Journal of Advances and Scholarly Researches in Allied Education
Vol. XV, Issue No. 5, July-2018, ISSN 2230-7540

A Study on Hybrid Clustering Scheme

Poonam Rani[1]* Dr. M. K. Bisht[2]

[1] Research Scholar, Pacific University, Rajasthan, India

[2] Assistant Professor, Department of Computer Science, SSJ Campus, Almora, UK, India

Abstract – In this Paper model Will help to converts data coming from different social media sites , such as traditional blogging sites , e-mail sites , Social web pages of Face book and intelligent tutoring systems into information that may be useful for researchers, professors, institutions and students on understanding and evaluating educational systems, aiming for improving the quality of the educational process. For that matter, the educational data mining process aims for developing and applying computational approaches such as data mining, web mining, psychometrics and statistical techniques, for allowing the automatic information extraction from huge amounts of data.

Keywords: Clustering, Scheme, Hybrid, Social Sites

----------X----------

1. INTRODUCTION

The framework is designed with different learning algorithms for artificial neural network which also allows control over different respective learning parameters. The main interface will provide three options to choose development in one of these three categories. The interface of framework is displayed in next Figure. . There are three options respectively for different approaches used for this study. Choosing any one of these will open an interface of the approach. The developed framework allows generation of Neuro, Fuzzy and Neuro-Fuzzy systems as follows:

Artificial Neural Network Layer

When the required system needs to be trained artificial neural networks are to be used. Artificial neural networks have self learning mechanism based on their learning strategy. The artificial neural networks are broadly divided into two types of learning supervised and unsupervised.

In supervised learning artificial neural network is trained offline under supervision, in such kind of system users have to specify input and output broad categories along with training sets. In un-supervised learning user needs to specify only input and output broad categories, here training sets are not required as neural network is trained at the time of execution. The generic framework provides facilities through its user friendly interface to collect data in textual form, which can be saved for later use. The artificial neural network part of the framework allows development of system that requires learning. The developed library contains built in algorithm and learning paradigms. The developers of such system need to specify the input and output broad categories and training data sets required to train the artificial neural network .The user can choose from available different learning paradigms of the library to decide learning strategy .Once the parameters of the chosen learning strategy are decided the system will train the artificial neural network which can be saved for future use of the same artificial neural network in MATLAB.

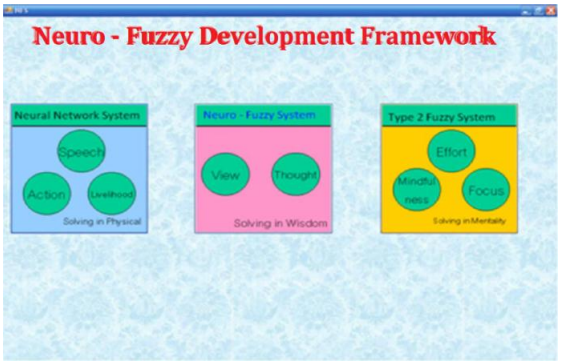

Fig 1 Neuro Fuzzy Development Framework

Fuzzy Logic System Layer

Those systems with imprecision and vagueness require extensive usage of logical conditions are developed under fuzzy logic systems. Fuzzy system compromise of type 1 fuzzy sets which are generalizations of crisp sets whose *membership grades* can only be 0 or 1. Hence type 2 fuzzy sets are used in the library to model the uncertainty. Type 2 fuzzy sets have advantage over conventional fuzzy sets as the membership grades of type 2 fuzzy sets are able to cover the second order derivatives which were not possible of type 1

fuzzy set. The framework developed using this library allows user to specify the rules in a interactive fashion. The framework allows its users to dynamically select the appropriate membership functions like linear, trapezoidal, gaussian and triangular. Based on the membership function the degree of uncertainty is handled by the fuzzy rule base through inference engine and appropriate output is given to user in human (English) language.

These rules are further processed by the inference engine which decides rule applicability in given situation. The rules are saved in database to document knowledge processing for future requirements. Hence, fuzzy logic systems are useful to support critical decision making in uncertain conditions which aids in development of expert systems.

Neuro-fuzzy System

The hybridization of artificial neural network and fuzzy logic system falls into these types of system. Fuzzy logic and artificial neural network systems are capable to generate experts systems in respective domain areas, however to extract advantages of both these systems, hybridization was done. Hence a neuro-fuzzy system has capability of logical conditioning to provide decision support and learning mechanism of artificial neural network. The generic framework facilitates automatic hybridization features for non-computer professionals and advance user can select the custom hybridization approach.

The neuro-fuzzy system uses libraries of fuzzy logic and artificial neural network. The framework provides different hybridization techniques to fuse artificial neural network with fuzzy logic to achieve the goal. The hybridization of artificial neural network and fuzzy logic generates expert advisory system that helps in decision support and .After collection of data it is hybridized in such a way that input and output broad categories of artificial neural network are merged with fuzzy rules to make interactive input and output for the generated system in a given domain area.

