Prior Knowledge

The following concepts are learnt to implement the project:

1. **Supervised and Unsupervised Learning**

Supervised learning is a subcategory of machine learning defined by its use of a ground truth (the actual nature of the problem which is the target of a machine learning model, reflected by the relevant data sets associated with the use case). A machine learning model is a tool or algorithm that has been trained to make decisions based on certain data sets. During supervised learning, these training data sets contain prior knowledge of what the output values of the samples should be.

During training, the model will have both data and its labels available and will attempt to learn a function from the data to the label. The features and output of labels are used to learn a mapping from one to the other using machine learning algorithms. After it has been trained, the model can take in new data and predict a label using its learned function. To evaluate the model, the labels it predicts on a test set of unseen samples are compared to the ground truth labels from the test set.

Unsupervised learning, the counterpart to supervised learning, is defined by its lack of ground truth. Unsupervised learning uses machine learning algorithms to analyze and cluster unlabelled datasets. It is capable of independently discovering similarities and differences in information. Unsupervised learning can take many forms, but the goal remains to learn a natural structure within the data set.

In contrast to supervised learning, the model does not take labels as input. Instead, the model takes in data and will attempt to infer an output. The output takes many forms, such as clusters in the data or features inside an image. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or groups in data.

2. **Regression, Classification and Clustering**

Regression means to predict the value using the input data. Regression models are used to predict a continuous value. It is mostly used to find

the relationship between the variables and forecasting. Regression models differ based on the kind of relationship between dependent and independent variables.

Classification is defined as the process of recognition, understanding, and grouping of objects and ideas into preset categories a.k.a “sub-populations.” With the help of these pre-categorized training datasets, classification in machine learning programs leverages a wide range of algorithms to classify future datasets into respective and relevant categories.

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them.

3. **Artificial Neural Network**

The term "Artificial neural network" refers to a biologically inspired sub-field of artificial intelligence modelled after the brain. An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain. Similar to how a human brain has neurons interconnected to each other, artificial neural networks also have neurons that are linked to each other in various layers of the networks. These neurons are known as nodes.

4. **Convolutional Neural Network**

   A convolutional neural network, or CNN, is a deep learning neural network sketched for processing structured arrays of data such as portrayals.CNN are very satisfactory at picking up on design in the input image, such as lines, gradients, circles, or even eyes and faces. This characteristic makes convolutional neural networks so robust for computer vision. CNN can run directly on an underdone image and do not need any preprocessing. A convolutional neural network is a feed forward neural network, seldom with up to 20. The strength of a convolutional neural network comes from a particular kind of layer called the convolutional layer.

5. **Flask**

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself.