List of Tables

2.1 A summary of common operations on fuzzy sets .................. 8
2.2 Basic structure of a GA ............................................. 11

3.1 Characteristics of the Partitional Clustering Algorithms ............ 36
3.2 Characteristics of the Hierarchical Clustering Algorithms .......... 37
3.3 Characteristics of the Grid-Based Clustering Algorithms .......... 38
3.4 Characteristics of the Density-Based Clustering Algorithms ........ 39

4.1 Algorithm for k-means clustering .................................... 53
4.2 Algorithm for fuzzy c-means clustering ............................. 57
4.3 GA based clustering Algorithm ..................................... 59
4.4 Algorithm for KSOM clustering ..................................... 65
4.5 Comparison of objective function value obtained by the k-means, FCM and the KSOM algorithm for the Iris and MS data set for varying number of clusters ........................................ 69
4.6 Comparison of cluster sizes obtained by the k-means, FCM and the KSOM algorithm for the Iris and the Diabetes data sets ........ 69
4.7 Execution time for the $k$-means, FCM, GA based FCM and the KSOM algorithm for the Iris data set.  
4.8 Number of clusters obtained by the $PBM$, $XB$, $DUNN$, $SC$ and the $SI$ validity indices using $k$-means clustering algorithm.  
4.9 Values of the $PBM$, $XB$, $DUNN$, $SC$ and the $SI$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Cancer, Iris and the Diabetes data sets using the $k$-means algorithm.  
4.10 Values of the $PBM$, $XB$, $DUNN$, $SC$ and the $SI$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Data_3_2, Data_4_3 and the Data_5_2, data sets using the $k$-means algorithm.  
4.11 Values of the $PBM$, $XB$, $DUNN$, $SC$ and the $SI$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Data_6_2 and the Data_9_2 data sets using the $k$-means algorithm.  
4.12 Number of clusters obtained by the $PBMF$, $XBF$, $PC$ and the $MPC$ validity indices using the FCM and GA based FCM clustering algorithm.  
4.13 Values of the $PBMF$, $XBF$, $PC$ and the $MPC$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Cancer and the Iris data sets using the FCM and the GA based FCM algorithm with $m = 2$.  
4.14 Values of the $PBMF$, $XBF$, $PC$ and the $MPC$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Diabetes and the Data_3_2 data sets using the FCM and the GA based FCM algorithm with $m = 2$.  
4.15 Values of the $PBMF$, $XBF$, $PC$ and the $MPC$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Data_4_3 and the Data_5_2 data sets using the FCM and the GA based FCM algorithm with $m = 2$.  

viii
4.16 Values of the $PBMF$, $XBF$, $PC$ and the $MPC$ validity indices in the range of $k = 2, 3, \ldots, 10$ for the Data_6_2 and the Data_9_2 data sets using the FCM and the GA based FCM algorithm with $m = 2$.

4.17 Comparison of the clustering techniques.

5.1 Algorithm to partition the original data set to form the simplified data set.

5.2 Result of the Data_6_2 data set for the FCM, $psFCM$ and the $pshFCM$ clustering algorithms with optimal cluster center $c = 6$, $m = 2.0$ and the number of split in the $k-d$ tree $= 6$.

5.3 Result of the Data_4_2 data set for the FCM, $psFCM$ and the $pshFCM$ clustering algorithms with optimal cluster center $c = 6$, $m = 2.0$ and the number of split in the $k-d$ tree $= 6$.

5.4 Result of the Data_4_3 data set for the FCM, $psFCM$ and the $pshFCM$ clustering algorithms with optimal cluster center $c = 4$, $m = 2.0$ and the number of split in the $k-d$ tree $= 6$.

5.5 Result of the Data_4_3 data set for the FCM, $psFCM$ and the $pshFCM$ clustering algorithms with optimal cluster center $c = 4$, $m = 2.0$ and the number of split in the $k-d$ tree $= 6$.

6.1 Number of clusters $c$ obtained by the validity indices for different data sets using the FCM clustering algorithm with $m = 1.2$.

6.2 Values of the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Iris and the Diabetes data sets using the FCM clustering algorithm with $m = 1.2$. 

ix
6.3 Values of the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Data_3-2 and the Data_4-3 data sets using the FCM clustering algorithm with $m = 1.2$.

6.4 Values of the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Data_5-2 and the Data_6-2 data sets using the FCM clustering algorithm with $m = 1.2$.

6.5 Number of clusters $c$ obtained by the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Data_3-2 and the Data_6-2 data sets using the FCM clustering algorithm with $m = 2.0$.

6.6 Values of the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Data_3-2 and the Data_6-2 data sets using the FCM clustering algorithm with $m = 2.0$.

6.7 Values of the validity indices in the range of $c = 2, 3, \ldots, 15$ for the Elliptical_10.2 data set using the FCM clustering algorithm with $m = 2.0$.

6.8 Number of clusters $c$ obtained by the validity indices for the DataX1, DataX2, DataX3 and the DataX1Ndata data sets using the FCM clustering algorithm with different values of $m$.

6.9 Values of the validity indices in the range of $c = 2, 3, \ldots, n - 1$ for DataX1 and DataX2 data sets using the FCM clustering algorithm with $m = 1.2$.

6.10 Values of the validity indices for DataX3 and DataX1N data sets using the FCM clustering algorithm with $m = 1.2$. 