**REDUCING SEARCH SPACE USING**

**BIG DATA**

**ABSTRACT**

Many existing data mining algorithms search interesting patterns from transactional databases of precise data. However, there are situations in which data are uncertain. Items in each transaction of these probabilistic databases of uncertain data are usually associated with existential probabilities, which express the likelihood of these items to be present in the transaction. When compared with mining from precise data, the search space for mining from uncertain data is much larger due to the presence of the existential probabilities. This problem is worsened as we are moving to the era of Big data. Furthermore, in many real-life applications, users may be interested in a tiny portion of this large search space for Big data mining. Without providing opportunities for users to express the interesting patterns to be mined, many existing data mining algorithms return numerous patterns—out of which only some are interesting. This paper is about handling large dataset, which is not possible with existing traditional clustering algorithms. In this paper we have used SNN (SHARED NEAREST NEIGHBOUR) algorithm and FUZZY-C algorithm which is an innovative approach for clustering which is one of the most efficient clustering algorithms which can handle most of the issues related to clustering, like, it can generate clusters of different sizes, shapes and densities. This paper is about handling large dataset, which is not possible with existing traditional clustering algorithms. Our algorithm greatly reduces the search space for Big data mining of uncertain data, and returns only those patterns that are interesting to the users for Big data analytics**.**