CBSE CLASS X

Artificial Intelligence (417)

ANSWER KEY

From previous CBSE Board Exam questions

Code: SIAL4J	Questions: 43	Maximum Marks: 68	Generated: 2026-06-21 02:53

SELECTIONS USED

Source	Previous-year board
Subject	Artificial Intelligence
Lessons	Evaluating Models
Year	2022-2026
Questions selected	43

Composition — Types: 23 MCQ · 7 Short · 6 Case-based · 4 fill_in_blank · 2 Very short · 1 true_false

Q1. § Unit Overview **[2]**

What is the purpose of Evaluation stage of AI project cycle? Discuss briefly.

Previously asked in: 2022 104 Q14

◆ Revisiting AI Project Cycle & Ethical Frameworks for AI | 1.1 AI Project Cycle | 3.1 Importance of Model Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

The **Evaluation stage** assesses how well the trained AI model performs on new/test data. Its purpose is to measure the model's accuracy and efficiency, identify errors or weaknesses, and improve the model before deployment. It ensures the AI solution produces reliable and acceptable results in real-world conditions.

Source: AI Project Cycle, Chapter 1, Section 1.1

Explanation

- The question is 2 marks, so ~40–50 words is ideal.
- Key points examiners look for: (1) testing on new data, (2) measuring accuracy/performance, (3) improving the model.
- Avoid writing about other stages; stay focused on **Evaluation** only.
- The word "briefly" confirms a concise answer is expected — no need for bullet elaboration.

Q2. § Test Yourself § Case study-based questions **[1]**

In a medical screening test for a specific disease, which scenario represents a True Negative?

(A) A person without the disease tests positive for the disease.
(B) A person with the disease tests positive for the disease.
(C) A person with the disease tests negative for the disease.
(D) A person without the disease tests negative for the disease.

Previously asked in: 2025 104/S Q5 (v)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(D) A person without the disease tests negative for the disease.

Explanation

A **True Negative (TN)** means the test correctly identifies a healthy person as healthy — no disease present, result is negative. This is the "correct rejection." Contrast with False Positive (healthy person tests positive) and False Negative (sick person tests negative).

Q3. § Assertion and reasoning-based questions **[2]**

Define Confusion Matrix.

Previously asked in: 2023 104 Q14

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

A **confusion matrix** is a table used to evaluate the performance of a classification model. It presents **actual values on the y-axis** and **predicted values on the x-axis**. Each cell contains the number of predictions falling into that category — True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Source: Chapter 3, Section 3.4 – Evaluation Metrics for Classification

---

Explanation

- Examiners expect the definition to mention: (1) it is a table/matrix, (2) actual vs. predicted values, and (3) the four cells (TP, TN, FP, FN).
- Avoid writing long descriptions of each cell — just name them for a 2-mark answer.
- The phrase "handy presentation of accuracy of a model with two or more classes" from the textbook is a good addition if space allows.

Q4. § 3.1: Importance of Model Evaluation § 3.3: Accuracy and Error § Assertion and reasoning-based questions **[1]**

_____ is one of the parameter for evaluating a model's performance and is defined as the fraction of positive cases that are correctly identified.

(a) Precision
(b) Accuracy
(c) Recall
(d) F1

Previously asked in: 2024 104 Q3 (i)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(c) Recall — Recall is defined as the fraction of positive cases that are correctly identified by the model.

Explanation

Recall (also called sensitivity) = TP / (TP + FN). It measures how many actual positives were correctly caught. Precision focuses on correctness of positive predictions; Accuracy covers overall correct predictions; F1 balances precision and recall.

Q5. § Case study-based questions **[1]**

In spam email detection, which of the following will be considered as "False Negative" ?

(a) When a legitimate email is accurately identified as not spam.
(b) When a spam email is mistakenly identified as legitimate.
(c) When an email is accurately recognised as spam.
(d) When an email is inaccurately labelled as important.

Previously asked in: 2024 104 Q3 (v)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(b) When a spam email is mistakenly identified as legitimate.

False Negative (FN) means the model wrongly predicts the positive class as negative — i.e., a spam email (positive) is incorrectly classified as not spam (legitimate).

Explanation

- **False Negative = actual positive, predicted negative.** In spam detection, "spam" is the positive class. So FN occurs when a spam email slips through and is labelled as legitimate.
- Option (a) describes **True Negative**; option (c) describes **True Positive**; option (d) is not a standard confusion matrix term.
- Remember the pattern: *False* = wrong prediction; *Negative* = predicted as the negative class (not spam).

Q6. § Assertion and reasoning-based questions **[1]**

Statement 1 : Confusion matrix is an evaluation metric.
Statement 2 : Confusion Matrix is a record which helps in evaluation.

(a) Both Statement 1 and Statement 2 are correct.
(b) Both Statement 1 and Statement 2 are incorrect.
(c) Statement 1 is correct and Statement 2 is incorrect.
(d) Statement 2 is correct and Statement 1 is incorrect.

Previously asked in: 2024 104 Q4 (i)

◆ Evaluating Models | 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(a) Both Statement 1 and Statement 2 are correct.

