Globally, it is witnessed that the databases in number and size in terms of data volumes are proliferating in an exponential manner. These databases contain a rich treasure of knowledge based on which an organization may take a strategic decision or initiate a pivotal plan of action. It is also seen that the number of human data analysts grows at a much smaller rate than the amount of data stored. Hence there is a need for automatic methods to extract knowledge from the stored data through Data Mining. The goal of Data Mining is to extract high-level knowledge from low-level data in the context of large data sets.

This thesis provides a comprehensive methodology of extracting useful information from various benchmark and real world datasets using evolutionary computation techniques. Due to the randomized nature of evolutionary computation techniques, they have been very suitable to explore and exploit the search space to extract useful information. In this work, the focus is on data clustering, which is an important task in the process of Data mining. Suitable approaches have been suggested to overcome the limitations of popular K-means approach. Two most popular evolutionary techniques, namely Particle Swarm Optimization (PSO) and Differential Evolution (DE) are used, to develop strategies to overcome difficulties like initial seed value and local optima problem encountered in K-means. Several simulations have been done to show the
effectiveness of PSO and DE to overcome these problems. In this work, a new method of adopting the parameters of DE has been suggested. The new DE known as Adaptive Differential Evolution (ADE) works well when compared to simple DE in clustering and this has been demonstrated with several simulation runs. The thesis also suggests two novel hybridizations of PSO and ADE to exploit the advantage of both the techniques. Not knowing the number of clusters beforehand has been a great challenge for Data Mining researchers for several years. In this work, a new dynamic clustering with ADE has been proposed to tackle this aspect. Results reveal that this approach can predict the number of clusters for several benchmark datasets accurately, compared to other approaches. Finally to demonstrate the effectiveness of the proposed Algorithm, the Dynamic Clustering with ADE is applied to a realistic Transportation dataset. The simulation results shown in this work demonstrate very useful groups with regard to the transportation problem.