











classify the Web-pages through just estimating the relationship within the learned type of classifier and also some latest Web-pages. This may have various issues in order to recover those complexities by applying the Hybrid type of GA or PSO approaches.

### 5.6 Firefly Algorithm:

It is a Meta heuristic approach that is inspired through the glowing nature of fireflies [6]. To attract other fireflies is the main aim of firefly's flash. In this analysis formulating the algorithm through supposing situations like entire fireflies are unisex, such that anyone firefly may attract the entire type of fireflies. Brightness make them attractive accordingly, and also for the any two types of fireflies if one is less glowing then one will attract (and thus move) to the brighter one; here, distance increases makes decreases of brightness. And in case no single fireflies are more glowing as compare to the given firefly then this may be randomly move.

### 5.7 PSO algorithm:

In the PSO algorithm, the birds in a flock are shown as particles in n dimension. Best fitness value of particle at a location in the n-dimensional problem space represents one solution for the problem. When a particle updates its position, another problem solution is generated and then new solution is evaluated by fitness function and the process is repeated until a stopping criteria is met. PSO has a random population matrix like the GA, but the rows in the matrix are called particles instead of chromosomes. Particles are potential solutions that move in a particular direction on the cost surface with a certain velocity. Particles update their positions and velocities using formulas based on the knowledge about the best solution achieved by each particle in its movements (i.e., personal best) and by the complete swarm of particles (i.e., global best).

## 6. Literature Review

In this paper [11] suggested an effective approach for the WPC technique of data-mining. This approach may be very efficient for the training-set which is the set following the way so that it may produce maximum sets. However the results of experiments of this analysis are also influencing so that this will be efficient in case it is applied with the huge sized data-sets along with the maximum type of classes. Also the available approach may needs to be some data used for the training also less time is required for computation of these approaches.

Within this paper [12] a significant amount of patterns can be retrieved through applying the approaches of DM in order to derive the information among the data from Web. Various types of approaches for DM have also been suggested within the past some years. And these approaches may consist of the frequent-item-set mining process, ARM, sequential-pattern-mining process, closed-pattern-mining, etc. Though, here presented the way to efficiently apply these types of discovered patterns is still an unsolved problem. Another typical issue is that only the statistic properties are used while evaluating the effectiveness of patterns. A comprehensive comparison of data mining approaches is

implemented for the Web-mining purpose is performed in this analysis. And the results of experiment have represented that the closed-pattern approaches like the SCPM or the NSCPM may have perform more efficiently because of the using pruning technique within the stages of pattern discovery process.

This paper [13] gave an idea about web mining and how it can be utilized in an efficient way to improve the business. Customer behavior is very important for an organization to enhance the way of providing information to attract them. Analysis of significant information will be helpful for organization to develop promotions that are more effective, internet accessibility, inter – company communication, structure and productive marketing skills through the web-usage type of mining process. Pattern extraction and the web-content type of mining process are the best tools to know the customer and web behavior.

According to this paper [14] suggested the semantic type of web features that may enable to add structure to the Web, while Web Mining can learn implicit structures. And the merged domain of the semantic type of web-mining process may enable the latest approaches to be applied in order to enhance both the domains. The semantic approach of web-mining may involve the integration of the concept of knowledge within the process of web-mining. Domain knowledge can be integrated within the process of web-mining in three ways- domain ontology acquisition, knowledge base construction, and knowledge-enhanced pattern discovery. And obtained benefits of research within various domains of industry like health-care, e-activities, privacy, knowledge-management, security, etc have presented here.

Within this paper [15] describe the issues of the widely used domain over the big-RDF data and also suggested here a latest summary-oriented solutions. This analysis may provides a brief summary over the level types from the RDF type of data while evaluation of query and this represented the summary in order to reduce the required area of the RDF data among the search region and also it formulate the queries of SPARQL for effectively using the data. Additionally, the suggested analysis may be regularly get upgraded with the updates in data. And the experiments performed over the RDF-benchmark and also over the real datasets of RDF have represented that solution provided here is more effective, portable over the RDF type of engines, scalable, etc.

