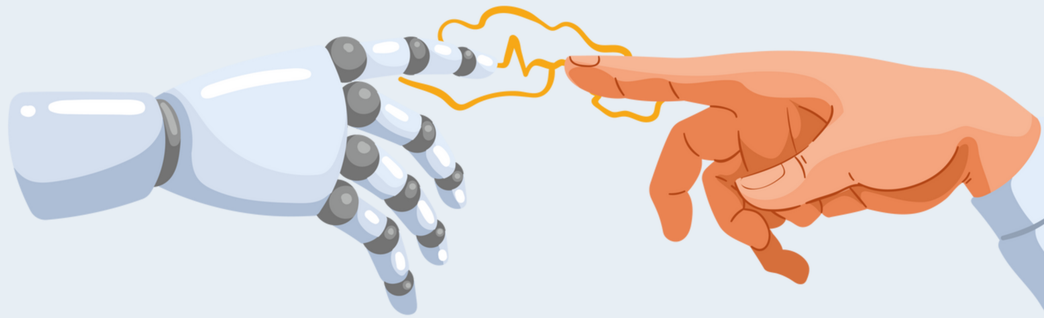
AI Day 2 Recap



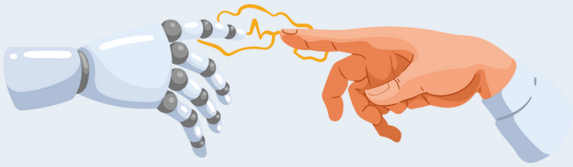
Choose the correct answer about Classification & Regression

a. Regressions is continuous while Classification is discrete

b. Classification is continuous while Regression is discrete

c. Both are discrete

d. Both are Continuous



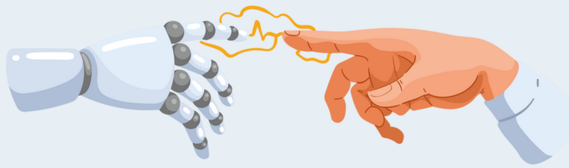
What of the following is an advantage of the Adam optimization technique?

a. Efficient for large datasets and noisy data

b. Faster Convergence

c. Efficient for sparse data

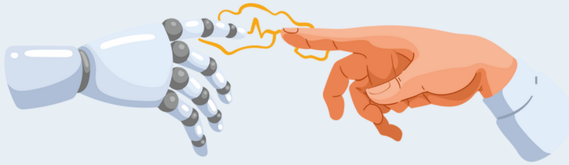
d. Can stop learning too early



Recall measures the ability of a model to identify all relevant instances in the dataset.

a. True

b. False



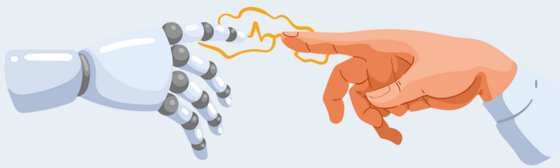
In logistic regression, what is the role of the sigmoid function?

a. To calculate the loss function

b. To map any real value to a probability between 0 and 1

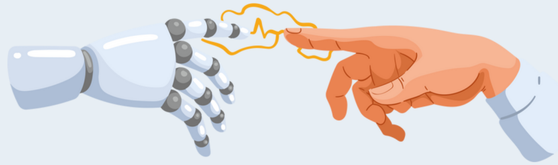
c. To model the relationship between features and target variables

d. To improve the convergence of gradient descent



Which optimization algorithm speeds up convergence by incorporating a term proportional to past gradients?

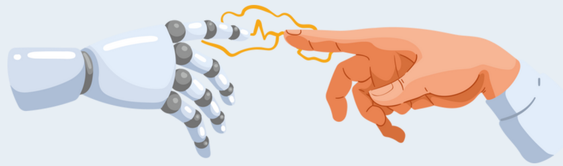
a. Adam
b. Stochastic Gradient Descent
c. Momentum
d. RMSProp



A high learning rate always guarantees faster convergence.

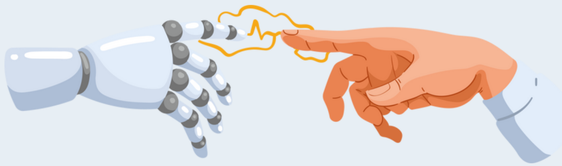
a. True

b. False



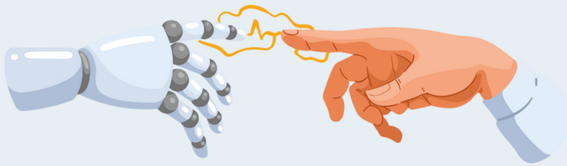
What type of aggregation in Random Forest models have the average of the numerical predictions?

a. Regression
b. Classification
c. Accuracy
d. Recall



Why is the F1 score less affected by class imbalance compared to accuracy?

a. It only considers true positives and true negatives
b. It averages precision and recall
c. It focuses on false positives only
d. It uses the harmonic mean of precision and recall



Which evaluation metric is most appropriate for an imbalanced classification problem?

a. Accuracy
b. Precision
c. F1-Score
d. Mean Squared Error

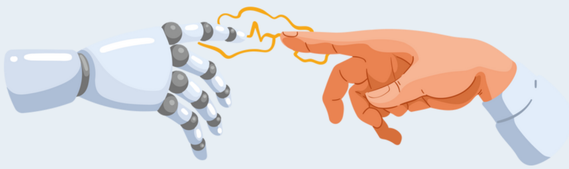
A Stochastic Gradient Descent has …….. variance and …… Convergence

a. Low, Slow

b. Low, Faster

c. High, Slow

d. High, Faster



Variant	Advantages	Disadvantages
Batch gradient descent	Guaranteed convergence to global optimum	Computationally expensive for large datasets, slow convergence
Stochastic gradient descent	Faster convergence, more efficient for large datasets	High variance, may not converge to global optimum
Mini-batch gradient descent	Balanced convergence speed and computational cost, efficient for large datasets	Choice of mini-batch size can be a challenge
Momentum gradient descent	Faster convergence, less likely to get stuck in local minima	May overshoot and oscillate around the optimum
Adagrad	Adaptive learning rate, efficient for sparse data	Can stop learning too early
RMSProp	Adaptive learning rate, efficient for non-stationary problems	Can stop learning too early, requires tuning of hyperparameters
Adam	Adaptive learning rate, efficient for large datasets and noisy data	Can converge to suboptimal solutions, requires tuning of hyperparameters