

Machine Learning In Python

Subject : Feature and Classification

Lecturer : Reza Akbari Movahed

Hamedan University of Technology

Winter 2020

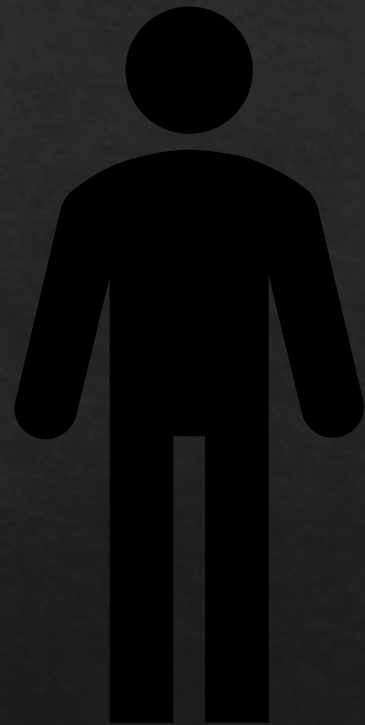
Feature and Classification

Feature

- A feature is a measurable property of the object you're trying to analyze.
- Features are also sometimes referred to as “variables” or “attributes.”
- Depending on what you're trying to analyze, the features you include in your dataset can vary widely.

Feature and Classification

Feature



Age

Height

Weight

Nationality

Sex

...

Feature and Classification

Feature

	A	B	C	D	E	F	G	H	I	J	K	L
1	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
2	1	0	3	Braund, Mr.	male	22	1	0	A/5 21171	7.25		S
3	2	1	1	Cumings, Mr.	female	38	1	0	PC 17599	71.2833	C85	C
4	3	1	3	Heikkinen, M	female	26	0	0	STON/O2. 31	7.925		S
5	4	1	1	Futrelle, Mrs	female	35	1	0	113803	53.1	C123	S
6	5	0	3	Allen, Mr. W	male	35	0	0	373450	8.05		S
7	6	0	3	Moran, Mr. J	male		0	0	330877	8.4583		Q

- The image above contains a snippet of data from a public dataset with information about passengers on the ill-fated Titanic maiden voyage.
- Each feature, or column, represents a measurable piece of data that can be used for analysis: Name, Age, Sex, Fare, and so on.

Feature and Classification

- In Supervised Learning, We have to basic data :

