video recognition, recommender systems, and natural language processing.

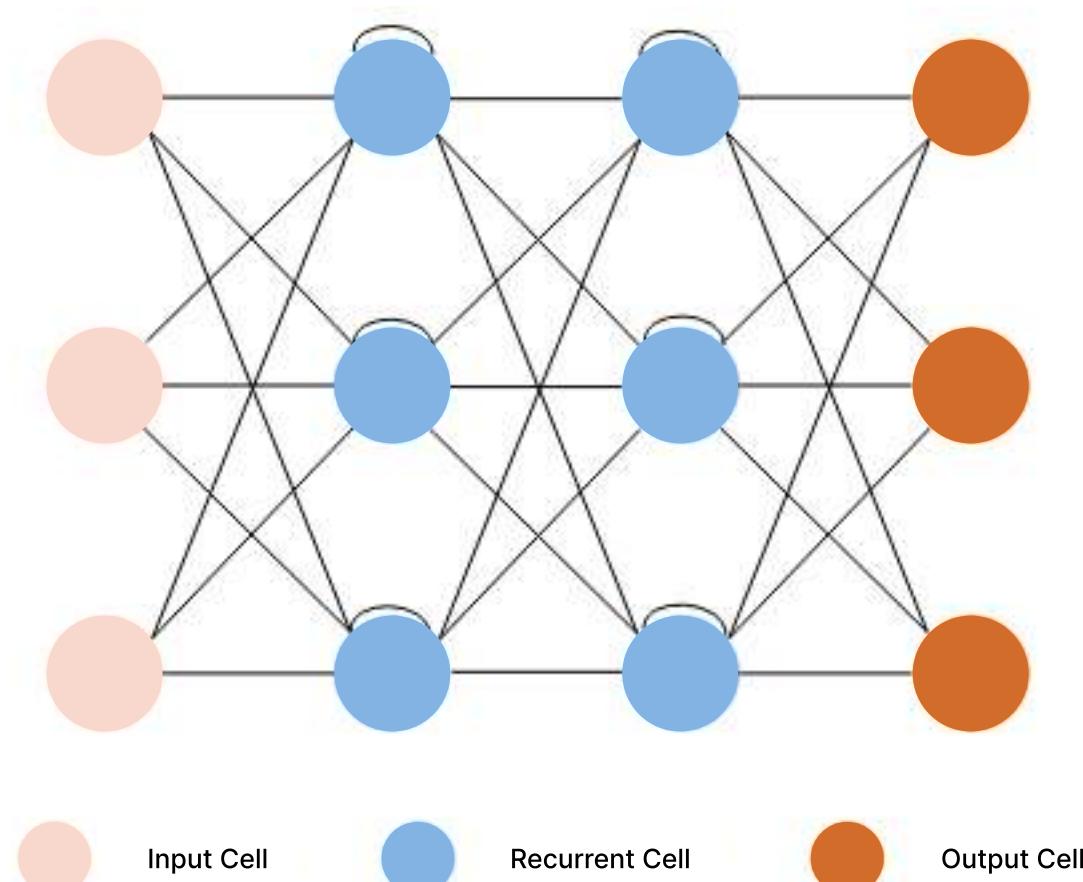


**Q. 7**

# How do Recurrent Neural Networks (RNNs) differ from CNNs?

**Ans. 7**

While CNNs are primarily used for spatial data (like images), RNNs are designed to work with sequence data (like text or time series). RNNs have loops allowing information to persist, meaning they can keep track of information in a sequence, making them ideal for tasks like language modeling and text generation. Unlike CNNs, RNNs can handle inputs of varying lengths.

