Abstract

During data analysis, often data needs to be grouped together based on similar looking or behaving. As the real world data features modulate with the Big data, where the data is unlabeled, the task of dividing the population or data points into a number of groups with similar points is of prime necessity. This method of identifying similar groups of data in a data set is called clustering. In simple words, the aim is to segregate groups with similar traits and assign them into clusters. This paper presents the importance of the K-means Clustering algorithm to understand the inner structure of the data to obtain the areas wherein based on the number of car rides booked in an area, optimum pickup point can be found using K-Means Clustering Algorithm.

References


**Index Terms**

Computer Science Algorithms

**Keywords**

Clustering, K-Means, Big Data, Pickup optimization