Chapter 6: Logistic Regression

- It is a supervised classification algorithm.
- Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables.

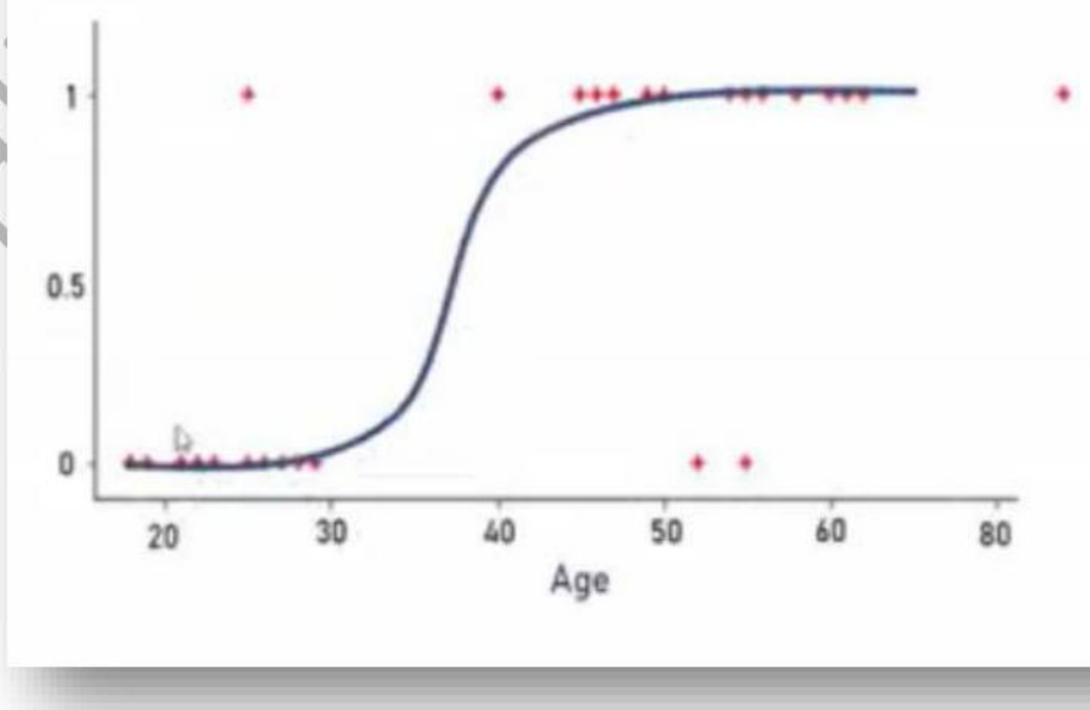
Types of Logistic Regression:

- Types of Logistic regression:
  - Binary Classifier
  - Multinomial Classifier

Binomial Classifier:

- If the logistic regression model is used for addressing binary classification problem (yes/no, 0/1, true/false) then its called binary logistic regression classifier.
- Binary classifier uses Sigmoid function.