2. LOGICAL SYSTEM ARCHITECTURE

The architecture of the proposed temporal rule based classification model. It consists of six major components namely, social network Data, Data collection agent, Classification module, Rule manager, Rule base, Temporal Data Manager and Administrator. The input to the Hybrid Neuro Fuzzy system is referred from the social network trace data which is used in this work for carrying out the experiments.

The data collection agent collects the necessary data from the social networks data set. These data are sent to the classification module for classification of the data where it classifies the data by the help of training and pruning agents. The training agent trains the data which are received from the data collection agent. The pruning agent prunes the data which are received from the training agent. The rule manager is responsible for rule extraction, rule selection and decision making.

First, rules are extracted based on the information provided by the classification module. They are stored in the rule base. During testing, the rule selection subsystem selects the suitable rules and executes to get the results which are sent to the decision making agent.

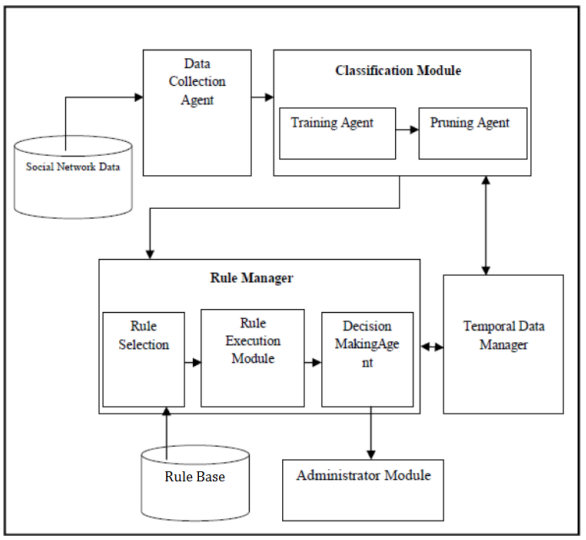

Fig 2 Logical System Architecture

Experiments

Fast and Appropriate Social Network Analysis (SNA) tools, techniques, are required to collect and classify opinion scores on social network sites, as a grouping on wrong opinion may create problems for a society or country. Social Network Analysis (SNA) is becoming an important tool for researcher as the number of users and groups increasing day by day on those social sites, and a large group may influence other. In this paper, we propose hybrid model of opinion recommendation systems, for individuals and for groups respectively, based on social liking and social influence network theory. By collecting the information of users' social networks and preference like, we designed a modified hybrid model to simulate the social influence by like and sharing the information among groups.

The purpose of this experiment and test results to analyze the suitability of ANN and Fuzzy sets method in a hybrid manner for social web sites classifications, First, we intend to use ANN methods in social media data classification by using some contemporary methods different than the traditional methods of statistics and data analysis, in next we want to propagate the fuzzy approach as a way to overcome the vagueness that is always present in social media analysis. We give a brief overview of the main ideas and recent results of social networks analysis , and we point to relationships between the two social network analysis and classification approaches .This experiment and test results

proposes a hybrid classification model using fuzzy and artificial neural network (HFANN). Information Gain and three popular social sites are used to extract information representing features that are then used to train and test the proposed methods. This novel approach combines the strength of ANN and Fuzzy sets in classification accuracy with utilizing social data and knowledge base available in the hate lexicons.

Use of Social network sites has already been popular even before the availability of websites like Facebook, Twitter or Whatsapp. In a comprehensive approach of modeling social network data as (UN) directed graphs has been proposed and has been widely accepted. Over the years a lot of research has been performed on e.g. cohesiveness of groups of members in social graphs or segmentation of social networks.

Although, the social weblog data analysis of the social media sites is an emerging field now , there exists various interesting WWW's phenomenon, several of them are listed by Barabási in his book "Linked: The New Science of Networks". According to him, "Networks are present everywhere. All we need is an eye for them".

Mostly social network analysis, similar to the ones discussed above, considers the links between its actors as binary (1 if present, or 0 if not). In reality, not all the actors are related with same degree. For example, hyperlinks between the two websites belonging to the same college will exhibit strong ties while these same websites will form weak ties with the third website belonging to some other college. But, based on the current methods these social links between all the three websites will be considered with equal importance.

MATLAB is a good tool to test non-linear, multi-dimensional, correlated, widely classified based data mining algorithm based on ANN Principles, for the testing of weblog based mining problems related to social media.

3. COMPARATIVE STUDY WITH OTHER METHODS

Social network sites are web sites that allow users to register, create their own profile page containing information about themselves (real or virtual), to establish public 'Friend' connections with other members and to communicate with other members. Communication typically takes the form of private emails, public comments written on each other's' profile pages, blog or pictures, or instant messaging. SNSs like Facebook and MySpace are amongst the ten most popular web sites in the world. SNSs are very popular in many countries and include Orkut (Brazil), Cyworld (Korea), and Mixi (Japan). The growth of SNSs seems to have been driven by the youth, with Facebook originating as a college site and MySpace having an average age of 21 for members in early 2008 .However, an increasing proportion of older members are also using these sites. The key motivating factor for using SNSs is sociability; however, this suggests that some types of people may never use social network sites extensively. Moreover, it seems that extraversion is beneficial in SNSs .and that female MySpace users seem to be more extraverted and more willing to self-disclose than male users .which suggests they may be more effective communicators in this environment.

