

| | |
|--------------|--|
| DATE | 1 NOV 2022 |
| TEAM ID | PNT2022TMID32429 |
| PROJECT NAME | Early Detection of Chronic Kidney Disease Using Machine Learning |

PRIOR KNOWLEDGE

DATA SCIENCE

SUPERVISED LEARNING:

- Suppose you had a basket and it is full with some fresher fruits your task is to arrange the same type fruits at one place.
- Suppose the fruits are apple ,banana,cherry,grape.
- So you already know your previous work that, the shape of each and every fruit so it is easy to arrange the same type of fruits at one place.
- Here your previous work is called as train data in data mining.
- So you already learn the things from your train data, this is because of you have a response variable which says you that if some fruit have so and so features it is grape, like that for each and every fruit.
- This type of data you will get from the train data.
- This type of learning is called as supervised learning.
- This type of solving problem comes under classification.
- So you already learn the things so you can do your job confidently.

UNSUPERVISED LEARNING :

- Suppose you had a basket and it is full with some fresh fruits your task is to arrange the same type fruits at one place.
- This time you don't know anything about that fruits , you are first time seeing these fruits so how will you arrange the same type of fruits.

- What you will do first you take on fruit and you will select any physical character of that particular fruit. Suppose you taken colors.
- Then the group will be something like this.
- RED COLOUR GROUP : apple & cherry fruits.
GREEN COLOR AND SMALL SIZE : grapes
- This type of learning is know unsupervised learning.

CLASSIFICATION :

- Classification is a process of categorizing a given set of data into classes, it can be performed on both structured or unstructured data.
- The process starts with predicting the class of given data points. The classes are often referred to as target, label or catagories.

REGRESSION:

- A technique for determining the statistical relationship between two or more variables where a change in a dependent variable is associated with , and depends on , a change in one or more independent variables.
- A regression problem is used when the output variable is a real or continuous value , such as”salary” or “weight”.

LOGISTIC REGRESSION :

- A type of classification algorithm.
- Based on linear regression to evaluate output and to minimize the error .
- Named after the method it uses to evaluate the outputs the logit function.
- Logistic regression just has a transformation based on linear regression hypothesis.
- For logistic regression, focusing on binary classification here , we have class 0 and 1.

- To compare with the target , we want to constrain predictions to some values between 0 and 1.
- That's why SIGMOID FUNCTION is applied on the raw model output and provides the ability to predict with probability.
- Logistic function , also called the sigmoid function was developed by statisticians to describe proper of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment .
- Its an s-shaped curve that can take any real -valued number and map it into value between 0 and 1, but never exactly at those limits: $1/(1+E^{-value})$.

FLASK IN PYTHON :

- Flask is a micro framework for python
- Easy to code
- Easy to configure
- Flask won't make many decisions for you , such as what database to use
- Has an excellent documentation
- RESTful
- Testable