X is the Feature Matrix, L is the Label Matrix

$$[X]_{N \times M}$$

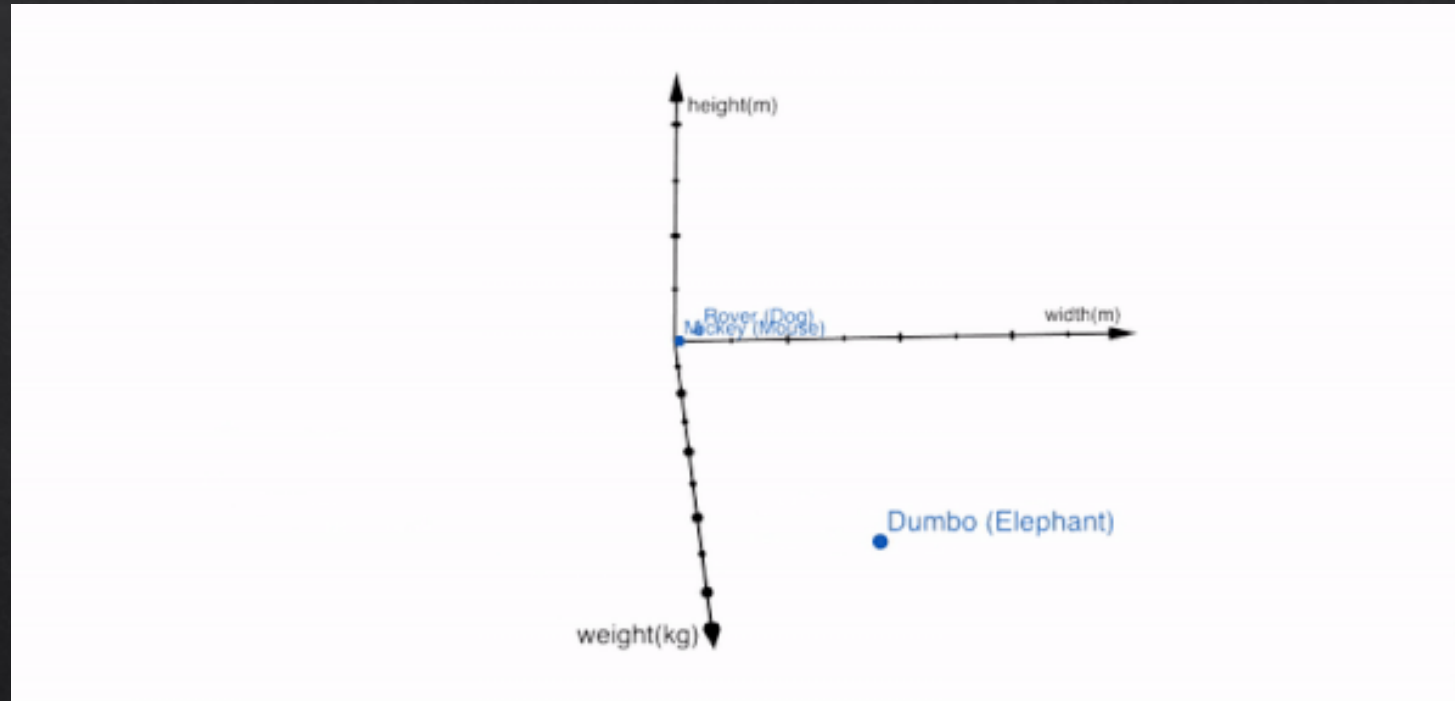
$$[L]_{N \times 1}$$

N is the number of samples.

M is the number of features.

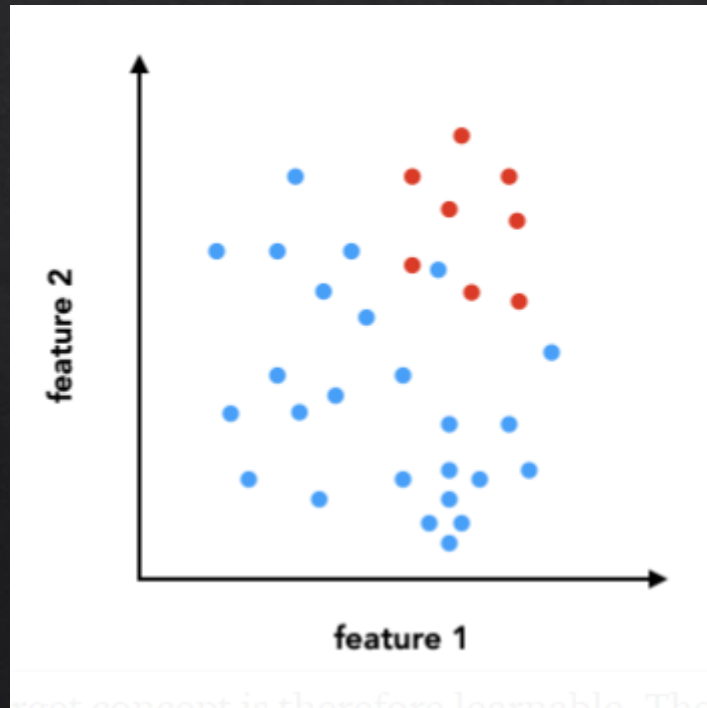
Feature and Classification

- The vector space associated with these vectors is often called the feature space.



Feature and Classification

- The vector space associated with these vectors is often called the feature space.
- The Feature space of a Feature matrix is depicted by Scatter plot



Feature and Classification

- Introduce the Iris plants dataset

7.2.2. Iris plants dataset

Data Set Characteristics:

Number of Instances:	150 (50 in each of three classes)
Number of Attributes:	4 numeric, predictive attributes and the class
Attribute Information:	<ul style="list-style-type: none">• sepal length in cm• sepal width in cm• petal length in cm• petal width in cm• class:<ul style="list-style-type: none">◦ Iris-Setosa◦ Iris-Versicolour◦ Iris-Virginica



Feature and Classification

What is Classification In Machine Learning

- Classification is a process of categorizing a given set of data into classes.
- Classification is a functions of Supervised Learning.
- For 2 class problems, we have Binary Classification.
- For multiple class problems, we have Multiclass Classification.
- The classes are often referred to as target, label or categories.