A confusion matrix is an evaluation metric (Statement 1) and it is also a handy presentation/record that helps evaluate model accuracy (Statement 2).

Explanation

Both statements are correct because the confusion matrix serves dual purposes: it is listed as a popular classification metric, and it is described as "a handy presentation of the accuracy of a model" — i.e., a record that aids evaluation. Neither statement contradicts the source material.

Q7. § Assertion and reasoning-based questions **[1]**

Statement 1 : To evaluate a models' performance, we need either precision or recall.
Statement 2 : When the value of both Precision and Recall is 1, the F1 score is 0.

(a) Both statement 1 and statement 2 are correct.
(b) Both statement 1 and statement 2 are incorrect.
(c) Statement 1 is correct, but statement 2 is incorrect.
(d) Statement 1 is incorrect, but statement 2 is correct.

Previously asked in: 2024 104 Q5 (iii)

◆ Evaluating Models | 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(b) Both statement 1 and statement 2 are incorrect.

Statement 1 is incorrect because both precision **and** recall are needed (combined as F1 score) to fully evaluate a model. Statement 2 is incorrect because when Precision = 1 and Recall = 1, F1 = 2(1×1)/(1+1) = **1**, not 0.

Explanation

- F1 = 2 × (Precision × Recall) / (Precision + Recall). When both = 1, F1 = 1 (perfect score).
- Using only precision or only recall gives an incomplete picture; both are needed together for proper evaluation.
- Examiners expect you to know the F1 formula and apply it to verify such statements quickly.

Q8. § 3.4: Evaluation metrics for Classification § Assertion and reasoning-based questions § Case study-based questions **[1]**

In a fire alarm system, if the model predicts "Fire Present" when there is actually no fire, this is classified as :

(A) True Positive (TP)
(B) True Negative (TN)
(C) False Positive (FP)
(D) False Negative (FN)

Previously asked in: 2026 104 Q4 (i)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(C) False Positive (FP)

When the model predicts "Fire Present" (positive) but there is actually no fire (negative), it is a **False Positive** — wrongly predicting the negative class as positive.

Explanation

FP = model says "yes" but actual answer is "no." Remember: the first word (False) tells correctness, the second word (Positive) tells what the model predicted. Examiners expect the correct option plus a one-line justification.

Q9. § Assertion and reasoning-based questions **[1]**

Precision is defined as :

(A) The ratio of correctly predicted positive observations to total observations.
(B) The ratio of correctly predicted positive observations to total predicted positive observations.
(C) The ratio of correctly predicted negative observations to total observations.
(D) The harmonic mean of true positives and true negatives.

Previously asked in: 2026 104 Q4 (v)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(B) The ratio of correctly predicted positive observations to total predicted positive observations.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Explanation

Precision measures how many of the model's *positive predictions* were actually correct. Examiners expect the formula alongside the definition. Do not confuse with Recall (TP/TP+FN) or Accuracy (correct predictions/total predictions).

Q10. § 3.1: Importance of Model Evaluation § Assertion and reasoning-based questions **[1]**

It is one of the parameters for evaluating a model's performance and is defined as the percentage of true positive cases versus all the cases where the prediction is true. Which of the following evaluation parameter is this ?

(A) Precision
(B) Recall
(C) F1 Score
(D) Accuracy

Previously asked in: 2024 104 Q4 (iii)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(A) Precision

Precision is defined as the percentage of true positive cases out of all cases where the prediction is positive (predicted true).

Explanation

Precision = True Positives / (True Positives + False Positives). It measures how many of the *predicted positive* cases are actually positive. Recall, by contrast, measures true positives out of all *actual* positives. Students often confuse the two — remember: Precision is about the quality of positive predictions.

Q11. § Introduction § 3.1: Importance of Model Evaluation **[1]**

With reference to evaluation process of understanding the reliability of any AI model, define the term True Positive.

Previously asked in: 2022 104 Q11

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

True Positive (TP): When the model correctly predicts a positive outcome — i.e., the actual value is positive and the model also predicts it as positive.

Explanation

This is a 1-mark definition question from the Confusion Matrix topic (Chapter 3). Examiners expect the key idea: **correct prediction of a positive class**. Mention both conditions — actual is positive AND predicted is positive. Avoid confusing it with False Positive (actual negative, predicted positive).

Q12. § Unit-3: Evaluating Models – Lesson Overview § Case study-based questions **[1]**

What is F1 score?

Previously asked in: 2022 104 Q12

◆ **Evaluating Models** 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

F1 Score is the harmonic mean of Precision and Recall, used when the dataset is unbalanced and we cannot decide whether FP or FN is more important.

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Source: Chapter 3, Section 3.4 – F1 Score

Explanation

Examiners expect the definition **and** the formula for full marks. Mention that it combines both Precision and Recall into a single measure — that's the key idea. Avoid writing a long paragraph; one sentence + formula is sufficient for 1 mark.

Q13. § Unit-3: Evaluating Models – Lesson Overview **[4]**

> Traffic Jams have become a common part of our lives now-a-days. Living in an urban area means you have to face traffic each and every time you get out on the road. Mostly, school students opt for buses to go to school. Many times the bus gets late due to such jams and students are not able to reach their school on time. Thus, an AI model is created to predict explicitly if there would be a traffic jam on their way to school or not. The confusion matrix for the same is

Explain the process of calculating F1 score for the given problem.

