Lecture Logistic Regression

Class	02_titanic
Created	@Jan 15, 2021 1:43 PM
Materials	
Reviewed	☐
Type	

Logistic Regression

0) Warmup

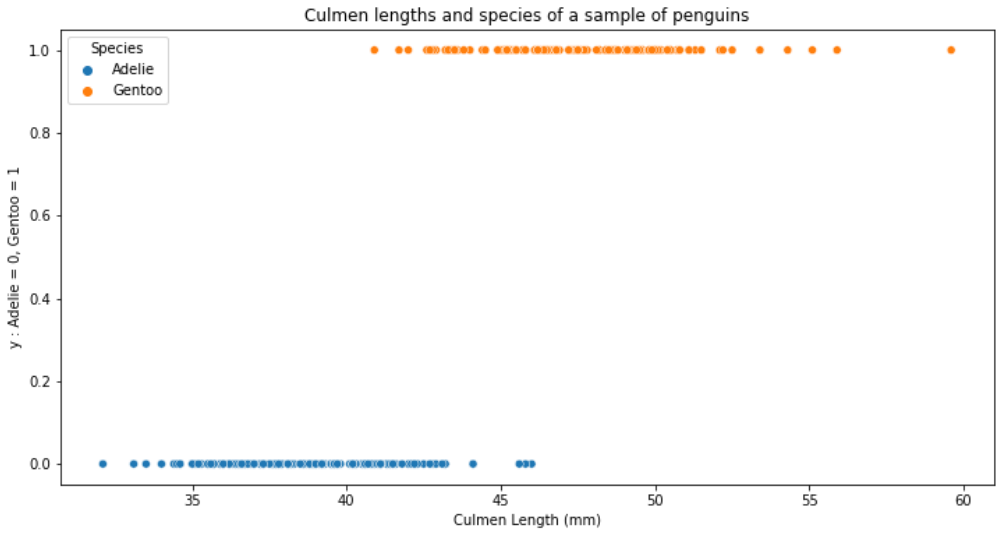
Imagine you are a biologist wanting to assign a penguin to its correct species (***y***). Assume that the only information available to you is the Bill Length of the penguin in mm (***x***). You have some **labeled data** that you can use to make predictions for penguins you are going to find on your next expedition.

Q: What type of machine learning problem is this?

- Classification Problem (Supervised Learning) → Focus of this week

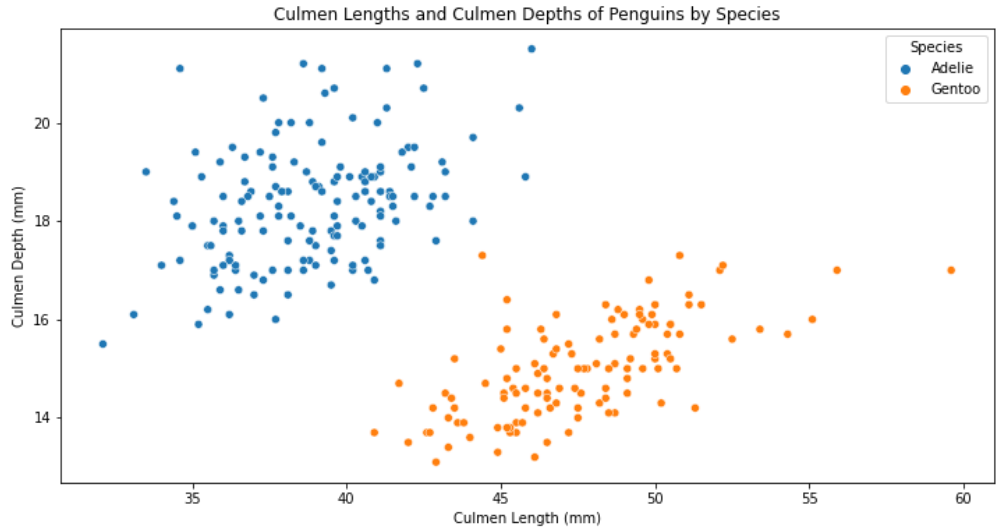
Q: Looking at the image below, what rule could you come up with to assign penguins to the correct class (species)?

