

# Quizz : Machine Learning - Setting Up the Scenes

Please answer the following questions.

\* Indicates required question

---

1. Nom \*

---

2. Adresse e-mail \*

---

3. Given a dataset of financial transactions where each transaction is labeled as 'fraudulent' or 'legitimate', which machine learning approach would be most appropriate, and why? 1 point

*Mark only one oval.*

- ☐ Supervised learning, because the dataset includes labeled outcomes that the model can learn from to predict future transactions.
- ☐ Unsupervised learning, to allow the model to identify hidden patterns in the data without relying on pre-labeled outcomes.
- ☐ Reinforcement learning, as it enables the model to learn optimal actions through trial and error with the environment.
- ☐ Semi-supervised learning, since financial transactions are too complex for fully supervised methods.

4. In the context of machine learning project lifecycle, what is the correct order of data handling steps ? \* 1 point

*Mark only one oval.*

- ☐ Train -> Test -> Preprocessing -> Validation
- ☐ Preprocessing -> Train -> Validation -> Test
- ☐ Test -> Preprocessing -> Train -> Validation
- ☐ Preprocessing -> Validation -> Train -> Test

5. Which metric is most appropriate for evaluating a machine learning model used for fraud detection in financial transactions, where false negatives are more costly than false positives? \* 1 point

*Mark only one oval.*

- ☐ Accuracy
- ☐ F1 Score
- ☐ Recall

6. When comparing two machine learning models in the context of financial fraud detection using the AUC of their ROC curves, which of the following statements are true? (Select all that apply) \* 2 points

*Tick all that apply.*

- ☐ A model with a higher AUC value is generally considered better at distinguishing between fraudulent and legitimate transactions
- ☐ If two models have identical AUC values, they are equally good, regardless of the distribution of false positives and negatives.
- ☐ Comparing models using AUC values alone is sufficient, without considering the context of their application or the cost of misclassifications
- ☐ It's important to also consider other metrics like precision and recall, especially in the context of imbalanced datasets, even if one model has a slightly higher AUC.
- ☐ AUC can be a useful metric for comparison, but understanding the models' operational thresholds is crucial for real-world applications

7. Did you understand the Programming Session ?

*Tick all that apply.*

☐ Yes

☐ No

8. Do you have any suggestions to improve the Programming Session ? \*

---

---

---

---

---

---

This content is neither created nor endorsed by Google.

Google Forms