**ABSTRACT**

In the modern world there is huge development in the field of networking technology which handles huge data at a time. This data can be structured, semi structured or unstructured. To perform efficient mining of valuable information from such type of data the big data technology is gaining importance nowadays. Data mining application is been used in public and private sectors of industry because of its advantage over conventional networking technology to analyze large real time data. Data mining mainly relies on 3 V’s namely, Volume, Varity and Velocity of processing data. Volume refers to the huge amount of data it collects, Velocity refers to the speed at which it process the data and Variety defines that multi-dimensional data which can be numbers, dates, strings, geospatial data, 3D data, audio files, video files, social files, etc. These data which is stored in big data will be from different source at different rate and of different type; hence it will not be synchronized. This is one of the biggest challenges in working with big data. Second challenge is related to mining the valuable and relevant information from such data adhering to 3rd V i.e. Velocity. Speed is highly important as it is associated with cost of processing. 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 Mat lab simulation surroundings and it is differentiated with everyone to identify the good methodologies beneath the different performance measures.