In this paper [16] suggested web-content oriented mining process as a component of Data Mining. Whenever in this discussed about data, and conclude that there is a vast range of data over WWW. And to manage this vast range of data, also often need many tools that can retrieve the data as per the criteria. There are various tools available on the internet which mines the data according to their types like whether they are in a structured format or the semi-structured type data or may be un-structured type of data. Here also discuss the project work which has an ability to mine the data from web efficiently. In this paper, the previous type of approaches of web-content oriented mining process will also be discussed and the tree structure of a webpage document will also be discussed.

This paper [17] has presented the approach of web-content based process of mining to be applied in order to derive the properties of the product and this may labels the attributes within the obtained result. Also the labeling process enables to recognize and also providing naming to the attributes when process of information-retrieval is completed. Then the information-gained may be applied here for analyzing the product along with its explorations. The approach of web-content type of mining is just the merging of the data among several types of resources through analyzing the views of customers. And this paper has also represents the analysis over the suggested concept applied for the mining and its applications for mining. Here also represented few growing approaches that are applied for deriving the data from the various online shopping portals.

In this paper [18] describe an analysis regarding the domain of Web-structure based mining concepts applied for structure and view of data. Here points some confusion in between the DM approach and the web-mining approach. Web data is growing at a significant rate. Web Mining is fertile area of research. Many Successful applications exist. In this also suggest the subtask of web mining & future of the WM process. Now also work for the process mining and try to combine the usage based mining process along with the structure oriented mining processes. In this analysis also go for the mining from cloud. Whenever work on mining over cloud computing that time researchers may hesitate for the cost but that come very less by cloud mining. So, also say that cloud mining can be seen as future of web mining.

In this paper [19] suggested an efficient method to address some of the problems during web content extraction. In the proposed method extract the required signatures through eliminating the noise which is present in the web document. Proposed method shows better performance when compared with existing methods. In future also plan to extend the work to construct DOM tree (Graphical representation) after extraction of useful patterns. Here suggested a new method for web data extraction. It has three phases. In the first phase list of web documents are selected, second phase documents are preprocessed, in the final phase results are presented to users. Experimental results are compared with existing methods. Performance of proposed system is better than existing methods.

In this paper [20] describe a definition of web mining, research direction and benefits of the web-mining process along with its taxonomies. Here identified some of the issues and problems in this area which can needs further analysis and development. Web mining is applied to various fields E-Commerce, Information filtering, Fraud detection Education and research.

In this paper [21] have analyzed the approach of web based mining of data that is the implementation of techniques of DM in order to derive the knowledge from the data of Web which may contains hyperlinks within the documents, some text documents, logs of usage of the web-sites, and so on. Now in today's advanced world, web becomes an important part of many of all organizations, businesspersons and daily individuals. As a web data is of very many different formats, we have studied the characteristics of web data. As it is very much important to mine particular data from web, we have

studied two effective techniques to mine this big data one along with the apache based Hadoop-Map-Reduce approach and second with visualization based technique called as Visual Web Mining.

## 7. Conclusion

Within this paper it is concluded with briefly describing the basics of computer technology with its contribution within various domains of mining of data over web along with its various types and also highlights few promising regions of the future analysis. In this paper given the brief description about the web-structure type of mining process (WSM) along with its various functioning and also explains types. This paper have described about the various techniques of the web- mining concept. This type of mining approaches has been proved very useful in the business world. The analysis have also discusses the techniques used for extracting information from different types of data available in the internet and how this extracted data can be used for mining purposes.

