1. What does classification refer to in machine learning?

-> Classification is a part of supervised machine learning in which we put labelled data for training, where the goal is to predict the categorical label or class of a given input based on historical data.

2. Can we use Accuracy all the time?

-> Yes

3. What is Precision in model evaluation, and why is it important in spam detection?

-> Precision is the ratio of True Positive cases to All predicted positive cases.

Precision is important in spam detection because:

A. Precision ensures spam filters accurately identifying spam, minimizing false positives where important emails are mistakenly blocked.
B. High precision improves user trust and avoids missed crucial messages.

4. How does the F1 Score help in evaluating a classification model?
-> The F1 Score helps by combining **precision** and **recall** into one number. It shows how well a model balances finding true positives and avoiding false alarms, giving a simple overall measure of accuracy.

5. Why is it important to split the dataset into training and testing sets?
-> HOMEWORK

B. Long answer type questions.

1. Explain Accuracy, Precision, Recall, and F1 Score with examples. When should each metric be used?

-> Accuracy is the ratio of correctly predicted instances (both positive and negative) to the total number of predictions.
Accuracy = (TP + TN) / (TP + TN + FP + FN)
Example:
Suppose you have a spam email classifier:
- True Positives (TP) = 80 (Spam correctly predicted as spam)
- True Negatives (TN) = 90 (Not spam correctly predicted)
- False Positives (FP) = 10 (Not spam wrongly marked as spam)
- False Negatives (FN) = 20 (Spam wrongly marked as not spam)

Accuracy = (80 + 90) / (80 + 90 + 10 + 20) = **85%**
Use it when:
- The classes are balanced (e.g., similar number of spam and not-spam emails).

- You care about overall correctness.

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.
Formula:
Precision = TP / (TP + FP)
Example:
From above:
TP = 80, FP = 10
Precision = 80 / (80 + 10) = **88.9%**
Use it when:
- False positives are **more costly** than false negatives.
- Example: **Cancer detection** — you don't want to wrongly tell someone they have cancer.

Recall is the ratio of correctly predicted positive observations to all actual positives.
Formula:
Recall = TP / (TP + FN)
Example:
TP = 80, FN = 20
Recall = 80 / (80 + 20) = **80%**
Use it when:
- False negatives are **more costly** than false positives.
- Example: **Disease screening** — better to flag a few healthy people than miss a sick person.

F1 Score is the **harmonic mean** of Precision and Recall.
Formula:
F1 Score = 2 * (Precision * Recall) / (Precision + Recall)
Example:
Precision = 88.9%, Recall = 80%
F1 Score = 2 * (0.889 * 0.8) / (0.889 + 0.8) ≈ **84.2%**
Use it when:
- There is **imbalance in class distribution** (e.g., 90% not-spam, 10% spam).
- You need a **balance** between Precision and Recall

2. Which metric is more important-Recall or Precision? Explain in detail.

-> The right metric depends on the problem you're solving.
Recall is the ratio of correctly predicted positive observations to all actual positives whereas **Precision** is the ratio of correctly predicted positive observations to the total predicted positive observations.

If the cost of a false positive is higher → Go for Precision.

Example : Email Spam Filter
You don't want important emails marked as spam (false positive).

If the cost of a false negative is higher → Go for Recall.
Example : Disease Screening (like COVID tests)
Missing a sick person (false negative) could cause them to infect others.

3. Describe the concept of overfitting and underfitting. How can they impact model evaluation, and how can we prevent them?
-> **Overfitting** and **Underfitting** are common issues in machine learning that impact model performance. **Overfitting** occurs when a model learns the training data too well, including noise and minor details, leading to poor performance on new, unseen data. In contrast, **Underfitting** happens when a model is too simple to capture the underlying patterns in the data, resulting in poor accuracy on both training and test data.

Evaluating these models gives a wrong idea about their performance-where overfitting will give high accuracy on training but low accuracy on testing, **underfitting** will offer low accuracy for both. Methods to prevent overfitting include cross-validation, **regularization**, choosing simpler models, dropout in neural networks, and adding more training data. To solve **underfitting**, one should work with more complex models, leverage better features, or train the models for some more period. One needs at an interface where the model will understand enough from the data and remain neither too complicated nor too simple.

4. Why is model evaluation important in machine learning? Discuss different techniques used to evaluate classification models.

-> **Model evaluation** is crucial in machine learning because it helps determine how well a model performs on unseen data. It ensures that the model is not just memorizing the training data (overfitting) or performing poorly due to simplicity (underfitting).

OR

Model evaluation is crucial in machine learning because -
A. It helps identify the best model among different options.
B. Ensures that the model generalizes well to new data.
C. Detect issues like overfitting or underfitting.

Techniques to Evaluate Classification Models:
A. Accuracy - Ratio of correctly predicted instances to total instances.
B. Precision - Measures how many predicted positives are actually correct.
C. Recall - Measures how many actual positives were correctly identified.
D. Confusion Matrix - A table showing TP, TN, FP, and FN to evaluate overall performance.