Fuzzy theory is not only applicable to the real life problems solutions but can be applied to all the social networks analysis in general. As argued by Brunelli and Fedrizzi that most social analyses tools represent adjacency relations in bidirectional (binary) form, and presented "A fuzzy approach to social network analysis". They mathematically formalized the fuzzy theory dimension to represent the relations in the social networks. They extended their fuzzy model to represent the ordered weighted averaging (OWA) operators such as mean on m-ary fuzzy adjacency relations.

SNA combines the concept of the sociogram (a visual representation of relationships in a social group) with elements of graph theory to analyze patterns of interaction among people in various kinds of networks, allowing quantitative comparisons between different network structures. There is a large body of scholarly literature describing the use of SNA. Much of this work addresses the basic science of SNA, that is the development of theoretical models of network organization and the mathematical derivation of quantitative measures of network characteristics. More recent work examines the association of these quantitative measures with organizational performance outcomes.

Cummings and Cross, for example, found that degree of hierarchy, core-periphery structure, and structural holes of leaders correlated negatively with performance in 182 work groups in a large telecommunications company and Aydin et al found that increased network communication density was associated with higher use of an electronic medical record system by nurse practitioners and physician's assistants. There have also been studies showing how network parameters change with time. Shah, for example showed that network centrality decreased after downsizing in a consumer electronics firm, whereas Burkhardt and Brass documented increased network centrality after introduction of a new computer system in a federal agency.

4. MATHEMATICAL MODELING FOR NEURAL NETWORKS BASED TEST MATLAB

"We used Neural Networks as analytic techniques after the hypothesis formation processes of opinion mining in a social media system.

The first step is to design a specific network architecture (that includes a specific number of "layers" each consisting of a certain number of "neurons"). The size and structure of the network needs to match the nature (the number of nodes and graphs) of the investigated area. Because the latter is obviously not known very well at this early stage, this task is not easy and often involves multiple "trials and errors."

The designed network is then used to the process of "training." In that phase, neurons apply an iterative process to the number of inputs (variables) to adjust the weights of the network in order to optimally predict (in traditional terms, we could say find a "fit" to) the sample data on which the "training" is performed. After the phase of learning from an existing data set, the new network is ready and it can then be used to generate predictions based on our algorithms.

5. CONCLUSION

Fast and hybrid social network analysis techniques are needed to mine opinion scores on social networks as a grouping on wrong opinion may create problems for a society or country . Social Network Analysis (SNA) can be used as an important tool for researchers, as the number of users and groups increasing day by day on that social sites, and a large group may influence others, but the necessary information is often distributed and hidden on social site servers, so there is a need to design some new approaches for collection and analysis the social web data. The test results are tested on Fuzzy set and ANN based models to improve finding relation and to prove that there is meaningful and good correlation between the like and share of social media users on a common issue while chatting or commenting, even like and dislike ratio is very much same and mean of net inputs we used for output layer. Each of test case results is meaningful and there are higher correlations.

6. REFERENCES

1. Ability Net (2008). State of the e-Nation. "Social Networking Websites. Available online at: http://www.abilitynet.org.uk/docs/enation/2008 SocialNetworkingSites.pdf [accessed 31st August 2011].

2. Discapnet (2010). Accessibility of Social Networking Services. Observatory of ICT Accessibility.

3. Hewett, R. and Douglas, G. (2010) Post-14 transition support – a survey of visiting teacher services for blind and partially sighted students. Birmingham. Visual Impairment Centre for Teaching and Research, School of Education, University of Birmingham.

4. Hewett, R., Douglas, G. and Williams, H. (2011). Post-14 transitions support – a survey of the transition experience of visually impaired young people. Technical report of findings to summer 2011. Birmingham. Visual Impairment Centre for Teaching and Research, School of Education, University of Birmingham.

5. Kasiviswanathan S.P., Melville P., Banerjee A., Sindhwani V. (2011). Emerging topic detection using dictionary learning. In: Proc of CIKM'11

6. Kavanaugh A., Fox E.A., Sheetz S., Yang S., Li L.T., Whalen T., Shoemaker D., Natsev P., Xie L. (2011). Social media use by government: from the routine to the critical. In: Proc of 12th annual international digital government research conference, College Park, Maryland, pp. 121–130

7. Krippendorff K. (2004). Content analysis: an introduction to its methodology. Sage, London.

Corresponding Author

Poonam Rani*

Research Scholar, Pacific University, Rajasthan, India