Logistic Regression

Joe

Introduction

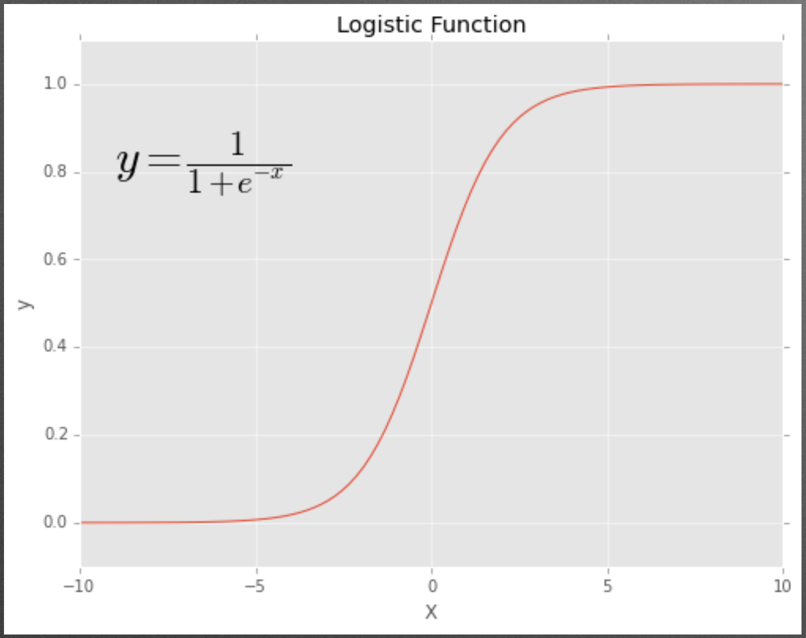
Session Objective

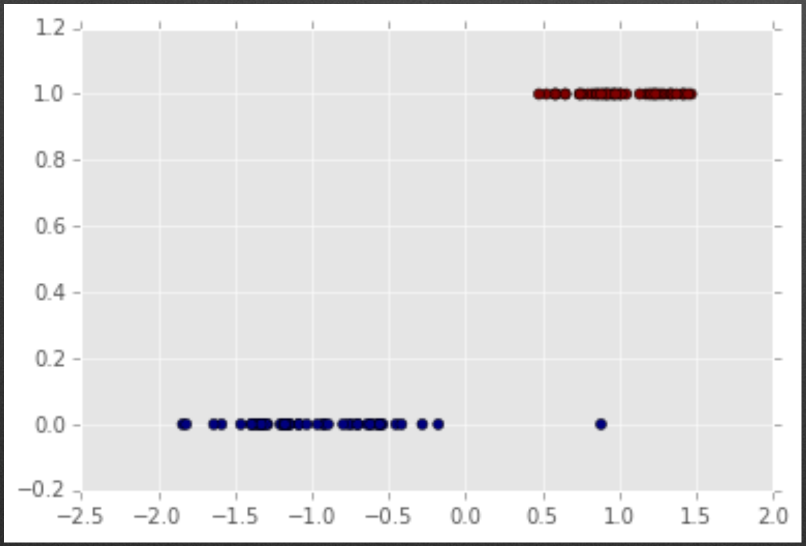
Please clone:
https://github.com/drJAGartner/log_reg_demo

1. Build a logistic regression model that describes the probability that an iris belongs to a particular category based on measurable parameters.

2. With the same model, create a confusion matrix to show the performance of the model.

The W’s of Logistic Regression





- What -
  a method of producing continuous predictions (i.e. regression) when the *dependent variable* is discrete

- Why -
  the output of logistic regression is bound between 0 and 1, and can be directly interpreted as an outcome probability

- When -
  data has classes that are linearly separable (i.e. can be split by a line, bottom)

  Although logistic regression is a continuous model, it is used as a classification model

- <u>Side note, the logistic function is also called the sigmoid, and is often used in neural networks as an activation function</u>

Classification Models

What is Classification

- Up to this point, all of the models we have produced numeric output, i.e. regression models
- Classification models are built for instances where the output have discrete meaning
  - e.g. image classifier: cat or dog
- Logistic Regression becomes a classifier by applying a decision threshold

Measures of Effectiveness (1)

