Chapter - 7
Observations and Findings

7.1 Experimental Observations

A detailed discussion of the performance of the proposed techniques is presented in this chapter. The results are obtained from the experimental results sections of the chapters 4, 5 and 6. In this chapter, a performance comparison of the proposed approaches is presented.

7.1.1 Page Blocks Classification Data Set

A. Reading Cost

The reading cost of the frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree and the Fuzzy Decision Tree on Page Blocks Classification dataset is provided in fig. 7.1. It is clearly observed from the fig. 7.1 that the proposed frequent itemset mining using Fuzzy Decision Tree provides very low reading cost when compared to frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree.
Fig. 7.1: Comparison of Reading Cost in Page Blocks Classification Data Set.

B. Frequent Itemset Extraction Time

The frequent itemset extraction time taken by the three proposed techniques for the Page Blocks Classification dataset is illustrated in fig. 7.2. It clearly shows that the proposed frequent itemset mining using Fuzzy Decision Tree takes very less time when comparing with the other two proposed approaches.
Fig. 7.2: Comparison of Frequent Itemset Extraction Time in Page Blocks Classification Data Set.

C. Memory Consumption

Fig. 7.3 reports the peak main memory needed by the three proposed frequent itemset mining techniques on Page Blocks Classification dataset. For all the three approaches, the buffer cache (with default size 512 Kbytes) is included in the memory held by the process. Since the Fuzzy Decision Tree is used in the third proposed approach, it consumes very low main memory space.
7.1.2 Letter Recognition Dataset

A. Reading Cost

The reading cost of the frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree and the Fuzzy Decision Tree on Letter Recognition dataset is provided in fig. 7.4. It is clearly observed from the figure that the proposed frequent itemset mining using Fuzzy Decision Tree provides very low reading cost when compared to frequent itemset mining Prefix Hash Tree, FP-Bonsai Tree.
Fig. 7.4: Comparison of Reading Cost in Letter Recognition Dataset.

B. Frequent Itemset Extraction Time

The frequent itemset extraction time taken by the three proposed techniques for the Letter Recognition dataset is illustrated in fig. 7.5. It clearly shows that the proposed frequent itemset mining using Fuzzy Decision Tree takes very less time when comparing with the other two proposed approaches.
Fig. 7.5: Comparison of Frequent Itemset Extraction Time in UCI Letter Recognition Dataset.

C. Memory Consumption

Fig. 7.6 shows the utilization of peak main memory by the three proposed techniques for the Letter Recognition dataset. It clearly shows that the proposed frequent itemset mining using Fuzzy Decision Tree utilizes very low main memory space when comparing with the other two proposed approaches.
Fig. 7.6: Comparison of Peak Main Memory Utilization in UCI Letter Recognition Dataset.

7.1.3 Census Data (1990)

A. Reading Cost

The reading cost of the frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree and the Prefix Hash Tree, FP-Bonsai Tree and the Fuzzy Decision Tree on Census Data (1990) dataset is provided in fig. 7.7. It is clearly observed from the figure that the proposed frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree and the Fuzzy Decision Tree provides very low reading cost when compared to frequent itemset mining using Prefix Hash Tree, FP-Bonsai Tree.
Fig. 7.7: Comparison of Reading Cost in US Census Data (1990) Dataset.

Frequent Itemset Extraction Time

The frequent itemset extraction time taken by the three proposed techniques for US Census Data (1990) Dataset is illustrated in fig. 7.8. It clearly shows that the proposed frequent itemset mining using Fuzzy Decision Tree takes very less time when comparing with the other two proposed approaches.
**Fig. 7.8: Comparison of Frequent Itemset Extraction Time in US Census Data (1990) Dataset.**

*Main Memory Utilization*

Fig. 7.9 shows the utilization of peak main memory by the three proposed techniques for US Census Data (1990) Dataset. It clearly shows that the proposed frequent itemset mining using Fuzzy Decision Tree utilizes very low main memory space when compared with the other two proposed approaches.
Fig. 7.9: Comparison of Peak Main Memory Utilization in US Census Data (1990) Dataset.

7.2 Overall Ranking Based On Performance

It is clear from the experimental results that the performance of the proposed approaches are better in terms of reading cost, frequent itemset extraction time and memory consumption than the existing frequent itemset mining using Prefix Hash Tree and FP-Bonsai Tree. The following section provides the ranking of the proposed approaches based on the above mentioned parameters for three datasets.
Table 7.1

Overall Ranking Performance of the Frequent Itemset Mining Approaches in all Datasets

<table>
<thead>
<tr>
<th>Proposed Approach</th>
<th>Reading Cost</th>
<th>Frequent Itemset Extraction Time</th>
<th>Main Memory Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent Itemset Mining using Prefix Hash Tree</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>FP-Bonsai Tree</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fuzzy Decision Tree</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.1 shows the overall ranking performance of the proposed approaches. This overall ranking is based on the experimental results of all the performance parameters. It is observed from the table that the third proposed frequent itemset mining using Fuzzy Decision Tree ranked first in the overall performance as it outperforms the other approaches.