Previously asked in: 2022 104 Q21

◆ **Evaluating Models** 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding stimulus

Model Answer

From the confusion matrix:

- TP = 50, FP = 50, FN = 0, TN = 0

Step 1 – Calculate Precision:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{50}{50 + 50} = \frac{50}{100} = 0.5$$

Step 2 – Calculate Recall:

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{50}{50 + 0} = \frac{50}{50} = 1$$

Step 3 – Calculate F1 Score:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.5 \times 1}{0.5 + 1} = \frac{1}{1.5} = 0.67$$

The F1 Score for this model is **0.67**.

Source: AI Chapter – Evaluating AI Models (Confusion Matrix & Performance Metrics)

Explanation

- Examiners expect **all three steps** to be shown: Precision → Recall → F1.
- Identify TP, FP, FN, TN correctly from the confusion matrix first — that is the foundation.
- Write the **formula** for each metric and substitute values clearly; step-by-step working earns full marks.
- F1 Score is the **harmonic mean** of Precision and Recall, used when there is a class imbalance — relevant here since TN = 0.

Q14. § 3.3: Accuracy and Error § Assertion and reasoning-based questions **[1]**

Which of the following is defined as the measure of balance between precision and recall?

(a) Accuracy
(b) F1 Score
(c) Reliability
(d) Punctuality

Previously asked in: 2023 104 Q2 (vi)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(b) F1 Score

Explanation

F1 Score is defined as the harmonic mean of precision and recall, combining both into a single metric. The source explicitly states: "F1-Score provides a way to combine both precision and recall into a single measure that captures both properties." Accuracy measures correct predictions overall; reliability and punctuality are not ML evaluation metrics.

Q15. § Assertion and reasoning-based questions § Case study-based questions **[1]**

While evaluating a model's performance, recall parameter considers _____.
(i) False positive
(ii) True positive
(iii) False negative
(iv) True negative
Choose the correct option:

(a) only (i)
(b) (ii) and (iii)
(c) (iii) and (iv)
(d) (i) and (iv)

Previously asked in: 2023 104 Q3 (v)

◆ Evaluating Models 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(b) (ii) and (iii)

Recall = TP / (TP + FN). It considers **True Positives (ii)** and **False Negatives (iii)**, measuring how well the model identifies all actual positives.

Explanation

Recall focuses on minimising missed positives (False Negatives). The formula TP/(TP+FN) uses only True Positives and False Negatives — not False Positives or True Negatives. Students often confuse recall with precision (which uses TP and FP), so remember: **Recall = catching all actual positives; Precision = correctness of positive predictions.**

Q16. § 3.3: Accuracy and Error § Reflection Time **[1]**

When a model is evaluated on the training data it always predicts correctly. This is known as _____.

Previously asked in: 2025 104/S Q4 (iii)

◆ Evaluating Models 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

When a model is evaluated on the training data, it simply remembers the whole training set and always predicts correctly. This is known as **overfitting**.

Explanation

The source passage explicitly states: *"our model will simply remember the whole training set, and will therefore always predict the correct label for any point in the training set. This is known as overfitting."* Examiners expect the exact term **overfitting** as the answer.

Q17. § Unit-3: Evaluating Models – Lesson Overview § 3.3: Accuracy and Error § Reflection Time **[1]**

Statement 1 : Overfitting occurs when a model memorizes the training data rather than learning patterns.
Statement 2 : Using the same data for training and evaluation helps the model give accurate results.

(A) Both statements are correct.
(B) Both statements are incorrect.
(C) Statement 1 is correct but statement 2 is incorrect.
(D) Statement 1 is incorrect but statement 2 is correct.

Previously asked in: 2026 104 Q2 (iv)

◆ Evaluating Models 3.1 Importance of Model Evaluation 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(C) Statement 1 is correct but Statement 2 is incorrect.

Statement 1 is correct as overfitting means the model memorizes training data. Statement 2 is incorrect because using the same data for training and evaluation leads to overfitting, not accurate results.

Explanation

The source passage explicitly states: *"It's not recommended to use the data we used to build the model to evaluate it... the model will simply remember the whole training set... This is known as overfitting."* So Statement 2 is clearly wrong. Statement 1 matches the textbook definition of overfitting directly.

Q18. § Introduction § 3.1: Importance of Model Evaluation **[1]**

What is the primary need for evaluating an AI model's performance in the AI Model Development process ?

(a) To increase the complexity of the model.
(b) To visualize the data.
(c) To assess how well the chosen model will work in future.
(d) To reduce the amount of data used for training.

Previously asked in: 2024 104 Q4 (vi)

◆ Evaluating Models 3.1 Importance of Model Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(c) To assess how well the chosen model will work in future.

Model evaluation helps find the best model that represents our data and how well the chosen model will work in the future.

Source: Chapter 3, Introduction & Section 3.1

Explanation

The passage explicitly states: *"Model Evaluation... helps to find the best model that represents our data and* ***how well the chosen model will work in the future****."* This directly matches option (c). For MCQs, always quote or paraphrase the source to show reasoning, but keep it to one line.