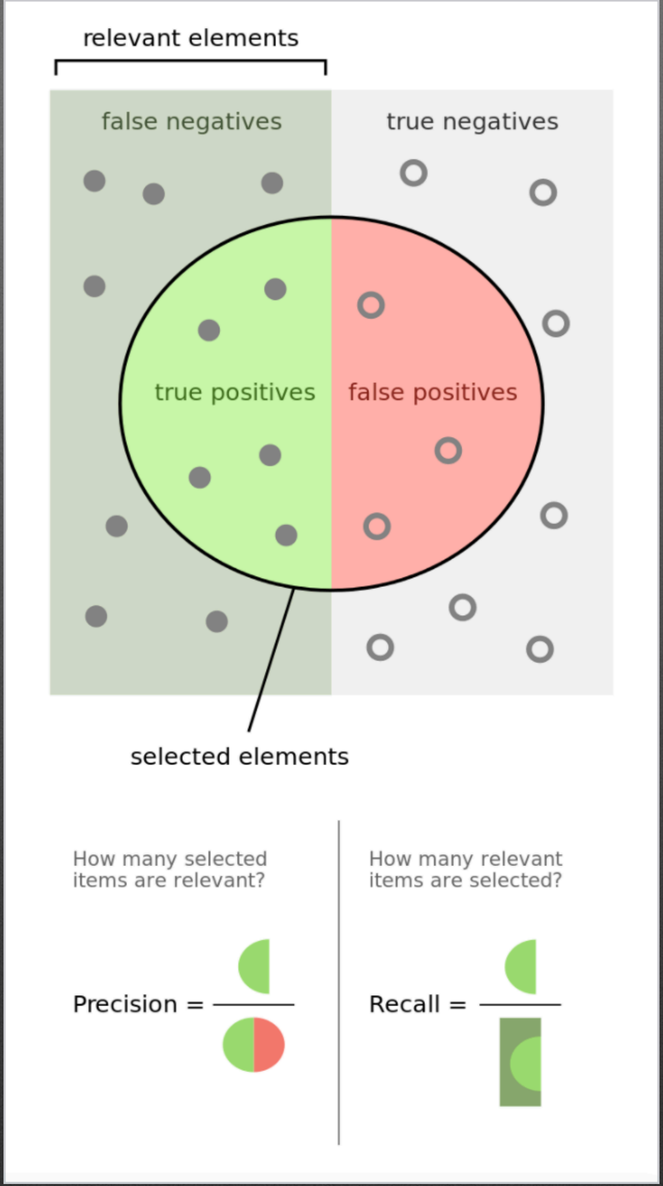
		True condition	
	Total population	Condition positive	Condition negative
Predicted condition	Predicted condition positive	**True positive**	**False positive** (Type I error)
	Predicted condition negative	**False negative** (Type II error)	**True negative**

Confusion matrix:
https://en.wikipedia.org/wiki/Confusion_matrix

Measures of Effectiveness (2)

		True condition		
	Total population	Condition positive	Condition negative	Prevalence $= \frac{\Sigma \text{ Condition positive}}{\Sigma \text{ Total population}}$
Predicted condition	Predicted condition positive	**True positive**	**False positive** (Type I error)	Positive predictive value (PPV), Precision $= \frac{\Sigma \text{ True positive}}{\Sigma \text{ Predicted condition positive}}$
	Predicted condition negative	**False negative** (Type II error)	**True negative**	False omission rate (FOR) $= \frac{\Sigma \text{ False negative}}{\Sigma \text{ Predicted condition negative}}$
		True positive rate (TPR), Recall, Sensitivity, probability of detection $= \frac{\Sigma \text{ True positive}}{\Sigma \text{ Condition positive}}$		
		False negative rate (FNR), Miss rate $= \frac{\Sigma \text{ False negative}}{\Sigma \text{ Condition positive}}$		

Another Visualization



Measures of Effectiveness (3)

True condition	
Condition positive	Condition negative
True positive	**False positive** (Type I error)
False negative (Type II error)	**True negative**
True positive rate (TPR), Recall, Sensitivity, probability of detection $= \frac{\Sigma \text{ True positive}}{\Sigma \text{ Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm $= \frac{\Sigma \text{ False positive}}{\Sigma \text{ Condition negative}}$
False negative rate (FNR), Miss rate $= \frac{\Sigma \text{ False negative}}{\Sigma \text{ Condition positive}}$	True negative rate (TNR), Specificity (SPC) $= \frac{\Sigma \text{ True negative}}{\Sigma \text{ Condition negative}}$

Measures of Effectiveness (4)

Accuracy (ACC) =

$$\frac{\Sigma \text{ True positive} + \Sigma \text{ True negative}}{\Sigma \text{ Total population}}$$

F_1 score =

$$\frac{2}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}}$$

Back to Logistic Regression

From Linear Model to Probability Space

Start from the assumption we want a function to describe probabilities [0, 1] based on a linear model [-∞, ∞]. The function that has this property is the logit

$$g(F(x)) = \ln\left(\frac{F(x)}{1 - F(x)}\right) = \beta_0 + \beta_1 x,$$

From Linear Model to Probability Space

Start from the assumption we want a function to describe probabilities [0, 1] based on a linear model [-∞, ∞]. The function that has this property is the logit

$$g(F(x)) = \ln\left(\boxed{\frac{F(x)}{1 - F(x)}}\right) = \beta_0 + \beta_1 x,$$

odds

From Linear Model to Probability Space

Start from the assumption we want a function to describe probabilities [0, 1] based on a linear model [-∞, ∞]. The function that has this property is the logit

$$g(F(x)) = \boxed{\ln\left(\frac{F(x)}{1-F(x)}\right)} = \beta_0 + \beta_1 x,$$

Logit

From Linear Model to Probability Space

Start from the assumption we want a function to describe probabilities [0, 1] based on a linear model [-∞, ∞]. The function that has this property is the logit

$$g(F(x)) = \ln\left(\frac{F(x)}{1 - F(x)}\right) = \beta_0 + \beta_1 x,$$

Here, F(x) is a probability (i.e. bound from 0-1). Let’s whiteboard this for a clearer picture.

Enough Slides!

We discussed how Logitic Regression can be used in theory, let’s put it to practice by walking through some example cases.

Let’s hop into a notebook and start looking!