# Interview Question Of Artificial Intelligence(AI) & Machine Learning(ML)

1. **What is the difference between AI, Machine Learning, and Deep Learning?**

   - **Answer**: ○ **AI** is the broader concept of machines mimicking human-like intelligence.
     - **Machine Learning** is a subset of AI that focuses on algorithms that allow machines to learn from data.
     - **Deep Learning** is a subset of machine learning that uses neural networks with many layers (deep networks) to analyze various factors of data. While ML can work with structured data, deep learning excels in handling unstructured data such as images, audio, and text.

2. **What is supervised learning? Give an example.**

   - **Answer**: Supervised learning is a type of machine learning where the model is trained on labeled data, meaning each training example is paired with an output label. The model learns to predict the output based on the input data. An example of supervised learning is a spam detection algorithm that learns to classify emails as spam or not based on labeled examples.



3. **What is overfitting, and how can you prevent it?**

   - **Answer**: Overfitting occurs when a model learns the training data too well, including noise and outliers, resulting in poor performance on unseen data. To prevent overfitting, you can:
     - Use techniques like cross-validation.
     - Prune decision trees or use regularization methods (like L1 and L2).
     - Use dropout in neural networks
     - Simplify the model by reducing its complexity or using fewer features.

4. **What is the purpose of a confusion matrix?**

   - **Answer**: A confusion matrix is a performance measurement tool for classification problems. It summarizes the correct and incorrect predictions made by a model by comparing the predicted labels with the true labels. It provides metrics such as accuracy,

precision, recall, and F1-score, which help in evaluating the model's performance more comprehensively.

## 5. Explain the difference between classification and regression.

- **Answer**: Classification and regression are both types of supervised learning.
    - **Classification** involves predicting a discrete label (e.g., whether an email is spam or not), where the output is categorical.
    - **Regression** involves predicting a continuous value (e.g., predicting house prices), where the output is a real number.

## 6. What are some common evaluation metrics for regression models?

- **Answer**: Common evaluation metrics for regression models include:
    - **Mean Absolute Error (MAE)**: The average of the absolute differences between predicted and actual values.
    - **Mean Squared Error (MSE)**: The average of the squared differences between predicted and actual values.
    - **Root Mean Squared Error (RMSE)**: The square root of MSE, which provides error in the same units as the output variable.
    - **R-squared**: A statistical measure that represents the proportion of the variance for the dependent variable that's explained by the independent variables in the model.



## 7. What is a neural network, and how does it work?

- **Answer**: A neural network is a computational model inspired by the human brain, consisting of interconnected nodes (neurons) organized in layers: input, hidden, and output layers. Each connection has an associated weight. The network processes input data through these layers, applying activation functions to introduce non-linearity, and produces an output. The model is trained using backpropagation, adjusting the weights based on the error between predicted and actual outputs.

## 8. What is feature selection, and why is it important?

- **Answer**: Feature selection is the process of identifying and selecting a subset of relevant features (variables, predictors) for building a machine learning model. It is important because:
- ○ It helps reduce the dimensionality of the data, improving model performance and reducing overfitting.
    - It enhances the model's interpretability by eliminating irrelevant or redundant features. ○ It decreases training time and computational cost.

## 9. What are hyperparameters, and how do you optimize them?

- **Answer**: Hyperparameters are parameters whose values are set before the learning process begins and control the learning process (e.g., learning rate, number of trees in a forest, or depth of a neural network). Hyperparameter optimization involves searching for the best combination of hyperparameters to improve model performance. Techniques for optimization include grid search, random search, and using algorithms like Bayesian optimization.

## 10. What is regularization, and why is it used?

- **Answer**: Regularization is a technique used to prevent overfitting by adding a penalty to the loss function during model training. Common forms of regularization include:
  - **L1 Regularization (Lasso)**: Adds the absolute value of the magnitude of coefficients as a penalty term to the loss function, promoting sparsity.
  - **L2 Regularization (Ridge)**: Adds the squared magnitude of coefficients as a penalty term, discouraging large weights. Regularization helps in creating simpler models that generalize better on unseen data.