## References

- [1] T. Sunil Kumar, Dr. K. Suvarchala, "A Study: Web Data Mining Challenges and Application for Information Extraction", IOSR Journal of Computer Engineering (IOSRJCE), Vol 7, Issue3, Nov-Dec 2012, pp 24-29.
- [2] D. Jayalatchumy, Dr. P. Thambidurai, —Web Mining Research Issues and Future Directions –A Surveyl, OSR Journal of Computer Engineering (IOSR-JCE), Volume 14, Issue 3 (Sep. -Oct. 2013), PP 20-27.
- [3] Pradnyesh Bhisikar, Prof. Amit Sahu, —Overview on Web Mining and Different Technique for Web Personalization l, International Journal of Engineering Research and Applications (IJERA) Vol. 3, Issue 2, March - April 2013, pp.543-545.
- [4] Darshna Navadiya, Roshni Patel, " Web Content Mining Techniques-A Comprehensive Survey", International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 10, December- 2012, pp.1-6
- [5] S. Balan, P. Ponmuthuramalingam, " A study of various techniques of Web Content Mining Research Issues and Tools, International Journal of Innovative Research and Studies (IJRIS), Volume 2, Issue 5, May 2013, ISSN: 2319-9725.
- [6] R. Lokeshkumar1, R. Sindhuja2, Dr. P. Sengottuvelan, "A Survey on Pre-processing of Web Log File in Web Usage Mining to Improve the Quality of Data" International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 8, August 2014
- [7] Abdelhakim Herrouz, Chabane Khentout, Mahieddine Djoudi, "Overview of Web Content Mining Tools", The International Journal of Engineering And Science (IJES), Volume 2, Issue 6, 2013.
- [8] Rosli Omar; Abu Osman Md Tap; Zainatul Shima Abdullah, "Web usage mining: A review of recent works", Information and Communication Technology for The Muslim World (ICT4M), 2014 The 5th International Conference.

- [9] Ananthi .J, ” A Survey Web Content Mining Methods and Applications for Information Extraction from Online Shopping Sites”, International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014
- [10] Amit Pratap Singh<sup>1</sup>, Dr. R. C. Jain <sup>2</sup>, ” A Survey on Different Phases of Web Usage Mining for Anomaly User Behavior Investigation” International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 3, Issue 3, May – June 2014 ISSN 2278-6856
- [11] Keyur J. Patel<sup>1</sup>, Ketan J Sarvakar, “Web Page Classification Using Data Mining”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 7, July 2013
- [12] Sheng-Tang Wu and Yuefeng Li, “Pattern-Based Web Mining Using Data Mining Techniques”, International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 3, No. 2, April 2013
- [13] Abdul Rahaman Wahab Sait<sup>1</sup> and Dr.T.Meyappan<sup>2</sup>, “WEB MINING – A CATALYST FOR E-BUSINESS”, Advanced Computing: An International Journal ( ACIJ ), Vol.3, No.6, November 2012
- [14] V. Sitha Ramulu, Ch. N. Santhosh Kumar, K. Sudheer Reddy, “A Study of Semantic Web Mining: Integrating Domain Knowledge into Web Mining”, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012
- [15] V. A. Chakkarwar and Amruta A. Joshi, “Semantic Web Mining using RDF Data”, International Journal of Computer Applications (0975 – 8887) Volume 133 – No.10, January 2016
- [16] Pooja Rohilla<sup>1</sup>, Ochin Sharma<sup>2</sup>, “Web Content Mining: An Implementation on Social Websites”, International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 7, July 2015
- [17] Ananthi.J, “A Survey Web Content Mining Methods and Applications for Information Extraction from Online Shopping Sites”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 4091-4094
- [18] Kaikala Anjani Sravanthi<sup>1</sup>, Yalamarthi Madhavi Lata, “Web Mining Using Cloud Computing”, International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 4, April 2013)
- [19] V. Shanmuga Priya<sup>1</sup>, S. Sakthivel, “An Implementation of Web Personalization Using Web Mining Techniques”, IJCSMC, Vol. 2, Issue. 6, June 2013, pg.145 – 150
- [20] M. Vengateshwaran and E. V. R. M Kalaimani, “Web Mining Research Direction and Open Source Tools”, Volume 4, Issue 7, July 2014 ISSN: 2277 128X, International Journal of Advanced Research in Computer Science and Software Engineering
- [21] Pranit B. Mohata and Prof. Sheetal Dhande, “Web Data Mining Techniques and Implementation for Handling Big Data”, IJCSMC, Vol. 4, Issue. 4, April 2015, pg.330 – 334