

COMP4702/COMP7703 - Machine Learning

Homework 9 - Evaluating Models and Decision Trees

Marcus Gallagher

Core Questions

1. I found a random page on the internet containing python code and when I ran it the following confusion matrix was produced (note that the correct classes run down the rows and the predicted classes across the columns as in the lecture and Alpaydin):

Confusion matrix

```
[[27  0  0  0  0  0  0  0  0  0]
 [ 0 37  0  0  0  0  0  0  0  0]
 [ 0  1 24  0  0  0  2  1  0  0]
 [ 0  0  0 29  0  3  0  1  1  1]
 [ 0  0  0  0 24  0  0  0  0  0]
 [ 0  0  0  0  0 32  0  0  0  2]
 [ 0  0  0  0  0  1 24  0  0  0]
 [ 0  0  0  0  1  3  0 31  0  0]
 [ 0  1  0  0  0  0  1  0 27  2]
 [ 0  0  0  0  1  2  0  0  0 21]]
```

- (a) What is the size of the training set used?
 - (b) Which class had the highest number of errors?
 - (c) What was the overall percentage correct?
2. Decision trees partition the feature space into regions which are (hyper)rectangles. Consider the decision tree shown in Figure 9.5 of the Hastie et al. textbook. How many hyper-rectangles does the model define in the feature space for the **spam** class?
 3. Consider again the decision tree shown in Figure 9.5 of the Hastie et al. textbook. How many parameters does this model have?

Extension Questions

In my opinion, one of best experimental comparisons of classification techniques ever done can be found in this book[1] (pdf version is freely available and I have put it on the course blackboard site). You might like it as an alternative reference for some of the material that we have studied in the course. The book is now old enough (1994) to have some historical value - it does not include support vector machines, ensemble methods and deep learning because it pre-dates them!

Grab a coffee, sit down and have a relaxing read of the start of Chapter 7 of this book (<3 pages, up to the end of 7.1.3 only). Do you think it is still correct and useful in 2020? Are there any statements that are out of date or irrelevant? Write a paragraph (approx. 250 words) as your answer.

References

- [1] D. Michie et al. (editors), Machine Learning, Neural and Statistical Classification. Ellis Horwood, 1994.