Q19. § 3.2: Splitting the training set data for Evaluation **[1]**

In supervised learning, what is the purpose of the testing dataset ?

(A) To train the model.
(B) To evaluate the model's accuracy.
(C) To create new features.
(D) To label the data.

Previously asked in: 2026 104 Q3 (vi)

◆ Evaluating Models 3.1 Importance of Model Evaluation 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(B) To evaluate the model's accuracy.

The testing dataset is used to evaluate the performance of a trained model by comparing its predictions against known expected values.

Source: Chapter 3, Section 3.2 (Train-test split)

Explanation

The key idea from the passage is that the test dataset's input is fed to the trained model, predictions are made, and those predictions are compared to expected values — this is evaluation. The training dataset is what teaches/trains the model (ruling out option A). Options C and D are unrelated to the purpose of a test split.

Q20. § 3.1: Importance of Model Evaluation **[2]**

What do you mean by Evaluation of an AI model ? Also explain the concept of overfitting with respect to AI model Evaluation.

Previously asked in: 2024 104 Q13

◆ Evaluating Models | 3.1 Importance of Model Evaluation | 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

Evaluation of an AI Model: Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance. It acts like a report card, helping identify a model's strengths, weaknesses, and suitability for a task, ensuring it is trustworthy and reliable.

Overfitting: Overfitting occurs when a model performs very well on training data but poorly on new/test data. It means the model has memorised the training data rather than learning general patterns, making it unreliable for real-world use.

Source: Chapter 3 – Model Evaluation, Section 3.1

Explanation

- Evaluation = metrics-based performance check (think: report card analogy — examiners love this).
- Overfitting is a key evaluation concept: the model is "too fitted" to training data. Contrast with training accuracy vs. test accuracy to score full marks.
- Even though overfitting isn't explicitly detailed in the given passages, it is a standard concept tied to train-test split in this chapter — always link it to poor performance on unseen data.

Q21. § 3.2: Splitting the training set data for Evaluation **[2]**

Explain Train-test split technique with respect to machine learning algorithm.

Previously asked in: 2026 104 Q15

◆ Evaluating Models | 3.1 Importance of Model Evaluation | 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

Train-test split is a technique used to evaluate the performance of a machine learning algorithm. It divides the dataset into two subsets:

- **Training dataset** – used to train/teach the model.
- **Testing dataset** – input is given to the trained model, predictions are made, and compared with expected values.

This technique is suitable when a sufficiently large dataset is available.

Source: Chapter 3, Section 3.2

Explanation

Examiners expect two things: (1) a clear definition stating it is an *evaluation technique*, and (2) the role of each subset (train = learning, test = comparing predictions vs actual). Mentioning that it avoids overfitting is a bonus but not mandatory for 2 marks. Keep it concise.

Q22. § 3.3: Accuracy and Error | § Reflection Time | § Assertion and reasoning-based questions **[1]**

Statement 1 : Overfitting is not recommended for evaluation of a model.
Statement 2 : This is because the model will simply remember the whole training set, and will therefore always predict the correct label for any point in the training set.
(A) Both Statement 1 and Statement 2 are correct.
(B) Both Statement 1 and Statement 2 are incorrect.
(C) Statement 1 is correct but Statement 2 is incorrect.
(D) Statement 2 is correct but Statement 1 is incorrect.

(A) Both Statement 1 and Statement 2 are correct.
(B) Both Statement 1 and Statement 2 are incorrect.
(C) Statement 1 is correct but Statement 2 is incorrect.
(D) Statement 2 is correct but Statement 1 is incorrect.

Previously asked in: 2024 104 Q4 (ii)

◆ Evaluating Models | 3.1 Importance of Model Evaluation | 3.2 Splitting the Training Set Data for Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(A) Both Statement 1 and Statement 2 are correct.

Explanation

Overfitting means the model memorises training data, so it performs well on training data but poorly on unseen data — making it unsuitable for model evaluation. Both statements correctly describe this problem.

Q23. § Unit-3: Evaluating Models – Lesson Overview | § Introduction | § Assertion and reasoning-based questions **[1]**

_____ helps to find the best model that represents our data and how well the chosen model will work in future.

Previously asked in: 2023 104 Q3 (iv)

◆ Evaluating Models | 3.1 Importance of Model Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

Model Evaluation helps to find the best model that represents our data and how well the chosen model will work in the future.

Source: Chapter 3, Section 3.1 – Importance of Model Evaluation

---

Explanation

The exact phrase comes directly from the introduction of Chapter 3. Examiners expect the term **"Model Evaluation"** as the answer. Remember: model evaluation acts like a report card for your AI model, measuring its performance using metrics like accuracy, precision, and recall.

Q24. § Assertion and reasoning-based questions § Case study-based questions **[4]**

> Confusion Matrix for an AI model that predicts whether there is a chance of earthquake or not.

Recently the country was shaken up by a series of earthquakes which has done a huge damage to the people as well as the infrastructure. To address this issue, an AI model has been created which can predict if there is a chance of earthquake or not. The confusion matrix for the same is given. Answer the following:

(i) How many total cases are True Negative in the above scenario?

