**Machine Learning Algorithm:**

**K-Means Clustering**

K-Means clustering is an unsupervised learning algorithm that, as the name hints, finds a fixed number (*k*) of clusters in a set of data.

A *cluster*is a group of data points that are grouped together due to similarities in their features. When using a K-Means algorithm, a cluster is defined by a *centroid*, which is a point (either imaginary or real) at the center of a cluster. Every point in a data set is part of the cluster whose centroid is most closely located. To put it simply, K-Means finds *k* number of centroids, and then assigns all data points to the closest cluster, with the aim of keeping the centroids small.

**The Algorithm**

K-Means starts by randomly defining *k* centroids. From there, it works in iterative (repetitive) steps to perform two tasks:

1. Assign each data point to the closest corresponding centroid, using the standard Euclidean distance. In layman’s terms: the straight-line distance between the data point and the centroid.
2. For each centroid, calculate the mean of the values of all the points belonging to it. The mean value becomes the new value of the centroid.

Once step 2 is complete, all of the centroids have new values that correspond to the means of all of their corresponding points. These new points are put through steps one and two producing yet another set of centroid values. This process is repeated over and over until there is no change in the centroid values, meaning that they have been accurately grouped. Or, the process can be stopped when a previously determined maximum number of steps have been met.

Here we are considering pin code data from GPS module as data points and forming cluster according to accident prone areas and forming centroids and when we get nearby pin code through GPS module we will activate ultrasonic sensor for monitoring collision condition.