Abstract

Now days, there are more number of corporations are gathering a more number of information, frequently produced incessantly as a series of measures and approaching from different types of positions. Big data defines a knowledge used to record and execute the data set and it has the structured, semi structured and unstructured data that has to be mined for valuable data. On the other hand, mining through the high dimensional data the search space from which an optimal feature subset is determined and it is enhanced in size, guiding to a difficult stipulate in computation. With respect to handle the troubles, the research work is generally based on the high-dimensionality and streaming structure of data feeds in big data, a new inconsequential feature selection methodology that can be used to identify the feature selection methods in the big data. Some of the research work illustrates the different kinds of optimization methods for data stream mining would lead to tremendous changes in big data. This research work is focused on discussing various research methods that focus on finding the efficient feature selection methods which is used to avoid main challenges and produce optimal solutions. The previous methods are described with their advantages and disadvantages, consequently that
the additional research works can be focused more. The tentative experiments were on the entire research works in Matlab simulation surroundings and it is differentiated with everyone to identify the good methodologies beneath the different performance measures.

References


Index Terms

Computer Science

Information Sciences

Keywords

Big Data, Feature Selection, Particle Swarm Optimization, Classification