Chapter 7

CONCLUSION

This thesis has attempted to cover most of the activities of the rapidly growing area of Web usage mining. The proposed framework "Online Miner" works well for developing prediction models by inspecting web traffic volume at a deeper level than any other web mining tool.

With the growth of web-based applications especially electronic commerce, there is a significant interest in analyzing web usage data for understanding the Web Usage, and apply same knowledge to serve the users. As the web mining area is growing fast, there is lot of demand for web usage mining and there is a need to develop a common framework like J2EE and .NET. Similar process models or framework needs to be developed for creating an interest among the new researchers or business strategies and developers. We need a systematic web-site design methodology to create new web pages, or modify existing web pages, such that different user's navigational patterns could be better mapped to answers to set of specific questions.

This has led to number of commercial offerings for doing such analysis. There is need to develop tools, which incorporate statistical methods, visualization, and human factors to help better understand the mined knowledge. Since the output of knowledge mining algorithms is often not in a form suitable for direct human consumption, there is a need to develop techniques and tools for helping an analyst to understand it. One of the open issues in data mining, in general and Web mining,
in particular, is the creation of intelligent tools that can assist in the interpretation of mined knowledge. Clearly these tools need to have specific knowledge about the particular problem domain to do any more than filtering based on statistical attributes of the discovered rules or patterns.

The system developed Online Miner is a prototype to demonstrate an application example for the topic of general web usage mining together with an application of a data mining algorithm with a database approach. Through the development of this system it is clearly seen the steps of a web mining system architecture; data preparation, data description, application of data mining techniques, knowledge extraction, evaluation of the results. This experience also revealed all the difficulties mentioned in the previous research studies such as the data being abundant, complex and unstructured.

Perhaps the greatest benefit of Online Miner is its ability to provide many features with suitable architecture and mining repository algorithm. The illustrated experimental results are proved to be effective and efficient.

Finally, we conducted several experiments that demonstrate the following

• High accuracy in generating quick reports.
• Capabilities of Online Miner in capturing meaningful clusters in real data by setting a rule set.
• Scalability of the mining repository algorithm.
• High precision of classification of data when data filters are applied.
However, Web Usage mining raises some hard scientific questions that must be answered before robust tools can be developed. This framework Online Miner has aimed at describing such challenges and hope research community will take up the challenge of addressing them.

This study provides vista for future research. Since the study has been carried out very meticulously, there is a substantial scope for conducting and exploring facts pertaining to the area, which may be quite useful in Web Usage Mining.

More research needs to be done in e-Commerce, bioinformatics, computer security, Web intelligence, intelligent learning, database systems, Finance, marketing, Healthcare, and telecommunications by using web usage mining and Web usage patterns and data mining will be the basis for a great deal in future research. There is need to develop tools, which incorporate statistical methods, visualization, and human factors to help better understand the mined knowledge. Since the output of knowledge mining algorithms is often not in a form suitable for direct human consumption, there is a need to develop techniques and tools for helping an analyst to understand it.