What is Classification?

Classification is a process of designating the objects or data points into classes/targets. The data can be structured or unstructured in nature.

The Classifier can predict if the data falls into a class using input data that train the algorithm. The main aim is to find which class/category the new data will fall into.

Let us try to understand with an example



Finding whether a person is obese or not obese is identified as the classification problem. Since there are only two targets i.e. obese/not obese this problem belongs to Binary classification. In this case The classification model, needs training data to understand how the given independent variables are related to the class. And once the classifier is trained accurately, it can be used to detect whether the person is obese or not obese.

Important Classification Terminologies

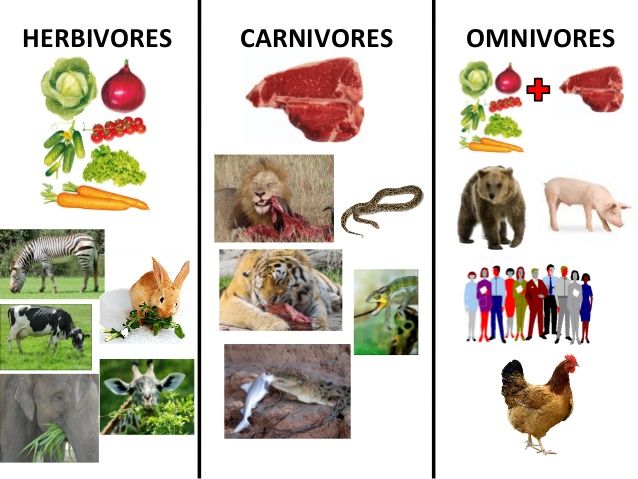
Classification Model- A model which predicts the destination class to the input data given for training.

Feature- Features are the individual measurable property that acts like an input.

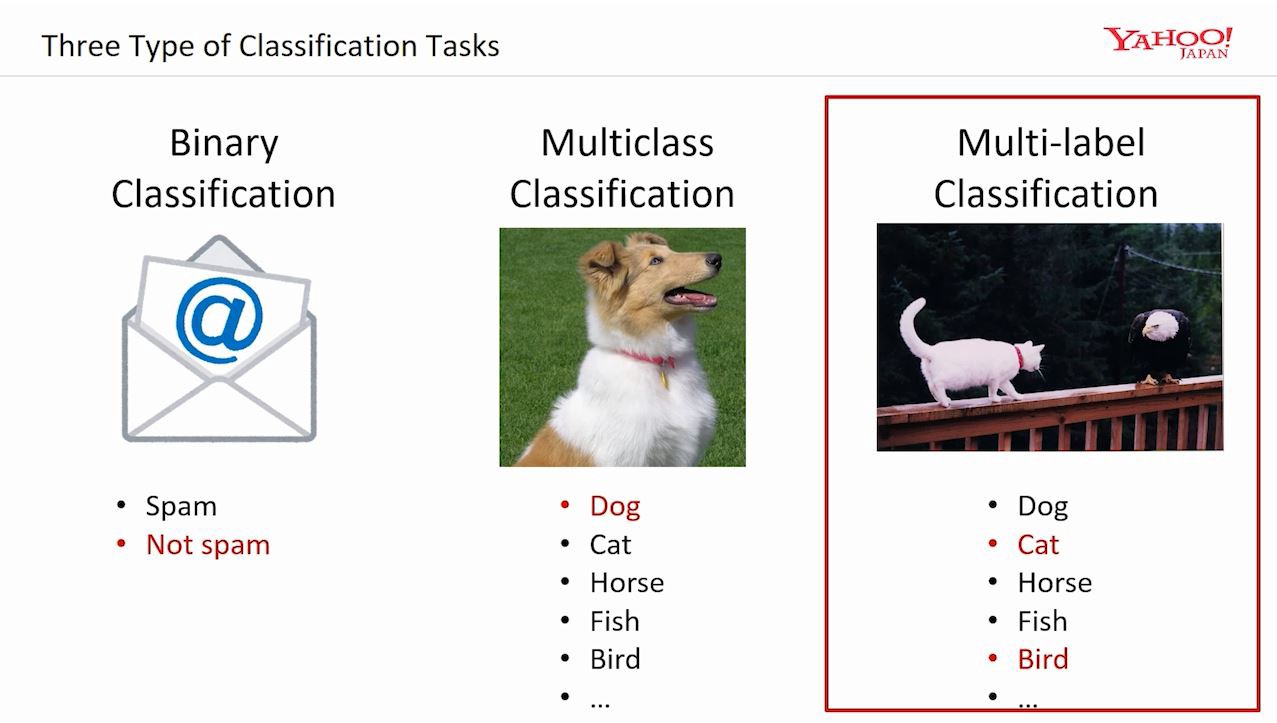
Binary classification- Classification with two target classes, example-spam/not spam, Yes/no, heads/tails.

Multiclass classification- Classification with more than two target classes, each data point is designated to one and only one class.

Example- You are in a forest and your job is to classify all the animals into three categories. Carnivorous, herbivorous and omnivorous.



Multi-label classification- Classification where each sample is assigned to set of labels or targets.



Classification vs Regression

Regression– Regression algorithms are supervised Learning algorithm. Systems where the predicted values belong to the continuous domain. Questions like “how much?” or “how many” are helpful here.

Example- Calculating the strike rate of the batsmen by taking different input conditions. The strike rate of a batsmen can be varied from 0-N where n can be any real value.

Classification- Classification algorithms are also supervised learning algorithm. Systems where the predicted value belong to discrete valued domain like(yes/no, 1/0 etc.). Questions like “Are you single” or “Is this tumor is benign or malignant” are helpful here.

Example- Categorizing the batsmen into form batsmen/inform batsmen by taking different input conditions.

Applications of Classification

[**Sentiment Analysis**](https://monkeylearn.com/blog/classification-algorithms/#sentiment)

[**Email Spam Classification**](https://monkeylearn.com/blog/classification-algorithms/#spam)

[**Document Classification**](https://monkeylearn.com/blog/classification-algorithms/#document)

[**Image Classification**](https://monkeylearn.com/blog/classification-algorithms/#image)

**Speech recognition**

**Classification Algorithms**

Logistic Regression.

Naïve Bayes Classifier

KNN

Decision Tree

Random Forest

SVM