Contents:

1. Introduction page 1

    1.1 Phases of gene expression data analysis page 1

    1.2 Existing Approaches page 2

    1.3 Motivation page 2

    1.4Aim of our work page 2

    1.5 Objectives of our work page 3

    1.6 Organization of the Dissertation page 3

2. Background of the work page 4

    2.1 Different types of clustering page 4

    2.2 Gene-based clustering page 5

    2.3 Different types of gene-based clustering page 6

    2.4 Different types of clusters page 6

    2.5 Proximity measures for gene expression data page 8

    2.6 Cluster validity measures page 10

3. Existing work done on clustering page 14

    3.1 K-means page 14

    3.2 Hierarchical Clustering page 15

    3.3 Density-Based Hierarchical Clustering page 16

    3.4Model-Based methods page 16

3.5 Graph-Based Clustering page 17

3.6 Nearest Neighbor Clustering page 19

3.7 Fuzzy-Clustering page 19

4. Proposed approach page 22

4.1 Link-Based Clustering (LBC) page 22

4.2 Concepts and Definitions in LBC page 22

4.3Fuzzy-Link Based Clustering (F-LBC) page 27

4.4 Concepts and Definitions in F-LBC page 27

5. Performance Evaluation page 32

5.1 Performance Evaluation page 32

5.2 Dataset Description page 32

5.3 Results page 33

6. Conclusion and Future work page 46

References page 47