(ii) Calculate precision, recall and F1 score.

Previously asked in: 2023 104 Q21

◆ **Evaluating Models** 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding stimulus

Model Answer

(i) True Negative (TN):

True Negative means the model predicted **No earthquake**, and in reality there was **No earthquake**.
From the table: **TN = 20**

(ii) Calculations:

From the confusion matrix:

- TP = 50, FP = 5, FN = 25, TN = 20
- Total = 100

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{50}{50 + 5} = \frac{50}{55} \approx 0.909$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{50}{50 + 25} = \frac{50}{75} \approx 0.667$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.909 \times 0.667}{0.909 + 0.667} \approx \frac{1.212}{1.576} \approx \mathbf{0.769}$$

Source: Confusion Matrix (AI), Class 10 IT/CS

Explanation

- **TN** is always the "Predicted No & Reality No" cell — easy 1 mark.
- For part (ii), correctly identify all four values (TP, FP, FN, TN) first — examiners award method marks for this.
- Show the formula, substitution, and final value for each metric to secure full marks.
- F1 Score is the harmonic mean of Precision and Recall — remember the formula exactly.

Q25. § Test Yourself § Assertion and reasoning-based questions **[2]**

Draw the confusion matrix for the following data :

(a) The number of true positive = 200

(b) The number of true negative = 50

(c) The number of false positive = 35

(d) The number of false negative = 55

Previously asked in: 2025 104/S Q15

◆ **Evaluating Models** 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

Confusion Matrix:

	Predicted Positive	**Predicted Negative**
Actual Positive	TP = 200	FN = 55
Actual Negative	FP = 35	TN = 50

- Total predictions = 200 + 50 + 35 + 55 = **340**

Source: Evaluation metrics for Classification, Chapter 3

Explanation

- The confusion matrix is a 2×2 table with **Actual values on rows** and **Predicted values on columns**.
- Place TP and TN on the diagonal (correct predictions); FP and FN off-diagonal (wrong predictions).
- Examiners expect the correct placement: TP (top-left), FN (top-right), FP (bottom-left), TN (bottom-right).
- Mentioning total is optional but shows understanding.

Q26. § Case study-based questions **[4]**

> Confusion Matrix showing the results of a sentiment analysis model tested on 500 movie reviews.

A sentiment analysis model was built to classify movie reviews as either Positive or Negative. The model was tested on a dataset of 500 reviews, resulting in the following confusion matrix :

(a) How many total cases are True Positive in the above scenario? [1]

(b) Calculate Precision, Recall and F1-Score. [3]

Previously asked in: 2025 104/S Q21

◆ **Evaluating Models** | 3.3 Accuracy and Error | 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding stimulus

Model Answer

(a) True Positives (TP) = **300**
(Cases where the model correctly predicted a review as Positive, and it actually was Positive.)

(b)

Given: TP = 300, FP = 60, FN = 40, TN = 100

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{300}{300 + 60} = \frac{300}{360} \approx 0.833$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{300}{300 + 40} = \frac{300}{340} \approx 0.882$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.833 \times 0.882}{0.833 + 0.882} = \frac{1.469}{1.715} \approx \mathbf{0.857}$$

Explanation

- **Part (a):** TP is simply read directly from the confusion matrix cell (Reality Yes, Predicted Yes).
- **Part (b):** Remember the three formulas — Precision measures how many predicted positives are correct; Recall measures how many actual positives were caught; F1-Score balances both. Always substitute values clearly step-by-step for full marks.

Q27. § 3.3: Accuracy and Error § Assertion and reasoning-based questions **[1]**

In the context of autonomous vehicle safety systems, which type of error would be most critical to minimize ?

(A) False Positive (detecting danger when there isn't any)
(B) False Negative (failing to detect actual danger)
(C) True Positive (detecting danger correctly)
(D) True Negative (correctly identifying that there is no danger)

Previously asked in: 2026 104 Q3 (i)

◆ Evaluating Models 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(B) False Negative (failing to detect actual danger)

In safety-critical systems, missing actual danger (False Negative) is most critical to minimize, as undetected hazards can cause accidents or loss of life.

Explanation

- The source passage highlights that in medical diagnosis, avoiding missed detections (false negatives) is prioritised over overall accuracy — the same principle applies to autonomous vehicle safety.
- **False Negative** = model predicts "no danger" when danger actually exists → catastrophic consequence.
- **False Positive** = model detects danger when there isn't any → inconvenient but not life-threatening (vehicle brakes unnecessarily).
- Examiners expect you to link the concept of minimising error to the real-world consequence of the error type, not just define the term.

Source: Chapter 3, Section 3.3 — Accuracy and Error

Q28. § 3.3: Accuracy and Error **[1]**

State True or False :
In machine learning, the error is used to see how accurately the model can predict data.

Previously asked in: 2026 104 Q5 (iv)

◆ Evaluating Models 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

True.

In machine learning, the error is used to see how accurately the model can predict data it uses to learn new, unseen data.

Explanation

The passage under the **Error** section explicitly states: *"In Machine Learning, the error is used to see how accurately our model can predict data it uses to learn new, unseen data."* So the statement is directly supported by the source. Just quote or paraphrase this line for full marks.