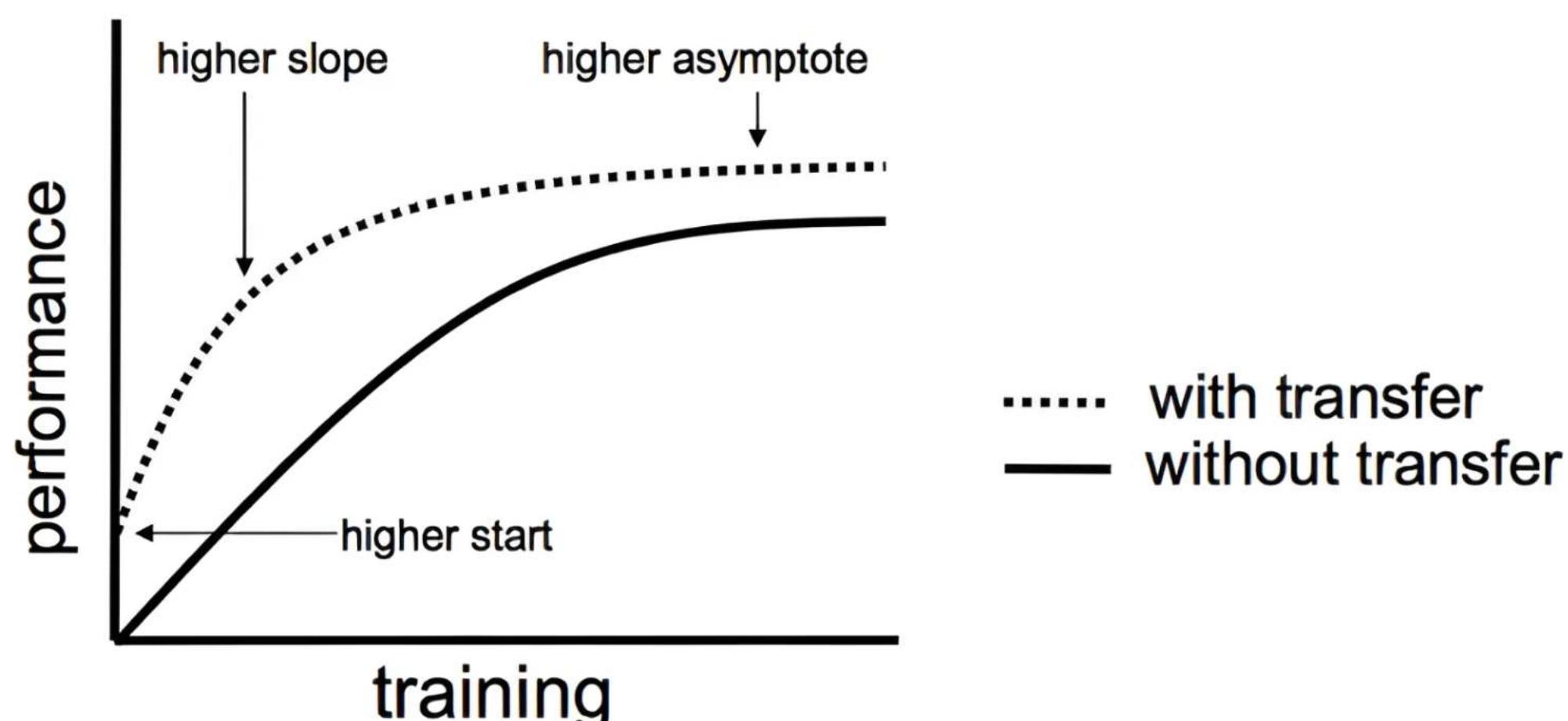


**Q. 8**

## Explain the concept of Transfer Learning.

**Ans. 8**

Transfer Learning involves taking a pre-trained model (trained on a large dataset) and fine-tuning it with a smaller dataset for a similar or different task. This approach allows leveraging learned feature maps without starting from scratch, saving time and computational resources. It's particularly useful in deep learning where large datasets and extensive training are usually required.

