6CCS3ML1 (Machine Learning)

Tutorial 02

(Version 1.1)

1. The following table gives some examples of recent films I have browsed using the WebMooviz app:

Examples	Attributes			Stream?
	New	Lang	Type	
X_1	N	Eng	Action	Y
X_2	N	Sp	Comedy	Y
X_3	Y	Eng	Drama	N
X_4	Y	Sp	Comedy	Y
X_5	N	Sp	Action	Y
X_6	N	Sp	Drama	Y
X_7	N	Fr	Comedy	N
X_8	Y	Sp	Action	N
X_9	Y	Eng	Drama	Y
X_{10}	Y	Fr	Action	N

   This records whether or not the film is a *New* release, what *Lang*uage the film is in (English, French or Spanish), and what *Type* the film is (Action, Comedy, Drama). The app also records whether or not I actually streamed the film (or just browsed it).

   (a) Describe how a decision tree could be learnt from this data.

   (b) Show how the idea of entropy could be used to pick the first node in the decision tree.

2. Answer Q1(b) using Gini impurity rather than entropy.

3. (a) What does the term "linearly separable" mean with respect to a data set?

   (b) Give an example of a data set that is linearly separable and one that is not.
   For this exercise find an example online that is different from the one that I showed you in the lecture.

   (c) If you had to design a classifier to work on a particular dataset, how would the knowledge that the dataset was not linearly separable affect the method that you chose to use?

4. You have the following training data:

Instance	x	y
E_1	1.5	1
E_2	3.5	3
E_3	3	2
E_4	5	3
E_5	2	2.5

(a) Use batch gradient descent with initial values:

$$w_0 = 0$$
$$w_1 = 0$$

and $\alpha = 0.01$ to compute the weights. Use the training set twice.

(b) Now start from the same values and use stochastic gradient. Use the training set once. How do your final values compare to the results from batch gradient descent?

For comparison, the function from which the results were "generated" is:

$$y = 1 + 0.5x$$

Of course, you should remember that the full method is to continue until the values converge, so you what you get will not necessarily be close to this after on iteration of the small amount of training data.

5. (a) Write down the algorithm you would use to learn a two variable linear regression model from data.

   (b) Explain the effect of learning rate on the performance of your algorithm.

Version list

- Version 1.0, January 18th 2020.
- Version 1.1, January 8th 2021.
- Version 1.2, January 15th 2022.