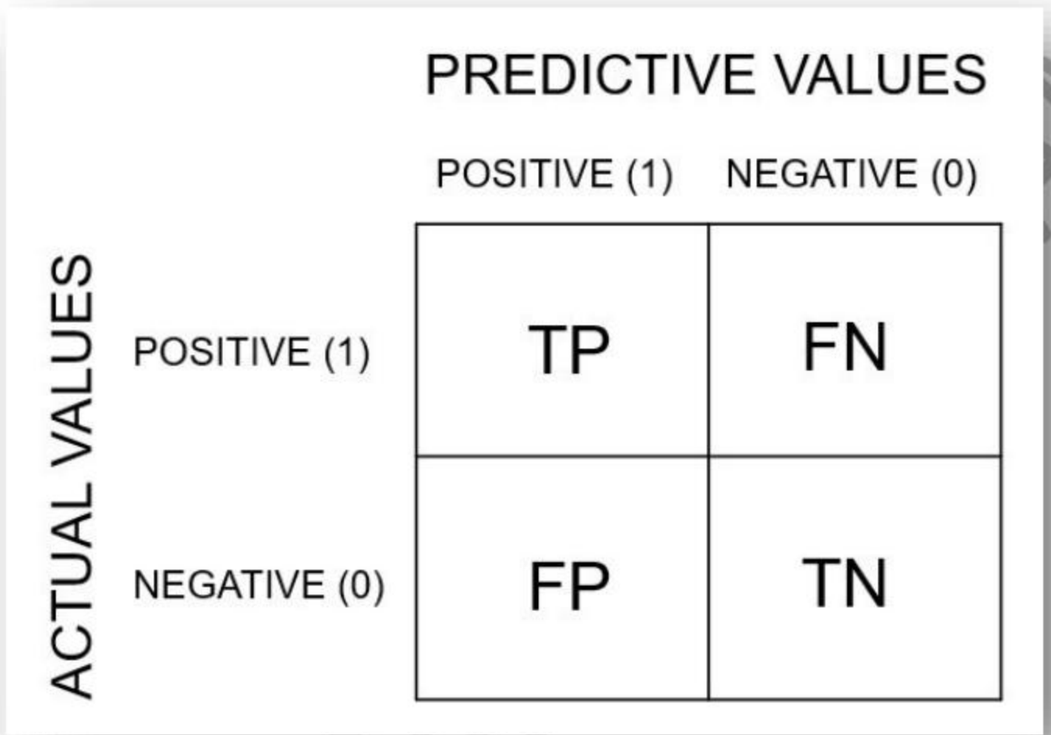
$$y = \frac{1}{1 + e^{-(m*x+b)}}$$



Confusion Matrix:

- Confusion matrix is used for summarizing the performance of a classification algorithm.
- Simply it is the number of times a model rightly identifies the truth (actual classes) and the number of times it gets confused in identifying one class from another.
- Calculating a confusion matrix can give you an idea of where the classification model is right and what types of errors it is making.

Confusion Matrix for binary classification:



- **True Positive:**
- You predicted positive and it's true. You predicted that an animal is a cat and it actually is.
- **True Negative:**
- You predicted negative and it's true. You predicted that animal is a dog and it actually is.
- **False Positive (Type 1 Error):**
- You predicted positive and it's false. You predicted that animal is a cat but it actually is not (it's a dog).
- **False Negative (Type 2 Error):**
- You predicted negative and it's false. You predicted that animal is not a cat but it actually is.

Precision:

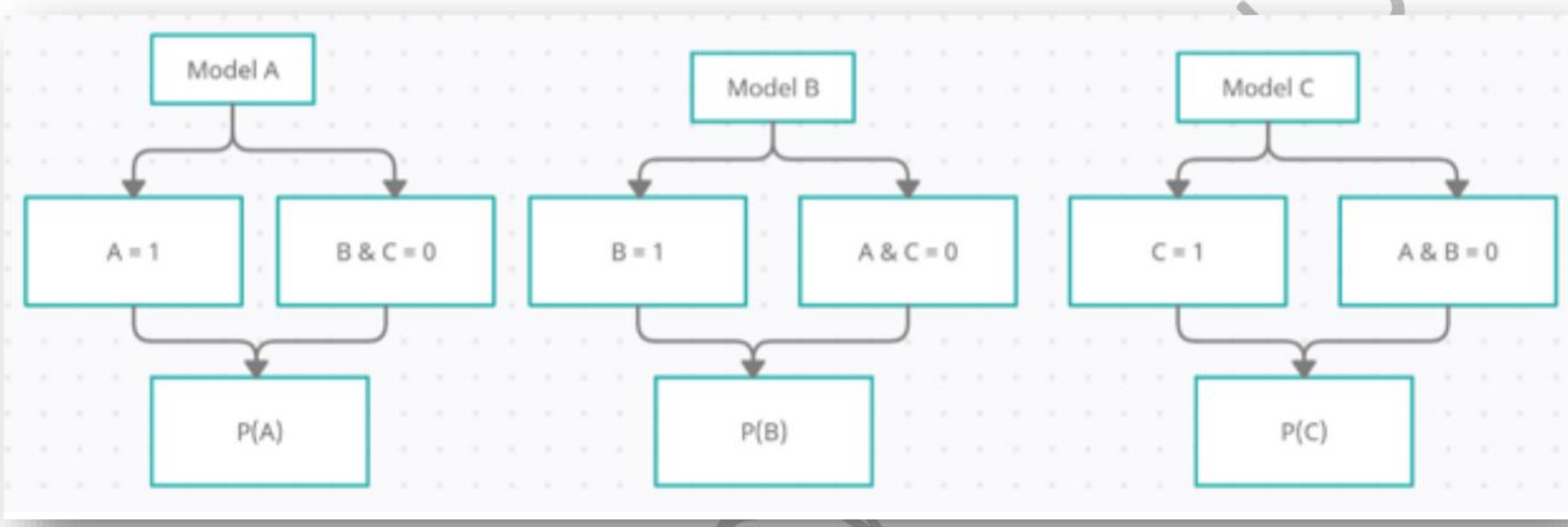
How often is the model right?

Recall :

How often model actually predicts the correct values?

Multinomial Classifier:

- k models will be built for k classes as a set of independent binomial logistic regression.
- For the given test data, a probability value would be created using each model.
- Model giving the highest probability would be selected and corresponding class would be the output.



	Linear Regression	**Logistic Regression**
Usage	It is used for solving regression/prediction problems	It is used for solving classification problems
Data	It works with continuous data such as price, salary etc.	It works with categorical data such as 0/1, yes/no, cycle/bike/car
How it works	We find the best fit line which can be used to predict the output.	We find the S-curve which can be used to classify the samples.