INDIAN SCHOOL AL WADI AL KABIR

| Woksheet No:1 | **ARTIFICIAL INTELLIGENCE (417) CLASS X** |
| | **CHAPTER 7: EVALUATION** |

1 Mark questions

1. Rajat has made a model which predicts the performance of Indian Cricket players in upcoming matches. He collected the data of players' performance with respect to stadium, bowlers, opponent team and health. His model works with good accuracy and precision value. Which of the statement given below is incorrect?
   (a) Data gathered with respect to stadium, bowlers, opponent team and health is known as Testing Data.
   (b) Data given to an AI model to check accuracy and precision is Testing Data.
   (c) Training data and testing data are acquired in the Data Acquisition stage.
   (d) Training data is always larger as compared to testing data.

2. Which of the following statement(s) is/are correct? (1)
   (i) Data collection helps us in determining the nature of the problem.
   (ii) Modelling is the process of implementing AI-enabled algorithms on a chosen model.
   (iii) Evaluation is the process of putting our model to test by using AI-enabled algorithms to generate accurate real time results.
   (iv) Data acquisition ensures that data obtained from many sources is accurate, dependable and traceable.
   (a) (i) and (ii) only   (b) (i), (ii) and (iii)   (c) (ii) and (iii) only   **(d) (ii), (iii) and (iv)**

3. Raunak was learning the conditions that make up the confusion matrix. He came across a scenario in which the machine that was supposed to predict an animal was always predicting not an animal. What is this condition called?
   (a) False Positive   (b) True Positive   **(c) False Negative**   (d) True Negative

4. Which of the following statements is not true about overfitting models?
   (a) This model learns the pattern and noise in the data to such extent that it harms the performance of the model on the new dataset.
   (b) Training result is very good and the test result is poor.
   (c) It interprets noise as patterns in the data.
   (d) The training accuracy and test accuracy both are low

5. _____is defined as the percentage of correct predictions out of all the observations.
   a) Predictions   **b) Accuracy**   c) Reality   d) F1 Score

6. What will be the outcome, if the Prediction is "Yes" and itmatches with the Reality? What will bethe outcome, if the Prediction is "Yes" and it does not match the Reality?
   a) True Positive, True Negative   b) True Negative, False Negative
   c) True Negative, False Positive   **d) True Positive, False Positive.**

7. Which two evaluation methods are used to calculate F1 Score?
   (a) Precision and Accuracy   **(b) Precision and Recall**   (c) Accuracy and Recall   (d) Precision, F1 score

8. Recall-Evaluation method is
   a) defined as the fraction of positive cases that are correctly identified.
   b) defined as the percentage of true positive cases versus all the cases where the prediction is true.
   c) defined as the percentage of correct predictions out of all the observations.
   d) comparison between the prediction and reality.

9. Which evaluation parameter takes into consideration all the correct predictions?
   Ans. Accuracy

10. The output given by the AI machine is known as_____.
    a. **prediction**   b. reality   c. hypothesis   d. statement

2 Mark questions

11. Define Evaluation.
Moving towards deploying the model in the real world, we test it in as many ways as possible. The stage of testing the models is known as EVALUATION.
OR
Evaluation is a process of understanding the reliability of any AI model, based on outputs by feeding the test dataset into the model and comparing it with actual answers.
OR
Evaluation is a process that critically examines a program. It involves collecting and analyzing information about a program's activities, characteristics, and outcomes. Its purpose is to make judgments about a program, to improve its effectiveness, and/or to inform programming decisions.

12. Which two parameters are considered for Evaluation of a model?
Prediction and Reality are the two parameters considered for Evaluation of a model.
The "Prediction" is the output which is given by the machine and the "Reality" is the real scenario, when the prediction has been made.

13. What is True Positive?
The predicted value matches the actual value
he actual value was positive and the model predicted a positive value

14. What is True Negative?
The predicted value matches the actual value
The actual value was negative and the model predicted a negative value

15. What is False Positive?
The predicted value was falsely predicted
The actual value was negative but the model predicted a positive value
Also known as the Type 1 error

16. What is False Negative?
The predicted value was falsely predicted
The actual value was positive but the model predicted a negative value
Also known as the Type 2 error

17. What is meant by Overfitting of Data?

Overfitting is "the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably".

