

DAILY ASSESSMENT FORMAT

Course:	Machine learning with Python	Name:	Bindu.N.R
Link:	https://cognitiveclass.ai/courses	USN:	4AL17EC101
Org By :	IBM	Semester & Section:	6-B
Github Repository:	bindunr-python	Date:	18/06/2020

Progress on 18-06-2020

• Topic Completed Today

Building Decision Trees (10:37)

Building Decision Trees (10:37)

Is 'Cholesterol' the best attribute?

Patient ID	Age	Sex	BP	Cholesterol	Drug
p1	Young	F	High	Normal	Drug A
p2	Young	F	High	High	Drug A
p3	Middle-age	F	High	Normal	Drug B
p4	Senior	F	Normal	Normal	Drug B
p5	Senior	M	Low	Normal	Drug B
p6	Senior	M	Low	High	Drug A
p7	Middle-age	M	Low	High	Drug B
p8	Young	F	Normal	Normal	Drug A
p9	Young	M	Low	Normal	Drug B
p10	Senior	M	Normal	Normal	Drug B
p11	Young	M	Normal	High	Drug B
p12	Middle-age	F	Normal	High	Drug B
p13	Middle-age	M	High	Normal	Drug B
p14	Senior	F	Normal	High	Drug A

Cholesterol

Normal: S: [9 B, 5 A], E = 0.940

High: S: [6 B, 2 A], E = 0.811

From the left side, we can see that we have one with the most "predictiveness," which results in two more pure branches. Let's first select the "Cholesterol" of the patient and see how the data gets split, based on its values. For example, when it is "normal," we have 6 for Drug B, and 2 for Drug A. We can calculate the Entropy of this node based on the distribution of drug A and B, which is 0.8 in this case. But, when Cholesterol is "High," the data is split into 3 for drug B and 3 for drug A. Calculating its entropy, we can see it would be 1.0. We should go through all the attributes and calculate the "Entropy" after the split, and then choose the best attribute. Ok, let's try another field. Let's choose the Sex attribute for the next check. As you can see, when we use the Sex attribute to split the data, when its value is "Female," we have 2 patients that responded to Drug B.

Logistic Regression vs Linear Regression (29:20)

[Bookmark this page](#)

Logistic Regression vs Linear Regression (15:30)

The training process

1. Initialize θ .
2. Calculate $\hat{y} = \sigma(\theta^T X)$ for a customer.
3. Compare the output of \hat{y} with actual output of customer, y , and record it as error.

$\sigma(\theta^T X) \rightarrow P(y=1|x)$

$\theta = [-1, 2]$

$\hat{y} = \sigma([-1, 2] \times [2, 5]) = 0.7$

Error = $1 - 0.7 = 0.3$

example, the age and income of the customer, for instance [2,5].

And θ is the confidence or weight that you've set in the previous step.

The output of this equation is the prediction value ... in other words, the probability that the customer belongs to class 1.

Step 3. Compare the output of our model, $y^{\hat{}}$, which could be a value of, let's say, 0.7, with the actual label of the customer, which is for example, 1 for churn.

Then, record the difference as our model's error for this customer, which would be 1-0.7, which of course equals 0.3. This is the error for only one customer out of all the customers in the training set.

Step 4. Calculate the error for all customers as we did in the previous steps, and add up these errors. The total error is the cost of your model.

and is calculated by the model's cost function. The cost function, by the way, basically represents how to calculate the error of the model, which is

Video
[Download video file](#)

Transcripts
[Download SubRip \(.srt\) file](#)
[Download Text \(.txt\) file](#)

• Progress Report