Q29. § Test Yourself § Reflection Time § Assertion and reasoning-based questions **[1]**

An AI model was tested with 1000 test samples. If True Positive (TP) = 200, True Negative (TN) = 600, False Positive (FP) = 100, False Negative (FN) = 100, how many total predictions were correct ?

(A) 300
(B) 600
(C) 800
(D) 900

Previously asked in: 2026 104 Q5 (vi)

◆ **Evaluating Models** 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(C) 800

Total correct predictions = TP + TN = 200 + 600 = **800**

Explanation

Correct predictions include True Positives (correctly predicted positive) and True Negatives (correctly predicted negative). FP and FN are incorrect predictions and are not counted. Always use: Correct = TP + TN.

Q30. § Case study-based questions **[4]**

> Confusion Matrix for a binary classification model classifying news articles as "Fake News" or "Real News", tested on 500 news articles. The confusion matrix cross-tabulates Predicted vs Reality.

A binary classification model has been developed to classify news articles as either "Fake News" or "Real News". The model was tested on a dataset of 500 news articles, and the resulting confusion matrix is as follows :

(A) How many total cases are True Negative in the above scenario ? [1]

(B) Calculate Precision, Recall and F1-Score. [3]

Previously asked in: 2024 104 Q21

◆ **Evaluating Models** | 3.3 Accuracy and Error | 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding stimulus

Model Answer

(A) True Negatives (TN):

True Negative = cases where the model predicted **No** (Real News) and the reality was also **No** (Real News).

TN = 420

(B) Precision, Recall and F1-Score:

From the matrix:

- TP = 45, FP = 15, FN = 20, TN = 420

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{45}{45 + 15} = \frac{45}{60} = \mathbf{0.75}$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{45}{45 + 20} = \frac{45}{65} \approx \mathbf{0.692}$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.75 \times 0.692}{0.75 + 0.692} = \frac{1.038}{1.442} \approx \mathbf{0.72}$$

Explanation

- **TN** = correctly predicted negatives (Real News predicted as Real News) → 420.
- **Precision** measures how many predicted "Fake" were actually fake (avoids false alarms).
- **Recall** measures how many actual "Fake" were correctly caught (avoids misses).
- **F1-Score** is the harmonic mean of Precision and Recall — useful when classes are imbalanced.
- Always identify TP, FP, FN, TN first before computing metrics; examiners award 1 mark per formula+substitution+result.

Q31. § Test Yourself § Assertion and reasoning-based questions § Case study-based questions **[4]**

> PQR Security Solutions has designed an AI Model to detect cyber attacks on E-Commerce websites. For this, various network activities were monitored and analyzed on one of the websites. The model was tested on a dataset of 1500 network activities. Out of these, the model correctly predicted that 1000 were cyber attacks. It also correctly identified that 250 were not cyber attacks. However, the model predicted that 200 were cyber attacks but actually they were not. Additionally, it predicted that 50 were not cyber attacks but they actually were.

Read the following paragraph and answer the questions that follow:

(A) Draw the confusion matrix based on the given scenario. [2]

(B) How many total cases are True Negative in the above scenario ? [1]

(C) Calculate Precision. [1]

Previously asked in: 2026 104 Q19

◆ **Evaluating Models** 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding stimulus

Model Answer

(A) Confusion Matrix:

	Predicted: Attack	**Predicted: Not Attack**
Actual: Attack	1000 (TP)	50 (FN)
Actual: Not Attack	200 (FP)	250 (TN)

(B) True Negative (TN):

True Negative = cases correctly identified as **not** cyber attacks = **250**

(C) Precision:

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{1000}{1000 + 200} = \frac{1000}{1200} \approx 0.833 \text{ (or } 83.3\%)$$

Explanation

- **TP** (True Positive): correctly predicted attacks → 1000
- **TN** (True Negative): correctly predicted non-attacks → 250
- **FP** (False Positive): predicted attack, actually not → 200
- **FN** (False Negative): predicted not attack, actually was → 50
- **Precision** measures "of all predicted positives, how many were actually positive" — examiners expect the formula and substitution shown clearly.

Q32. § Assertion and reasoning-based questions **[1]**

Which condition of evaluation does the following diagram indicate ?

(A) False Positive
(B) False Negative
(C) True Positive
(D) True Negative

Previously asked in: 2024 104 Q3 (vi)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(A) False Positive

A False Positive occurs when the model predicts a positive outcome, but the actual/true value is negative (incorrect positive prediction).

Explanation

Since the diagram is not fully visible, the standard exam approach is: **False Positive (FP)** is the most commonly tested condition in MCQs of this type, where the model predicts "Yes/Positive" but the actual label is "No/Negative." Examiners expect students to know all four confusion matrix conditions — TP, TN, FP, FN — and identify them from diagrams showing predicted vs. actual labels.

Q33. § 3.1: Importance of Model Evaluation | § 3.3: Accuracy and Error | § Assertion and reasoning-based questions **[1]**

With respect to evaluation, for which of the following does the prediction and reality match ?