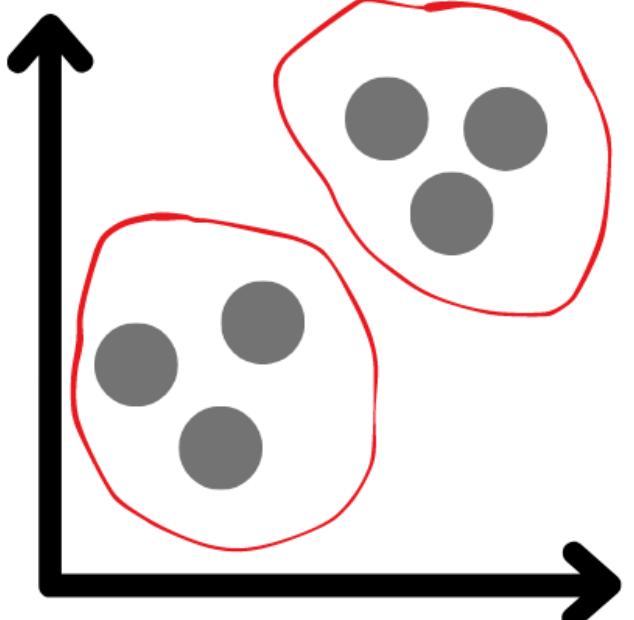
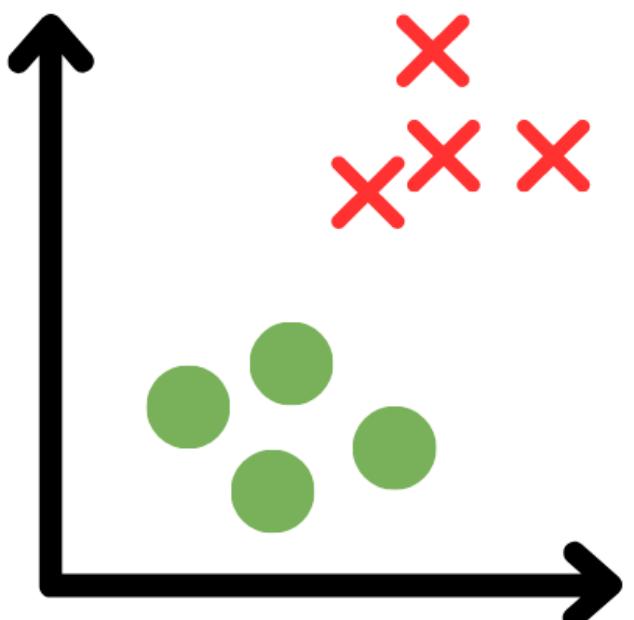


**Q. 9**

# What is the difference between Supervised and Unsupervised learning?

**Ans. 9**

| <b>Supervised learning</b>   | <b>Unsupervised learning</b>   |
|--|--|
| Input data is labeled  | Input data is unlabeled  |
| Has a feedback mechanism   | Has no feedback mechanism  |
| Data is classified based on the training dataset                                 | Assigns properties of given data to classify it                                    |
| Divided into Regression & Classification   | Divided into Clustering & Association  |
| Used for prediction  | Used for analysis  |
| Algorithms include: decision trees, logistic regressions, support vector machine | Algorithms include: k-means clustering, hierarchical clustering, apriori algorithm |
| A known number of classes  | A unknown number of classes  |

