Logistic regression is a fundamental classification technique. It belongs to the group of **linear classifiers** and is somewhat similar to polynomial and linear regression. Logistic regression is fast and relatively uncomplicated, and it’s convenient for you to interpret the results. Although it’s essentially a method for binary classification, it can also be applied to multiclass problems.

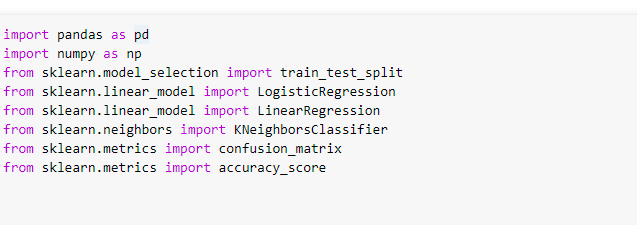
**Methodology**

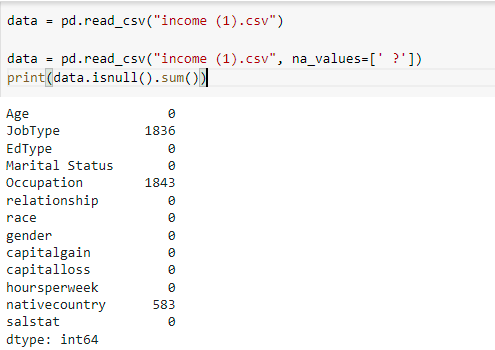
Logistic regression is a linear classifier, so you’ll use a linear function 𝑓(𝐱) = 𝑏₀ + 𝑏₁𝑥₁ + ⋯ + 𝑏ᵣ𝑥ᵣ, also called the **logit**. The variables 𝑏₀, 𝑏₁, …, 𝑏ᵣ are the **estimators** of the regression coefficients, which are also called the **predicted weights** or just **coefficients**.

The logistic regression function 𝑝(𝐱) is the sigmoid function of 𝑓(𝐱): 𝑝(𝐱) = 1 / (1 + exp(−𝑓(𝐱)). As such, it’s often close to either 0 or 1. The function 𝑝(𝐱) is often interpreted as the predicted probability that the output for a given 𝐱 is equal to 1. Therefore, 1 − 𝑝(𝑥) is the probability that the output is 0.

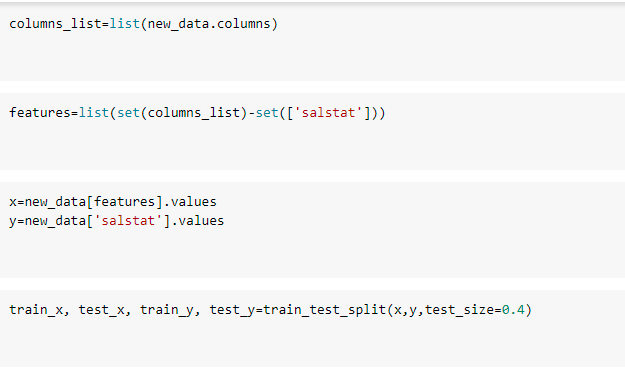
|  |
| --- |
| **For iNCOME dataset** |

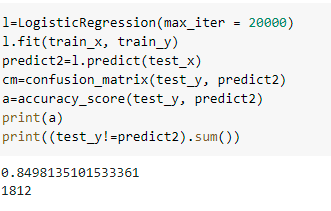
|  |
| --- |
| **SCREEN SHOTS:** |

****

****

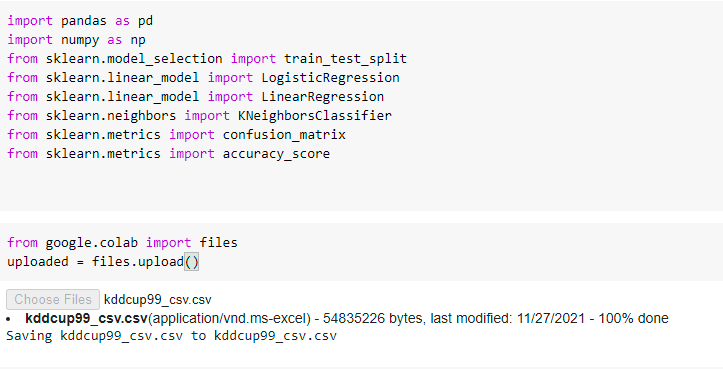
****

****

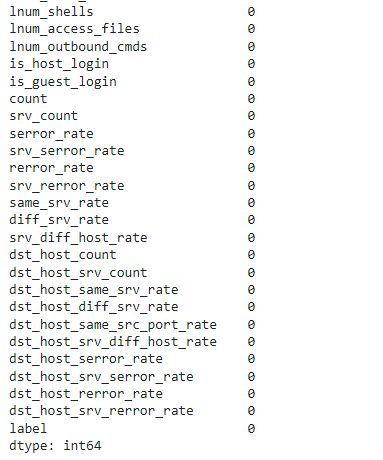
****

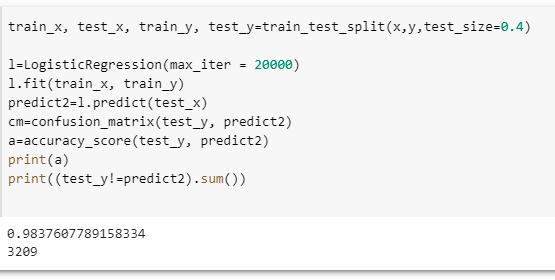
Accuracy for the income dataset is **84.98 % .**

|  |
| --- |
| **For IDS dataset** |

****

****

****

****

Accuracy for the IDS dataset is **98.37 % .**