

## **ABSTRACT**

The Big data is based on the 3V challenges that are the Volume, the Variety, and Velocity. Big data is collected from various sources and it is seen that data comes in a various format in high speed that are gathered together rapidly as well as they are created as an ancient batch models where it is infeasible to process in real time. Big Data has become an imminent part of all industries and business sectors today. All organizations in any sector like energy, banking, retail, hardware, networking, etc all generate huge quantum of heterogenous data which if mined, processed and analyzed accurately can reveal immensely useful patterns for business heads to apply to generate and grow their businesses. There is a very critical challenge in big data that demands data mining to be performed with high-speed information in big technology which has been getting a lot of importance today. The technique of feature selection has been employed for the mining of data stream on the fly for the big data to improve efficiency and for minimizing the process load on mining and its model.

In this paper, for achieving an accuracy and to have a minimum processing time for a query and for the reduction of the processing load a Particle Swarm Optimization (PSO), a Grammatical Evolution (GE) and the Hybrid PSO-GE methods has been proposed. The techniques of classification help users in retrieving required information from big database of transactions in a simpler manner. Database management systems have been indispensable to enterprises for decades. As the amount of data dramatically increased, database aggregation has encountered a dilemma between privacy and performance. In traditional database aggregation, all attributes have been

encrypted to protect the privacy of data. The results of the experiments demonstrate the efficiency of the hybrid PSO-GE method compared to existing methods.