Setosa

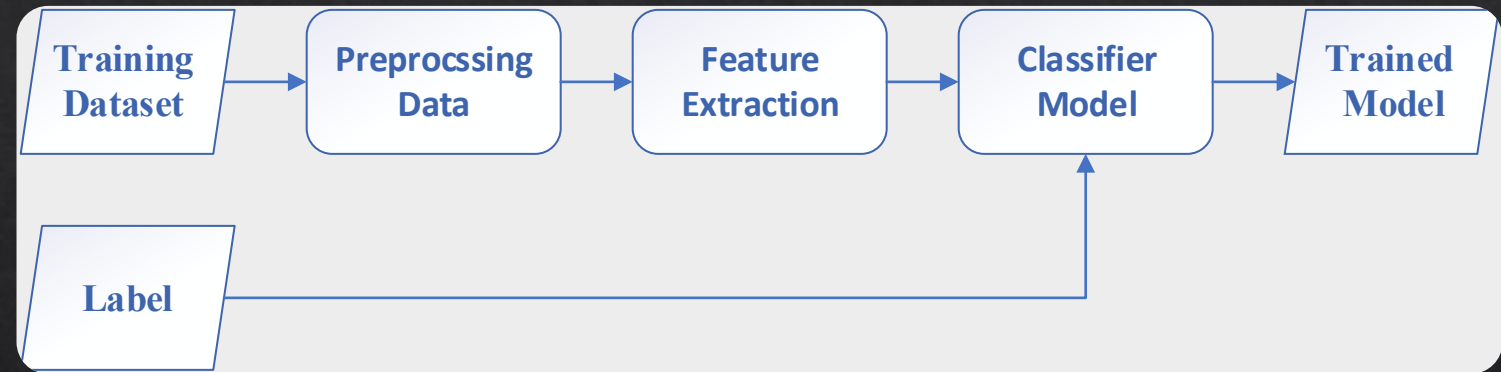
Versicolour

Virginica

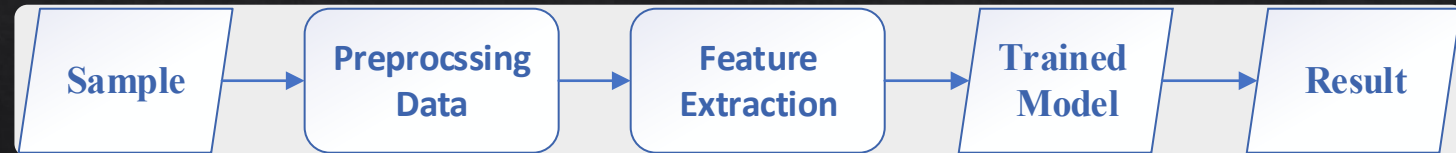
Feature and Classification

Classification In Supervised Learning Framework

Training Phase



Testing Phase



Feature and Classification

Classifier Models :

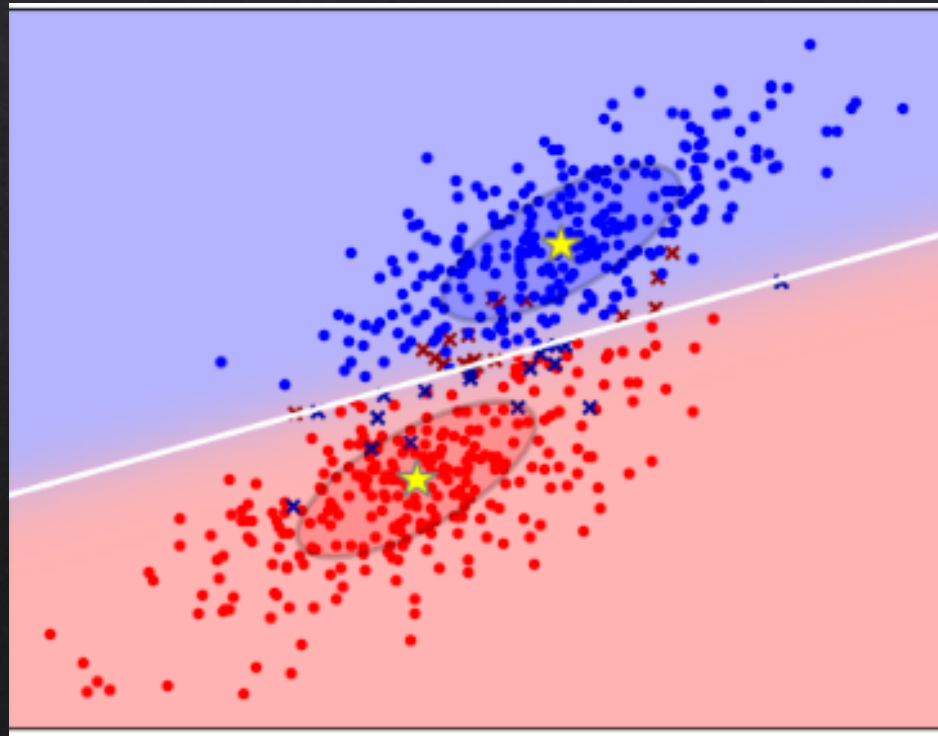
Linear Discriminant Analysis (LDA)

- It is the simplest classifier model in supervised learning.
- In the training stage, It finds an optimum hyperplane with the most separability between classes
- It divides feature space into the different classes.

Feature and Classification

Classifier Models :

Linear Discriminant Analysis (LDA)



Feature and Classification

Classifier Models :

Linear Discriminant Analysis (LDA)