We could introduce a cutoff value for Bill Length, for example 42 mm, and classify all penguins with a smaller bill as Adelie and all penguins with a larger bill as Gentoo



Q: Does the certainty of your prediction depend on the value of culmen length?

Q: How would you change your approach if you had additional data on the culmen depth of each penguin?



1) Goal of the model

We want to predict the species of a penguin

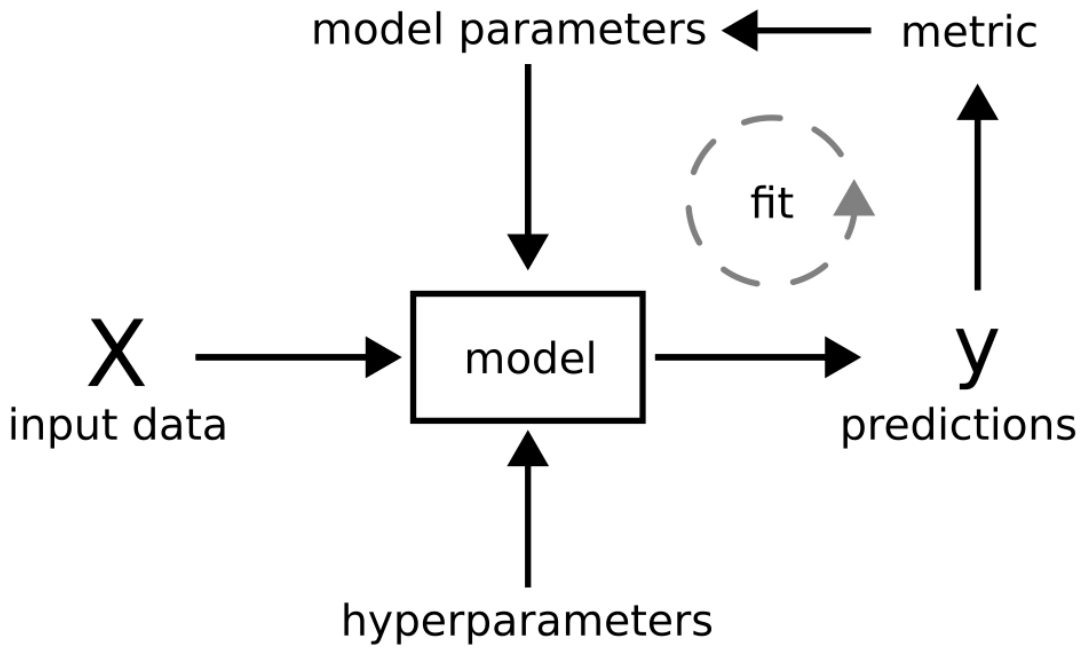
💡 For the project dataset (titanic) the goal of the model is to predict whether a passenger died or survived. The model will need to use input features X in order to make these predictions

2) How does the model come up with the predictions?

- $\hat{p}(y = 1|X) = \frac{1}{1+e^{-(w_0+w_1*x)}}$ (in case of one feature x)
- $\hat{p}(y = 1|X) = \frac{1}{1+e^{-(w_0+w_1*x_1+...+w_p*x_p)}}$ (in case of p features X)

x values are our input parameters. We know them.

$w_0, w_1, ..., w_p$ are the coefficients of the model. Those are the parameters that the model **learns** for us.



3) How do we come up with the coefficients?

The implementation of Logistic Regression in scikit-learn does that for us. For today we will not deal with this question. On Friday and next week you are talking about the general mechanics behind this.

4) How do we implement this in code?

Let's go to the notebook.