Recap

TOTAL POINTS 6

1. What back propagation is usually used for in neural networks? 1 point

- ( ) Make several random perturbations of parameters and go back to the best one
- ( ) To propagate signal through network from input to output only
- (•) To calculate gradient of the loss function with respect to the parameters of the network
- ( ) Select gradient update direction by flipping a coin

2. Suppose we've trained a RandomForest model with 100 trees. Consider two cases: 1 point

1. We drop the first tree in the model
2. We drop the last tree in the model

We then compare models performance *on the train set*. Select the right answer.

- (•) In the *case 1* performance **will be roughly the same** as in the *case 2*
- ( ) In the *case 1* performance **will drop less** than in the *case 2*

- ○ In the *case 1* performance **will drop more** than in the *case 2*

3. Suppose we've trained a GBDT model with 100 trees with a fairly large learning rate. Consider two cases: 1 point

   1. We drop the first tree in the model
   2. We drop the last tree in the model

   We then compare models performance *on the train set*. Select the right answer.

   - ○ In the *case 1* performance **will be roughly the same** as in the *case 2*
   - ○ In the *case 1* performance **will drop less** than in the *case 2*
   - ◉ In the *case 1* performance **will drop more** than in the *case 2*

4. Consider two cases: 1 point

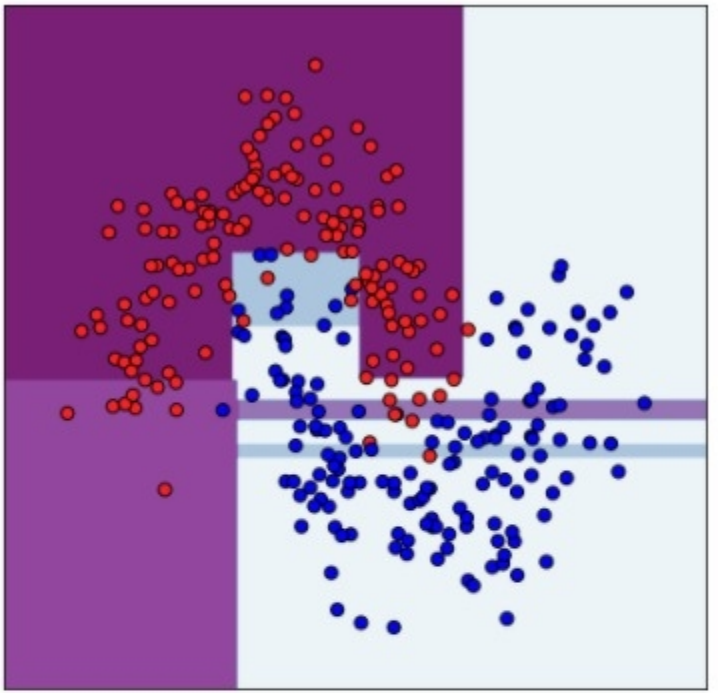
   1. We fit two RandomForestClassifiers 500 trees each and average their predicted probabilities on the test set.
   2. We fit a RandomForestClassifier with 1000 trees and use it to get test set probabilities.

   All hyperparameters except number of trees are the same for all models.

Select the right answer.

- (●) The quality of predictions in the *case 1* **will be roughly the same** as the quality of the predictions in the *case 2*
- ( ) The quality of predictions in the *case 1* **will be higher** than the quality of the predictions in the *case 2*
- ( ) The quality of predictions in the *case 1* **will be lower** than the quality of the predictions in the *case 2*

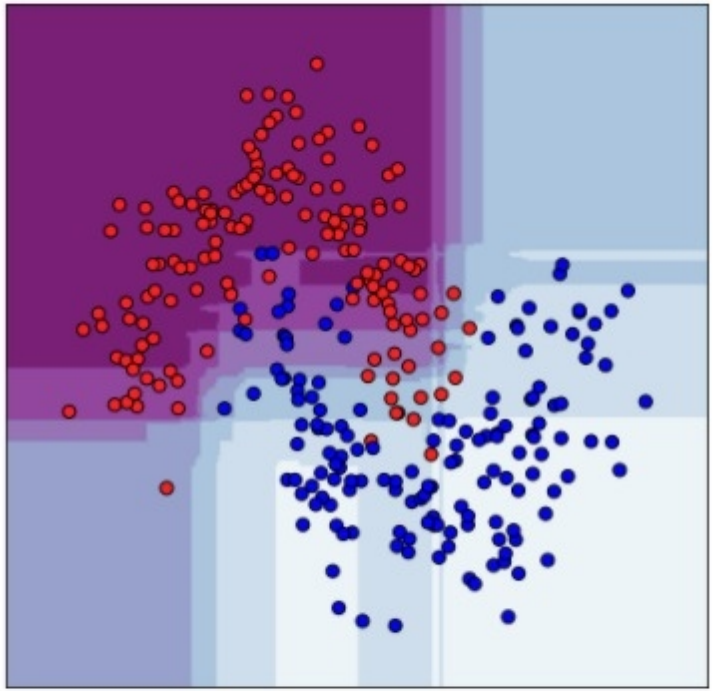
5. What model was most probably used to produce such decision surface? Color (from white to purple) shows predicted probability for a point to be of class "red". 1 point



- (●) Decision Tree
- ( ) Linear model
- ( ) Random Forest
- ( ) k-NN

6. What model was most probably used to produce such decision surface?

1 point



- (●) Random Forest
- ( ) Linear model
- ( ) Decision Tree
- ( ) k-NN