**Introduction**

**Numpy** :- NumPy is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.

NumPy was created in 2005 by Travis Oliphant. It is an open-source library and you can use it freely. NumPy stands for Numerical Python.

**How to install NumPy :-**

Open command prompt and write the command

**pip install numpy**

Or we can use a python distribution that already has NumPy installed like, **Anaconda, Spyder** etc.

**K-Means Algorithm**

K-Means Clustering is an Unsupervised Learning algorithm, which groups the unlabelled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

**“It is an iterative algorithm that divides the unlabelled dataset into k different clusters in such a way that each dataset belongs only one group that has similar properties.”**

The algorithm takes the unlabelled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm.

**How to install K-Means :-**

To use k-means we have to install scikit-learn and to install scikit-learn open command prompt and write the command

**pip install -U scikit-learn**

**Introduction to Machine Learning**

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so

**How machine Learning Model works:**

A Machine Learning system **learns from historical data, builds the prediction models, and whenever it receives new data, predicts the output for it**. The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately.



CODE:

def extractDominantColor (image, number\_of\_colors=1, hasThresholding=False):

if hasThresholding == True:

number\_of\_colors += 1

img = image.copy()

img = cv2.cvtColor(img, cv2.COLOR\_BGR2RGB) img =img.reshape((img.shape[0]\*img.shape[1]), 3) estimator=KMeans(n\_clusters=number\_of\_colors, random\_state=0)

estimator.fit(img)

colorInformation = getColorInformation(

estimator.labels\_, estimator.cluster\_centers\_, hasThresholding)

return colorInformation

In the function extractDominantColor we pass parameter image,number\_of\_colors and hasThresholding.

First we make a copy of image and then we use cv2 to find the RGB value of the image and then we reshape the image and save this in variable img.

Now we use KMeans algorithm to predict the image so we called fit () to train the model.

In Machine Learning fit () used for training the model.

estimator.fit(img) is used to train the image.