(OR)

An Overfitted Model is a statistical model that contains more parameters than can be justified by the data. Here, to evaluate the AI model it is not necessary to use the data that is used to build the model. Because AI Model remembers the whole training data set, therefore it always predicts the correct label for any point in the training dataset. This is known as Overfitting

(OR)

Models that use the training dataset during testing, will always results in correct output. This is known as overfitting.

18. What is Accuracy? Mention its formula.

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

A prediction is said to be correct if it matches reality. Here we have two conditions in which the Prediction matches with the Reality, i.e., True Positive and True Negative. Therefore, Formula for Accuracy is

$$Accuracy = \frac{Correct\ prediction}{Total\ cases} * 100\%$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

19. Why is evaluation important? Explain.

Importance of Evaluation

Evaluation is a process that critically examines a program. It involves collecting and analyzing information about a program's activities, characteristics, and outcomes. Its purpose is to make judgments about a program, to improve its effectiveness, and/or to inform programming decisions.

- Evaluation is important to ensure that the model is operating correctly and optimally.
- Evaluation is an initiative to understand how well it achieves its goals.
- Evaluations help to determine what works well and what could be improved in a program

20. How do you suggest which evaluation metric is more important for any case?

F 1 Evaluation metric is more important in any case. F1 score sort maintains a balance between the precision and recall for the classifier. If the precision is low, the F1 is low and if the recall is low again F1 score is low.

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

The F1 score is a number between 0 and 1 and is the harmonic mean of precision and recall.

When we have a value of 1 (that is 100%) for both Precision and Recall. The F1 score would also be an ideal 1 (100%). It is known as the perfect value for F1 Score. As the values of both Precision and Recall ranges from 0 to 1, the F1 score also ranges from 0 to 1.

21. Which evaluation metric would be crucial in the following cases? Justify your answer.
a. Mail Spamming
b. Gold Mining
c. Viral Outbreak
Here, Mail Spamming and Gold Mining are related to FALSE POSITIVE cases which are expensive at cost. But Viral Outbreak is a FALSE NEGATIVE case which infects a lot of people on health and leads to expenditure of money too for check-ups. So, False Negative case (VIRAL OUTBREAK) are more crucial and dangerous when compared to FALSE POSITIVE cases.

22. Calculate Accuracy, Precision, Recall and F1 Score for the following Confusion Matrix on Heart Attack Risk. Also suggest which metric would not be a good evaluation parameter here and why

The Confusion Matrix	Reality: 1	Reality: 0
Prediction: 1	50	20
Prediction: 0	10	20

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

$$Accuracy = \frac{Correct\ prediction}{Total\ cases} * 100\%$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$$

Where True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).
Accuracy = (50+20) / (50+20+20+10)
= (70/100)
= 0.7

Precision:
Precision is defined as the percentage of true positive cases versus all the cases where the prediction is true

$$Precision = \frac{True\ Positive}{All\ Predicted\ Positives} * 100\%$$

$$Precision = \frac{TP}{TP + FP} * 100\%$$

= (50 / (50 + 20))
= (50/70)
= 0.714

Recall: It is defined as the fraction of positive cases that are correctly identified.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$Recall = \frac{TP}{TP + FN}$$

= 50 / (50 + 60)
= 50 / 110
= 0.5

F1 Score:
F1 score is defined as the measure of balance between precision and recall.

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

= 2 * (0.714 *0.5) / (0.714 + 0.5)
= 2 * (0.357 / 1.214)
= 2* (0.29406)
= 0.58
Therefore,
Accuracy= 0.7 Precision=0.714 Recall=0.5
F1 Score=0.588
Here within the test there is a trade-off. But Recall is not a good Evaluation metric. Recall metric needs to improve more.
Because, False Positive (impacts Precision): A person is predicted as high risk but does not have heart attack. False Negative (impacts Recall): A person is predicted as low risk but has heart attack. Therefore, False Negatives miss actual heart patients, hence recall metric need more improvement. False Negatives are more dangerous than False Positives.
 ****