CHAPTER 6

CONCLUSION AND FUTUREWORK

6.1 CONCLUSION

Unlike conventional data, web pages typically contain a large number of information that is not part of main data. Such information is considered as irrelevant to the main data of the web page in this work. This research work first identify the important information of the webpage in a specific domain and it has proposed a technique for extracting the necessary information and making the extracted information in sequence using two systems and an intelligent agent. Adopting agent methodology with the extraction system will increase intelligence for searching and analysis of information.

The agent-based view provides a powerful repertoire of tools, techniques that have the potential to considerably improve the way in which people conceptualize and implement many types of application.

The first system which is information gathering system gathers the required data from the documents that are obtained from the web search engine. Java based programming tools provide a multi platform facility for supporting any kind of search engine. So extraction process starts with the consideration and performance of search engine.
The required data is transferred to the second system which is extraction system using data transfer agent to avoid data loss and make the data transmission secured. Extraction system depends on the result from information gathering system and it uses agent for communicating and transferring the data from information gathering system. The extraction system arranges the extracted data in three sequences such as locative disease pattern, disease affecting sequence and disease spreading sequence. These three sequences are mainly considered as the parameter of finding the significance of disease in the selected region. The sequence of information would be helpful to understand the entire details about the topic. The performance of the technique is evaluated by setting different threshold values based on the extraction of number of patterns, time and on memory utilization. It is shown by plotting the values in graph with different parameters.

6.2 FUTURE WORK

The information extraction could extract the complete information within the domain only. So the domain has to be considered for the development of extraction system. This work can be extended as the development of extraction system with independence of any domain. It could be achieved by the multi agent system framework by simplifying and automating the complexity of distributed computing environment.