(A) True positive and False positive
(B) True positive and True negative
(C) False positive and False negative
(D) True positive and False negative

Previously asked in: 2024 104 Q5 (ii)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(B) True positive and True negative

In both cases, the model's prediction matches reality — correctly predicting positive as positive, and negative as negative.

Explanation

"True" in True Positive and True Negative means the prediction **matches** the actual outcome. "False" (as in False Positive/Negative) means there is a **mismatch**. Examiners expect you to know these definitions from the confusion matrix topic.

Q34. § Assertion and reasoning-based questions § Case study-based questions **[1]**

Which of the following scenarios might have a high False Negative (FN) cost ?

(A) Viral Disease Outbreak
(B) Spam
(C) Mining
(D) Image Search

Previously asked in: 2024 104 Q5 (v)

◆ Evaluating Models 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(A) Viral Disease Outbreak

A viral disease outbreak has a high FN cost because missing an infected person (predicting negative when actually positive) can lead to undetected spread of disease, causing serious public health consequences.

Explanation

- **False Negative (FN)** = model predicts "No Disease/No Risk" but the actual case IS positive.
- In a viral outbreak, failing to identify an infected person is extremely dangerous (they spread the disease further), so FN cost is very high — **Recall** matters most here.
- Spam (FP cost is higher), Mining and Image Search are not life-critical, so FN cost is comparatively low.
- Examiners expect you to correctly identify the option and give a one-line justification.

Q35. § Introduction § 3.3: Accuracy and Error § Reflection Time **[2]**

With reference to evaluation stage of AI project cycle, explain the term Accuracy. Also give the formula to calculate it.

Previously asked in: 2022 104 Q17

◆ Evaluating Models | 3.1 Importance of Model Evaluation | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

Accuracy is an evaluation metric that measures the total number of predictions a model gets right. The accuracy of the model and its performance are directly proportional — the better the performance, the more accurate the predictions.

Formula:

$$\text{Accuracy} = 1 - \text{Error Rate}$$

where $\text{Error Rate} = \frac{\text{Error (Actual - Predicted)}}{\text{Actual Value}}$

Accuracy% = Accuracy × 100%

Source: Chapter 3, Section 3.3 — Accuracy and Error

Explanation

- The examiner expects **two things**: a definition of accuracy AND the formula — both are needed for full 2 marks.
- The formula is presented in the textbook's activity table, so use the step-by-step form: Error → Error Rate → Accuracy = 1 – Error Rate.
- Do **not** confuse this regression-style accuracy formula with classification accuracy (correct predictions / total predictions) — stick to what the source passage shows.

Q36. § Case study-based questions **[2]**

Suppose you are developing an AI model to detect fraudulent financial transaction risk. Describe False Positives and False Negatives in this context.

Previously asked in: 2024 104 Q16

◆ Evaluating Models | 3.5 Ethical Concerns Around Model Evaluation | False Positives and False Negatives in Model Evaluation

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

False Positive (FP): A legitimate transaction is incorrectly flagged as fraudulent. This causes inconvenience to genuine customers whose valid transactions are blocked unnecessarily.

False Negative (FN): A fraudulent transaction is incorrectly classified as legitimate. This is more dangerous as actual fraud goes undetected, causing financial loss.

Explanation

The examiner expects you to clearly define both terms **in context** (fraud detection), not in general. Always mention the real-world consequence — for fraud detection, FN is costlier because actual fraud is missed. This shows conceptual understanding, which CBSE rewards. Remember: FP = wrongly accused legitimate; FN = missed actual fraud.

Q37. § Test Yourself § Assertion and reasoning-based questions § Case study-based questions **[4]**

> Reality vs Prediction Confusion Matrix:
> - Prediction Yes / Reality Yes: 110 (True Positive)
> - Prediction Yes / Reality No: 60 (False Positive)
> - Prediction No / Reality Yes: 50 (False Negative)
> - Prediction No / Reality No: 410 (True Negative)
> Total: 630 tests

An AI model has been developed to test specimens of blood/urine/cough etc. to diagnose ailments (diabetes/liver infection etc.). The model was tested on a data-set of about 630 tests and the resulting confusion matrix is as follows :

(A) How many total cases are True Negative in the above scenario ? [1]

(B) Calculate Precision, Recall and F1 Score. [3]

Previously asked in: 2024 104 Q21

◆ **Evaluating Models** 3.3 Accuracy and Error 3.4 Evaluation Metrics for Classification

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding stimulus

Model Answer

(A)

True Negative (TN) = **410**

These are cases where the model predicted **No** and the reality was also **No**.

(B)

Given: TP = 110, FP = 60, FN = 50, TN = 410

Precision = TP / (TP + FP) = 110 / (110 + 60) = 110 / 170 ≈ **0.647**

Recall = TP / (TP + FN) = 110 / (110 + 50) = 110 / 160 ≈ **0.6875**

F1 Score = 2 × (Precision × Recall) / (Precision + Recall)
= 2 × (0.647 × 0.6875) / (0.647 + 0.6875)
= 2 × 0.4448 / 1.3345
≈ **0.667**

Source: Confusion Matrix & Model Evaluation, AI Chapter

Explanation

- **TN** is straightforward — read directly from the matrix (Prediction No / Reality No).
- For **Precision**, use only the "Prediction Yes" row (TP + FP in denominator).
- For **Recall**, use only the "Reality Yes" column (TP + FN in denominator).
- **F1 Score** is the harmonic mean of Precision and Recall — always show the formula first, then substitute values. Examiners award marks for correct formula + correct substitution + correct answer.

