Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:10(3), October, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed : Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available : www.ijmer.in

MINING AND ANALYSIS OF SOCIAL NETWORK SITES

K. Divya
Lecturer in Computer Science, GRP. Government Degree College, Bhainsa Dist. Nirmal

Abstract

With the availability of all types of huge data and a growing number of users for social networking sites, it would seem this area is the perfect environment for extensive data mining, research, and designing new algorithms to improve data mining on various aspects. While the usage rates, public availability and uploading of huge amounts of data on social media provide us a new area for research, there are several impediments to capitalising on data mining algorithms with best fit strategies for this area. From an academic researcher perspective, studying social area networks, growth rates, and social implications of social networking sites is likely to draw strong willingness for designing and testing new approaches of data mining from social area network sites, either by owners or by users. Social media is one of the forms of content on the Web. The important features of social media are the dynamism of new content and instant interaction among users. Social media is revolutionising the way information is provided and communicated over conventional media. This paper will demonstrate a hybrid recommendation system for matching similarity between people in social networks with the use of clustering. The proposed clustering algorithms amalgamate the similarity measurement and constraints validation in one step and execute it at the cluster level, which makes these algorithms suitable for large datasets, such as social networks. The proposed system and methods are evaluated using the data collected from a real-life social network.

Kew Words: Social Networking Sites, Algorithm, Social media, Revolutionizing, Conventional media

Introduction:

The social network data preparation requires huge attention because it processes raw data and convert them into usable forms for data mining and analysis tasks, in our work we stress upon the relevance of data preparation for social network analysis and mining tasks and further designing some new approaches with the help of soft computing methods. The main research problem covered in this study is how to prepare good-quality social network data for data analysis and mining by collecting web data from online social networks such as Twitter, Face book, Micro blogging sites, and Mobile social media sites like WhatsApp etc. that are rapidly grown in popularity. As we have studied that old conventional mining algorithms all alone cannot operate or generate accurate results on the giant and complex data of social networks. With the focus being profile extraction and designing algorithms for better mining on social media sites to frame relationships among users and groups on various social media sites. Since the data for the proposed work has to be obtained from social media Web only, Web mining techniques for Social Network extraction will be studied as well. In this proposed work, relationship information between various groups, social media communities and users entities obtained from multiple user profiles of social websites data sources have to be extracted which may lead to various ambiguities. As we studied that Soft Computing techniques can be used to solve ambiguity problems. We intend to develop some Soft Computing techniques based algorithms using ANN, Fuzzy sets and we will use ACO tools also. We also intend to perform a comparative study to evaluate the effectiveness of these algorithms in social network extraction, particularly in an academic social network extraction, where user profile data of teachers, students or other workers of a University or college will be used for testing our results.

Review of Literature:

During the past years, social media has seen an exponential growth in the number of users. For example, there are more than 1.44 billion members belonging to the Facebook network (Facebook2015), while Twitter now has more than 320 million monthly active users. Social media applications or platforms which in general can be classified into “weblogs, micro blogs, social network sites, social networks, discussion forums, wikis, podcast networks, picture and video sharing

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:10(3), October, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed : Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available : www.ijmer.in

platforms, ratings and reviews communities, social bookmarking sites, and avatar based virtual reality spaces". In a wider context, social media refers to "a conversational, distributed mode of content generation, dissemination, and communication among communities" (Zeng et al. 2010).

Social media applications have transformed the way in how people share, develop, and disseminate information. In specific, the methodology of information absorption and transformation is closely interrelated and interdependent with the process of creating and sharing information (Zeng et al. 2010).

Corporate have started using social media applications for communication internally as well as externally for a various reasons. Internal social media applications enhance communication and connectivity between coworkers. External communication enhances knowledge management, and product/service innovation development. Numerous corporate have started to establish social media-based communication with business partners and have also begun to involve themselves in public social media activities for the purposes of brand marketing, public relations (PR), customer relationships management, reputation management, and consultancy.

In the political field, social media is believed to have the potential to increase political participation and awareness by citizens and voters which can aid in decision and policy making (Wattal et al. 2010).

While Twitter can be taken as an ideal public application for communication political information and opinions in real time and to larger audience (Stieglitz and Dang-Xuan 2013a, 2013b; Bruns and Highfield 2013), politicians have also begun to use various social media applications for entering into active participation and discussions with citizens of the country.

