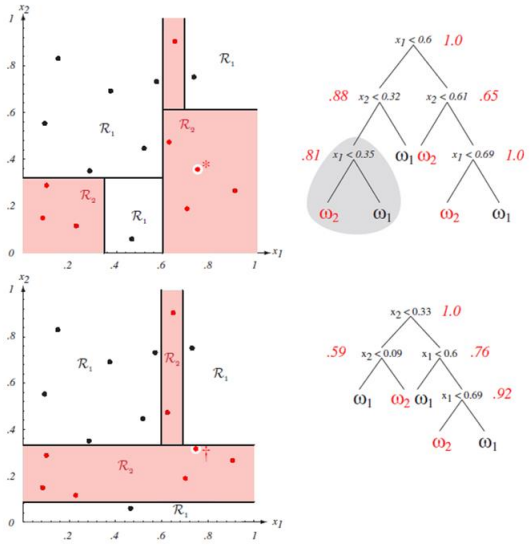
REI602M **Quiz 5.2.2021**

Hand in on Gradescope before 22:00 on Feb. 6 (Saturday). Each question will be given 1, 0.5 or 0 points as follows. If the question is more or less correct it gets 1 point. If it is partly correct it gets 0.5, and if it is missing or completely wrong it gets 0 points.

1) The data set below has two classes (black and red). Two trees were constructed for this data set using the CART algorithm. The top tree corresponds to the last example in the list having coordinates (0.75; 0.36) while the bottom tree corresponds to the last example having coordinates (0.75; 0.32), i.e. the last point has been perturbated slightly before constructing the lower tree. The corresponding decision boundaries are also shown.

ω_1 (black)	
x_1	x_2
.15	.83
.09	.55
.29	.35
.38	.70
.52	.48
.57	.73
.73	.75
.47	.06

ω_2 (red)	
x_1	x_2
.10	.29
.08	.15
.23	.16
.70	.19
.62	.47
.91	.27
.65	.90
.75	.36* (.32†)



It is frequently stated that one of the major features of decision tree classifiers is that they give insight into the classification process. Comment briefly on this statement in light of the above example.

Answer:

The information we get from a decision tree shows the most important features of the data when classifying it to a classifier, let's say we have a classifiers A and B, we input a training data and see when value of an attribute is equal to something then It's classified as A 95% of the time, that would mean that this is an interesting value.

2) For each of the following statements, comment briefly on whether it is TRUE or FALSE. Give a 1 – 2 sentence justification of your answer.

a) A Random Forests classifier can give a good estimate of accuracy (or error) in classification without the need for a separate test set.

True, there is no need for a separate test set since the classifier consists of a large number of individual decision trees that operate with an ensemble learning.

b) Random Forests classifiers are most often trained using Stochastic gradient descent.

False, the training algorithm for Random_forsets uses Bagging which is a technique of bootstrap aggregation.

c) By increasing the depth of a decision tree, the fit to the training data improves but robustness decreases (sensitivity to perturbations in the input data).

True, this is an example of overfitting, if we use more data the outliers could change the results drastically.

3) “Bootstrap aggregation” (bagging) aðferðafræðinni er oftast notuð með trjáflokkurum en það er hægt að nota hana með hvaða flokkara sem er. Ef bagging er notuð með SVM fæst hins vegar lítill eða engine bæting í nákvæmni. Hvers vegna telur þú að svo sé?

because SVM only weights instances not features, bagging uses features so if we would normalize the data in SVM we would be able to use bagging with better accuracy.