Q38. § Assertion and reasoning-based questions **[1]**

Two conditions when prediction matches with the reality are true positive and _____.

Previously asked in: 2023 104 Q4 (ii)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

Two conditions when prediction matches with the reality are **True Positive** and **True Negative**.

Explanation

The confusion matrix has four outcomes: TP, TN, FP, FN. Only TP (predicted positive, actually positive) and TN (predicted negative, actually negative) are cases where prediction matches reality. FP and FN are errors.

Q39. § 3.3: Accuracy and Error | § Assertion and reasoning-based questions **[1]**

_____ is defined as the percentage of correct predictions out of all the observations.

(A) Precision
(B) Accuracy
(C) Recall
(D) F1

Previously asked in: 2025 104/S Q3 (ii)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(B) Accuracy

Accuracy is defined as the percentage of correct predictions out of all the observations.

Explanation

The question directly defines **Accuracy** — the ratio of correct predictions to total predictions, expressed as a percentage. Precision = TP/(TP+FP); Recall = TP/(TP+FN); F1 combines both. Don't confuse them.

Q40. § 3.3: Accuracy and Error | § 3.4: Evaluation metrics for Classification | § Assertion and reasoning-based questions **[1]**

When the prediction matches the reality, the condition is termed as _____.

Previously asked in: 2023 104 Q5 (ii)

◆ Evaluating Models | True Positive / Correct Prediction (Prediction matches Reality)

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

When the prediction matches the reality, the condition is termed as **True Positive (TP)** (or a **correct prediction** / **accurate prediction**).

Source: Chapter 3, Section 3.3 – Accuracy and Error; Section 3.4 – Confusion Matrix

---

Explanation

The passage defines True Positive as "the outcome of the model correctly predicting the positive class," meaning prediction matches reality. More broadly, the chapter links this to **accuracy** — when predictions match actual values, they are counted as correct (TP or TN). Examiners accept either "True Positive" (if context is confusion matrix) or "accurate/correct prediction" (if context is general accuracy). Write whichever fits the context of your question paper.

Q41. § Assertion and reasoning-based questions | § Case study-based questions **[1]**

In a quality control system for manufacturing, which scenario represents a false negative?

(A) When a defective product is correctly identified as defective.
(B) When a non-defective product is inaccurately identified as defective.
(C) When a non-defective product is correctly identified as non-defective.
(D) When a defective product is mistakenly identified as non-defective.

Previously asked in: 2025 104/S Q3 (v)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

(D) When a defective product is mistakenly identified as non-defective.

A false negative means the model fails to detect a positive (defective) case, incorrectly labelling it as negative (non-defective).

Explanation

- **False Negative (FN):** The actual condition is positive (defective), but the model predicts negative (non-defective) — the defect is *missed*.
- Option (A) = True Positive; Option (B) = False Positive; Option (C) = True Negative.
- Remember the key rule: "False Negative = missed detection of a real positive case." This concept is directly linked to Recall/Sensitivity, which penalises false negatives.

Q42. § Assertion and reasoning-based questions **[1]**

Which of the following talks about how true the predictions are by any model?

(a) Accuracy
(b) Reliability
(c) Recall
(d) F1 score

Previously asked in: 2023 104 Q5 (vi)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:18 · grounding rag

Model Answer

(a) Accuracy

Accuracy measures how true (correct) the predictions made by a model are — i.e., the total number of predictions a model gets right.

Explanation

The passage explicitly states: *"Accuracy is an evaluation metric that allows you to measure the total number of predictions a model gets right."* Examiners expect you to recall this definition directly. Reliability is not a standard ML metric; Recall measures how many actual positives were found; F1 score balances precision and recall — none of these describe "how true" predictions are.

Q43. § Assertion and reasoning-based questions **[1]**

Which of the following is not true about Confusion Matrix?

(A) It allows us to understand prediction results.
(B) It is a Model Training Matrix.
(C) It helps in evaluation of machine learning models.
(D) It is used to record comparison between prediction and reality.

Previously asked in: 2025 104/S Q4 (v)

◆ Evaluating Models | 3.3 Accuracy and Error

Generated by claude-sonnet-4-6 · 2026-06-21 03:19 · grounding rag

Model Answer

The correct answer is **(B) It is a Model Training Matrix.**

A Confusion Matrix is used for **model evaluation**, not model training. It records comparisons between predictions and actual values, helping us understand and evaluate model results.

Explanation

The examiner expects you to identify the **false statement** about a Confusion Matrix. Options A, C, and D are all true — the Confusion Matrix helps understand predictions, evaluates ML models, and compares prediction vs. reality. Option B is false because a Confusion Matrix is an **evaluation tool**, not a training tool. Remember this distinction clearly.

Available for free from:
https://cbsegrade10studyguide.com
https://github.com/orgs/cbse-free-resources/repositories