"Autonomic Computing" is an emerging paradigm of computing, inspired by autonomous human nervous systems. Autonomic computing was conceived to reduce complexity, and to drive computing into a new era that may better utilize its potential to support higher order thinking and decision making. It refers to the self-managing characteristics of the system resources, adapting to changes while hiding the complexities involved to the users.

Autonomic Computing is a prominent challenge, requiring advances in several fields of science and technology, particularly systems, software architecture and engineering, human-system interfaces, policy, modeling and optimization. Integrating these technologies and embedding them in suitable system architecture so as to achieve the desired self-management properties is a research challenge in itself. The ultimate goal of autonomic computing system is the reduced or almost zero human intervention.

Machine Learning and Data Mining were and still are considered as core enabling technologies for Autonomic Computing. Machine Learning deals with the construction and study of systems that can learn from the available data. It can learn normal and anomalous patterns from training data and generate classifiers and association rules, which can be used to capture the knowledge or information of interest. The core of machine learning is representation and generalization. Some machine learning systems attempt to eliminate the need for human intervention or involvement in data analysis, while others adopt a collaborative approach between human and machine.

The objective of this research is to study and analyze the autonomic components to improve the performance of some database applications. That is, to investigate a high-level approach to controlling and automating the impact of complexity involved in solving and analyzing huge datasets with various parameters, and in various domains. Thus human resources and the available data could be utilized in a much better and sensible manner.
Optimization challenges in autonomic computing overlap a good deal with those of machine learning. Due to either extrinsic effect such as changing workload or intrinsic effect such as adaptative behavior by other autonomic elements, an individual autonomic element's optimization scenario is bound to change over time. Many traditional optimization techniques assume implicitly that the objective function is static, and do not bother to resample it.

The method to address the problem to handle large queries, involves, dividing a large query into an equivalent set of smaller queries and an implementation of the Query disassembler. Execution time taken by the system or the response time for a query increases as the workload in terms of number of queries increases. This is due to the fact that the processor needs to access more number of resources (in terms of tables in the database) parallelly.

Online index selection is designed with the motivation to understand if the query pattern changes over time. The major problem associated with huge databases is indexing and retrieving the frequently viewed products. This is required to increase the performance level which is done with the help of online indexing and minimum support parameter. For these concepts to be realized, apriori algorithm was implemented. It was found that while the minimum support value is increased, the number of frequent itemset is decreased.

In Optimized Association Rule Mining, a genetic algorithm is applied over the rules fetched from apriori association rule mining. Learning association rules is basically, finding the items that are purchased together more frequently than others. Implementation of association rule mining of data using genetic algorithm leads to improved performance of accessing information from log file maintained at server machine. Also from the given data set, using genetic algorithm, more precise and useful information were extracted from the routine data. The study shows how using apriori algorithm and genetic algorithm, the system could be optimized. The analysis was carried out considering the minimum support value as 20%. The information
gathered at the output is useful to identify the values of the number of groups formed based on the frequent items bought, buying pattern of the customers, frequency of buying a particular product, and also the most different pattern of purchase. This becomes the critical information for further policies of decision making. It was observed that with the change in the dataset size from 1000 to 5000, the number of groups changed nominally from 34 to 35. However, the density of the datasets in each group increased drastically. This indicates that larger size of dataset gives more insights about the historical data.

The Decision Tree hybrid classifiers were tested and implemented for three applications i.e., determining the potability of water samples, health check of horses and quality of wine. These benchmark datasets were used for experimentation. The combination of genetic algorithm and decision tree learning system was able to outperform the decision tree without genetic algorithm. This was due to the fact that the hybrid approach is able to focus on relevant features and eliminate unnecessary or distracting features. The maximum training set accuracy achieved by using a decision tree built with GA based feature was 98.67% for water dataset sample and 100% and 99.33% for horse and wine dataset sample respectively. It was found that on an average, the difference in the accuracy level was in the range of 3-15% for different sample data sets. Here the decision maker also has an option and flexibility to set or choose the features of the sample, test set accuracy and training set accuracy as expected or desired.

Thus, this thesis work has successfully demonstrated the improvement in classification and association rule mining algorithms for various examples, using the two main components of autonomic systems i.e., machine learning and data mining. This improvement in turn will lead to better realization of autonomic computing systems in future.