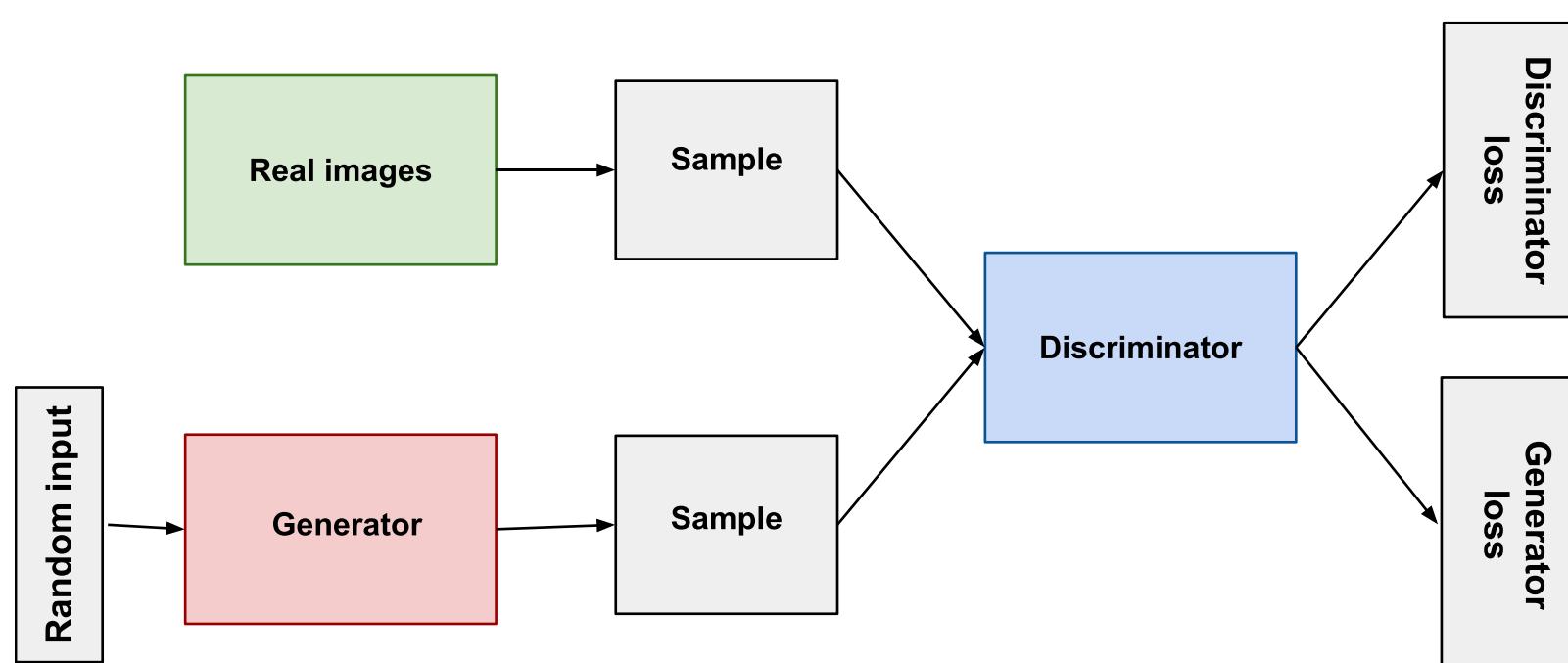


# Q. 10

# What are GANs, and how do they work?

# Ans. 10

Generative Adversarial Networks (GANs) consist of two models: a generative model that captures the data distribution, and a discriminative model that estimates the probability that a sample came from the training data rather than the generative model. The two models are trained simultaneously in a game; the generator tries to produce data that is indistinguishable from real data, while the discriminator tries to distinguish between real and fake data.



# Case-Study Based Questions

**How would you approach building a model to predict stock prices for the next day?**

## Approach

This problem requires analyzing time series data. One approach could be to use RNNs or LSTM (Long Short Term Memory) networks to capture temporal dependencies and trends in historical price data. Feature engineering will also be crucial, incorporating not just price but also volume, historical averages, and potentially external data like economic indicators.

# Imagine you need to classify emails into spam and not-spam. How would you design this system?

## Approach

This is a classic example of a supervised learning classification problem. One could use Naive Bayes, SVM, or deep learning models like CNNs for text classification. The key is to convert emails into a suitable format for these models, using techniques like TF-IDF or word embeddings. Performance can be improved by including more contextual features and fine-tuning the model.

# You're tasked with designing a recommendation system for a streaming service. What approach would you take?

## Approach

A collaborative filtering approach could be initially adopted, leveraging user-item interactions. Matrix factorization techniques like SVD can be employed here.

For a more sophisticated solution, one could use deep learning-based models that can also incorporate content-based features (like genre or director of a movie) to make recommendations even for new users or items (the cold start problem).

# Develop a strategy for a model that can translate spoken language in real-time.

## Approach

This problem involves both speech recognition and machine translation. An end-to-end deep learning approach could be used, where an RNN model first transcribes speech to text, and then another model translates the text to the target language.

Attention mechanisms and Transformer models would be crucial for handling long sequences and improving translation accuracy.

# How would you build a model to identify objects in a video in real-time?

## Approach

This requires a combination of object detection and video processing techniques. CNNs, specifically models like YOLO (You Only Look Once) or SSD (Single Shot Detector), could be used for object detection. These models need to be optimized for real-time performance, considering the computational constraints. Additionally, incorporating temporal information using RNNs or 3D CNNs could improve detection accuracy by considering object movement over time.

Each of these case-study questions touches on different aspects of machine learning and data science, requiring a combination of theoretical knowledge and practical application. Approaching these problems involves understanding the data, choosing the right model, feature engineering, and iterative improvement based on performance metrics.



## WHY BOSSCODER?

 **750+** Alumni placed at Top Product-based companies.

 More than **136% hike** for every **2 out of 3** working professional.

 Average package of **24LPA**.

The syllabus is most up-to-date and the list of problems provided covers all important topics.

Lavanya  
 Meta



Course is very well structured and streamlined to crack any MAANG company

Rahul .  
 Google



[EXPLORE MORE](#)