CHAPTER 7

SUMMARY OF FINDINGS AND CONCLUSIONS

7.1 INTRODUCTION

The main objective of the research work is to develop data classifier model using Fuzzy Logic which can be used as a helping tool for the user in the decision making process. The basic characteristic of the design process such as the accuracy, comprehensibility, interpretability, compactness, robustness, versatility, modifiability, and coherence with previous knowledge, are considered due to the fact they may be useful to allow the system to be accepted for use.

In the preceding chapters, a research study carried out on the Intelligent Optimization Techniques for Fuzzy Logic Based Classification is presented. The purpose of this concluding chapter is to:

i) give a summary of findings of the research study.

ii) list out the major contributions of study.

iii) suggest the possible extensions for future work in this area of research.

The details are presented in the following sections.
7.2 SUMMARY OF THE RESEARCH FINDINGS

Formation of if-then rules and the membership functions are the two major tasks of designing a classifier model using Fuzzy Logic. In this research work, the design of a fuzzy classifier system is formulated as a search problem in the solution space where each point represents a rule set, membership function and the corresponding system behavior. Then an optimization algorithm is applied to search for an optimal location of this solution space which hopefully represents the near optimal solution.

Genetic Algorithm is proposed as tool for optimal design of fuzzy classifier model. The conventional Binary-coded GA has Hamming cliff problem which sometimes may cause difficulties in the case of coding continuous variables. Also, for discrete variables with total number of permissible choices not equal to $2^k$, it becomes difficult to use a fixed length binary coding to represent all permissible values. To address these issues, Real- coded GA using Tournament selection, BLX-$\alpha$ crossover and Non-uniform mutation is proposed to design the fuzzy classifier.

While GA can rapidly locate good solutions, even for difficult search spaces, it has the tendency to converge towards local optima, if the fitness function is not defined properly. For specific optimization problems, and given the same amount of computation time, simpler optimization algorithms may find better solutions than GA. Operating on dynamic data sets using GA is a difficult tedious task. To overcome this difficulty, Particle Swarm Optimization approach is proposed to obtain the optimal rule set and the membership function for a fuzzy classifier.

Even though PSO produces improved result than RGA, premature convergence and local optima are the two problems that make PSO to perform poor for complex data. For better classification of complex data, two kinds of
hybrid GA-PSO approaches are proposed by combining the strength of GA and PSO. The first one proposes a modified form of PSO based mutation for Genetic Fuzzy System and the second one incorporates BLX-α crossover and Non-uniform mutation into Swarm Fuzzy System.

Fuzzy system design usually comes with two contradictory requirements; namely Accuracy and Interpretability. It is often found that a fuzzy system with highest classification accuracy shows poorer interpretability and vice versa. The IGA approach that has reasonable interpretability and produces compact rule set shows inferior classification accuracy than EPSO. Controversially the EPSO approach is better in accuracy but poorer in interpretability and compactness than IGA.

In view of this interpretability-accuracy tradeoff, a novel Water Swirl Algorithm (WSA) inspired by the swirl motion of water inside a sink that searches for the drain is proposed for obtaining accurate and interpretable fuzzy classifier model.

When the drain of the sink is opened, a swirling motion is started in the water mass near the drain that leads to an important phenomenon called vortex and result in release of water through the drain. According to vortex particle theory, every water particle in the vortex has position and strength that interacts with each other and evolves during swirl motion. Based on this fact, suitable update equations are proposed to locate the optimum solution iteratively from the initial randomly generated solution space.

The performances of all the proposed approaches are tested using ten benchmark datasets available in the UCI machine learning repository. All the proposed approaches are developed using Matlab R2009 and executed in a PC with Pentium IV processor with 2.40 GHz speed and 256 MB of RAM. Experiments are conducted to examine both the learning ability as well as the
generalization ability of all the proposed approaches in the formation of rule base and membership function.

In order to visualize the performance of all the proposed approaches, and to assess whether significant differences exist among their results, ROC analysis and Wilcoxon’s signed-rank test are employed. All these test results show that the decision boundary is continuously improved to cover the samples of overlapping classes by various approaches and the one formed by WSA covers more number of samples accurately with highly interpretability and compact rule set.

### 7.3 SIGNIFICANT RESEARCH CONTRIBUTIONS

The main contribution of this research work is a developing data classifier model using Fuzzy Logic. A typical fuzzy logic based classification system consists of collection of rules (Rule Base) and membership function for each input variable (Data Base) that are used as a Knowledge Base upon which quantitative reasoning is performed to derive the class label.

As against the human experts and data driven approaches, an optimization approach makes the fuzzy classification system as a self learning system. In this respect, the following are the major contributions of this research work.

1. Real-coded Genetic Algorithm for optimal fuzzy classifier design.
2. Particle Swarm Optimization approach for improved performance of fuzzy logic based classification system.
3. Hybrid GA-PSO approach for classification of complex datasets.

5. Water Swirl Algorithm for design of accurate and interpret Table fuzzy classifier model.

Each one of the above approach contributes a distinct methodology for addressing the problems in its domain and this is done in a cooperative, rather than a competitive, manner.

7.4 CONCLUSION

Data classification is one of the most important Data Mining tasks that are required in most of the Engineering domains for in depth analysis of the data. Among the various approaches available in the literature, this research works uses the concept of Fuzzy Logic for designing a data classification system since it provides a facility to model uncertainty, the human way of thinking, reasoning and making decisions in complex situation.

The design of fuzzy logic based classification system is formulated as an optimization problem and various novel optimization algorithms are proposed to obtain the near optimal rule set and membership function. Real-coded Genetic Algorithm, Particle Swarm Optimization, Intelligent Genetic Algorithm, Enhanced Particle Swarm Optimization and Water Swirl Algorithm are the proposed optimization approaches for designing the optimal fuzzy classifier model.

All these approaches focused on developing data classifier model using Fuzzy Logic that can be used as a helping tool for the user in the decision making process. Each approach follows its own individual methodology for tackling the issues in its field and this is carried out in a supportive rather than an aggressive manner. The result is a more intelligent
and robust classifier system providing an accurate, compact and human-interpretable solution as compared to traditional techniques.

7.5 SUGGESTION FOR FUTURE WORK

The objective of the proposed fuzzy classification is to form an optimal class separating boundary in the form of if-then rules and membership function that must be appropriately account for unseen data. In view of the accomplishing this objective, novel population based stochastic optimization algorithms are proposed that works intelligently for optimal design of fuzzy classifier. The possible extensions of the present work are listed below:

- Construction of a fuzzy system with good generalization ability in an imbalanced feature space by including Boosting/Bagging techniques.

- Design of adaptable fuzzy system for dealing with changing external environments by incorporating the learning concept of ANN or SVM.

- Development of effective fuzzy classifier for high dimensional data using Dimensionality reduction techniques.

- Devising of new Biological/Nature inspired optimization algorithms for improved performance over plenty of datasets.

- Multi objective approaches to alleviate the interpretability-accuracy trade-off in fuzzy classifier design.

- Addressing the sub problems of classification namely association rule mining, subgroup discovery and pattern identification.