Integration of Bayesian Theory and Association Rule Mining in Predicting User's Browsing Activities – Survey Paper

Geoffrey Gitonga
Department of Computing
School of Computing and
Information Technology
Jomo Kenyatta University of
Agriculture & Technology
Nairobi, Kenya

Wilson Cheruiyot
Department of Computing
School of Computing and
Information Technology
Jomo Kenyatta University of
Agriculture & Technology,
Nairobi, Kenya

Waweru Mwangi
Department of Computing
School of Computing and
Information Technology
Jomo Kenyatta University of
Agriculture & Technology,
Nairobi, Kenya

Abstract: Bayesian theory and association rule mining methods are artificial intelligence techniques that have been used in various computing fields, especially in machine learning. Internet has been considered as an easy ground for vices like radicalization because of its diverse nature and ease of information access. These vices could be managed using recommender systems methods which are used to deliver users' preference data based on their previous interests and in relation with the community around the user. The recommender systems are divided into two broad categories, i.e. collaborative systems which considers users which share the same preferences as the user in question and content-based recommender systems tends to recommend websites similar to those already liked by the user. Recent research and information from security organs indicate that, online radicalization has been growing at an alarming rate. The paper reviews in depth what has been carried out in recommender systems and looks at how these methods could be combined to from a strong system to monitor and manage online menace as a result of radicalization. The relationship between different websites and the trend from continuous access of these websites forms the basis for probabilistic reasoning in understanding the users' behavior. Association rule mining method has been widely used in recommender systems in profiling and generating users' preferences. To add probabilistic reasoning considering internet magnitude and more so in social media, Bayesian theory is incorporated. Combination of this two techniques provides better analysis of the results thereby adding reliability and knowledge to the results.

1. INTRODUCTION

The use of internet has been growing tremendously over the years accelerated by the ease in internet access. This has led to growing dependent on the internet for chores like business, work and other social responsibilities like sharing of information, files and opinions using applications like Facebook, Twitter and Friendster among others. As a result of this enormous development of the internet and web based applications, new vices, like radicalization, in relation to internet usage have also been introduced that were not considered in the internet usage and accessibility campaign. Online radicalization which is a process by which internet users are brainwashed to accept extreme religious, social or political ideologies. Radicalization therefore led to other internet vices like cybercrime, child trafficking and even hacking among others.

Due to internet development and the increase in the number of users, recommender systems for web applications were introduced to anticipate users' preferences in terms of content and information based on their personalized content. This is based on the users' interaction with the system such that the information is analyzed over time to determine the preferences. Preferences could also be determined based on similarities discovered between preferences from different user groups (community opinion) where the trend could also be used as the basis for recommendation. Recommender systems are mainly based on two categories; collaborative and content-based filtering methods.

Using this idea, internet menace could also be curbed by being able to understand users' online activities. Radicalization does not happen overnight, and therefore with accumulated information about user's browsing activities, the system would be able to identify and relate the trend that leads to radicalization. The relationships between websites and users behavior is analyzed in order to identify the trend in regard to internet usage. The importance of websites to users will be considered based on time spent, frequency of accessibility to determine the probability of radicalization using Bayesian theory and association rule mining methods. The following sections will review what has been done by other scholars in this field, concentrating on the recommender systems techniques and how they could be applied in curbing the menace associated with the use of web applications.

2. THEORETICAL REVIEW

Users' activities on the website are guided by the guiding principles of ethics. Users tend to misuse online sites knowingly thereby ending up radicalized as a result of a continuous use of some specific online sites. Ethics has been defined in various ways by different scholars. According to Mohd & Sembok [7], these are moral values that help guide human behavior, actions, and options. They can be defined in two ways normative and prescriptive. For the normative description, ethics are considered to be well-based standards of right and wrong that suggest what humans should to do,

usually in terms of rights, obligations, benefits to society, fairness, and specific virtues. Ethics also involves values that related to virtues of honesty, compassion, and loyalty. For the prescriptive description, ethics refers to the study and nurturing of personal ethical standards, as well as community ethics, in terms of behaviour, feelings, laws, and social habits and norms which can deviate from more universal ethical standards [7].

