

Contents

Contents	viii
List of Figures	ix
List of Tables	x
1 Introduction	1
1.1 Project profile	2
1.1.1 Project title	2
1.1.2 Person Involved	2
1.1.3 Guide	3
1.1.4 About the Organization	3
2 Motivation	4
2.1 Problem definition.	4
2.1.1 Problem objectives.	4
2.2 Related Work	4
2.2.1 Apriori	4
2.2.2 FP-Growth	5
2.2.3 Other FP-Growth based approach	6
3 Proposed Algorithm	8
3.1 Build the SBFP-Tree :	9
3.2 Mining the SBFP-Tree :	14
4 Experimental Results	17
4.1 Environment Used	17
4.2 Dataset Used	17
4.3 Discretization of dataset	18
4.4 Results	19
5 Conclusion and future work	23

List of Figures

3.1	Tree after TID=1	9
3.2	Linked list after TID=1	10
3.3	Linked list after TID=2	10
3.4	Tree after TID=2	11
3.5	Tree after adding node 'c'	11
3.6	Linked list after adding node 'c'	12
3.7	Tree after TID=3.	12
3.8	Linked list after TID=3.	12
3.9	Tree after TID=4.	13
3.10	Linked list after TID=4.	13
3.11	Tree after TID=5.	13

List of Tables

3.1	Trasaction dataset	8
3.2	<i>NodeTable</i> after TID=1	10
4.1	Description of Datasets	18
4.2	File size of different datasets(Market Basket)	19
4.3	No. of nodes generated in different Datasets	20
4.4	SBFP-Tree file size in secondary storage for datasets in Market Basket(Equal Width Interval Binning)	20
4.5	SBFP-Tree file size in secondary storage for datasets in Market Basket(Recursive Minimal Entropy Partitioning)	21
4.6	Percentage of compression after the market basket datasets being converted to SBFP-Tree (allowing 100% of the total nodes created to be in physical memory)	22