Based on literature review we have concluded that a little work has been done in the area of Soft Computing techniques based algorithms and tools for social media mining and data extraction. In this work, we intend to study the existing social media mining approaches, analyze them, and propose a new model based on Social Web mining techniques with a mix of soft computing techniques that shall be an improvement over the existing ones. Thus, we hope to improve the efficiency and usefulness of the extracted social network

Research Objectives:

The following are the basic aims and objectives of the study:-

1. To study of various earlier social web mining and profile extraction algorithm.
2. To study of various earlier Soft Computing technique used for data mining and web mining.
3. To developing Soft Computing techniques based social web mining and profile extraction algorithm
4. To developing Soft computing techniques based on ANN, Fuzzy sets and ACO for disambiguation algorithm(s).
5. To developing a Social web sites Network Extraction system that will be using multiple Groups communities, user profiles that may be multi-relational in nature, and to test them on real data logs based on spectral and temporal considerations where possible.

Research Methodology:

The Secondary and Primary data analysis research methods are followed to achieve the goals and objectives defined earlier for the proposed study.

As we collected social media datasets and studied some simulation packages like MATLAB, the web log datasets for the proposed work are mostly available online and are free for research purposes. The following sources are used to.

1. University / College Institution Websites and Users of Social media sites of that
2. Digital Social Web Log Libraries.
3. Digital Theses Libraries for Literature Review and Early work Study

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:10(3), October, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed : Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available : www.ijmer.in

4. Web Links of Referred papers
5. Blog Post and Tweets
6. E-mail Communication Logs.
7. Online and offline survey conducted for this study

Web Log data for online social media communities are based on materials that are intentionally public and therefore considered a form of publication or mass communication. For simulation purposes packages like MATLAB will be used and, where necessary required software/simulation package may be downloaded from freely available sources from the Internet for the study at hand.

Research Outcomes:

In this study, some new soft computing based algorithms namely neural genetic clustering algorithm, hybrid fuzzy ant clustering based classification algorithm, neuro fuzzy link based classification algorithm, and hybrid intelligent multiuser fuzzy association with classification algorithm, have been proposed. All these algorithms have been implemented and now tested by the researcher using MATLAB and other tools as discussed earlier. From the experimental results, we will prove that the performance of this proposed algorithm is better than the other related algorithms. Hence, they are useful for web mining and to trace good relations, associations among social media users while providing their opinion and like, dislike formation in social networks.

Followings are the outcomes of this study:

a. Proposal of soft computing based new approached for social network analysis model.
b. Proposal of a multi factor analysis model for effective social media users clustering.
c. Design and implementation of the new social media mining algorithms named by us as Neural genetic clustering algorithm, Hybrid Ant Neuro Fuzzy Clustering based algorithm, Fusion of Link Based Neuro Classification Fuzzy Paradigm, and soft computing based ensemble techniques .
d. Provision of a new model for association and clustering the social media groups using the proposed hybrid clustering based classification algorithm for effective and fast results.

Conclusion:

The majority of academic researchers present their research work on the Web. This trace can be used to derive useful information about their past & present activities and forecast their future intentions. It also provides an insight into what is going on presently in a particular research area and research objectives. Thus, extraction/construction of social network of academic researchers can be of important value for scientific community. For an extracted network to be useful all the possible relations have to be analyzed. This relationship information can be obtained from various data sources available on the Web (social Web pages links, e-mail archives, Blog posts, tweets). From each of these data sources a single-relational network can be extracted and for a network to be multi-relational, relationship information from various data sources have to be considered. The techniques available so far use one or two of these data sources. Thus there is a need to develop a system that will be multi-relational with the relationship information obtained from various data sources.

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:10(3), October, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed : Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available : www.ijmer.in

References:

1) Ralph Gross, Alessandro AcquistiH. John Heinz III, "Information Revelation and Privacy in Online Social Networks (The Facebook case)", Pre-proceedings version. ACM Workshop on Privacy in the Electronic Society (WPES), 2005.
2) Huan Liu and Lei Yu, "Toward Integrating Feature Selection Algorithms for Classification and Clustering",IEEE Transactions on Knowledge and Data Engineering Volume 17 Issue 4,April 2005
3) Andrea Esuli_ and Fabrizio Sebastiani†, "SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining", Proceedings of the 5th Conference on Language Resources and Evaluation, 2006.
4) Marcelo Maia, Jussara Almeida, Virgílio Almeida, "Identifying User Behavior in Online Social Networks", SocialNets'08, April 1, 2008, Glasgow, Scotland, UK Copyright 2008 ACM ISBN 978-1-60558-124-8/08/04.
5) Meenu Sharma, "Clustering in Data Mining: A Brief Review", International Journal of Core Engineering & Management (IJCEM) Volume 1, Issue 5, August 2014.