Hilty [5] also defines the term ethics in two ways; philosophical reflection and the other to more practical governing principles, which is the principle of moral conduct governing an individual or a group.

Conclusion from the above definitions can therefore be derived that, ethics are standards of right or wrong as perceived by human being in the use of technology. What is perceived as wrong in the human eye could remain wrong as long as the person in question perceive it so. It does not have to be controlled by government institutions laws or standards for them to be seen so.

In relation to technology, and more so internet, there are many issues that come as a result of development in technology. These could be good or bad depending on the situation and information use. Social sites like Myspace, Facebook, Friendster, Twitter always enable users to learn something new. In Friendster, for example, students can share comments and pictures as well as ideas which could be beneficial or negative depending on the perception of the user. Negative influence such as addiction to pornographic, radicalization and even exposure to fraud [12]. Ethical issues in Information Technology (IT) involves three "rights" which are;

- Right to know: This is the extent to which information can be accessed from IT infrastructure.
- Property right: This is the right to that IT facilities are protected. For example, users can protect their computers from viruses
- Right to privacy: This refers to right to user's privacy. Every technology user should be responsible to protect their information like passwords from being used by other people.

As much as these rights guarantee freedom of information access and protection, some people, using skills acquired in a bid to enforce these rights, violate them. They therefore ends up in fraud, radicalization or even hacking.

Other challenges that face the industry, especially in academic institutions is the homogeneous society that makes it impossible to come up with umbrella technology ethical standards that can be used across board. In higher learning institutions, there are many people with different age groups, class, gender and affiliations. People of different occupations, age groups and educational backgrounds who live in different countries representing a variety of cultures use resources available through the global computer network. This, however, complicates the problem of developing universal standards of behaviour and a system of ethical norms, which could be widely recognized in the World Wide Web. Meanwhile many believe that such a system is really needed today. This is as mentioned by Gordon & Johnson [4]. However, this heterogeneous society brings along vices that are shared among the users thereby ending up affecting them in the long run.

According to Hilty [5], much has been done in terms of enforcing IT ethics. The adoption of the "ACM Code of Ethics and Professional Conduct" by the ACM Council in 1992. Also, IFIP was introduced SIG 9.2.2, its Special Interest Group on a Framework for Ethics of Computing. He also mentioned that, there have been many challenges for IT ethics for over a decade. For example in privacy protection, data security, trust in complex systems, online communication, freedom of speech, intellectual property and sustainable development

As much as technology was introduced in learning institutions over a decade ago, concentration has been on the application of technology and less on ethics and influence effects. As mentioned by Mohd & Sembok [7], technology ethics have been negatively been affected mainly because of lack of awareness of technology information security issues, the rapidly evolving complexity of systems that are difficult to implement and operate, the ease at which one can access and reach information and communication technology, the anonymity provided by these technologies, and the transnational nature of communication networks. Cybercrime and online radicalization have been issues of the global ethical aspects that have not been resolved fully despite government regulations and policies. Even where awareness is growing and where legislation may be adequate, capacity to use information security technologies and related procedures, as well as to protect against, detect and respond effectively, to cybercrime and online radicalization, is low.

Globally, internet and communications cut across territorial boundaries, thereby creating a new realm of human activities, and undermining the feasibility and legitimacy of applying laws based on geographic boundaries. The new boundaries, which are manifested in the monitor screens, firewalls, passwords, intruder detection, and virus busters, have created new personalities, groups, organizations, and other new forms of social, economic, and political groupings in the cyber world of bits. Traditional border-based law making and law enforcing authorities find this new environment of cyber boundaries very challenging.

The following are some of the main malpractice issues that have not been adequately addressed by the current measures in place, as mentioned by Mohd & Sembok [7];

Cybercrime: Research conducted by Computer Security Institute (CSI, 2003) in Asia-Pacific region indicate that the threat from computer crime and information security breaches continues persistently thereby resulting to financial losses 50 – 60 % between 2002 to 2004.

Pornography: Internet brings great benefits, but also enormous risks to children. Among the worst forms of cybercrime is the global trade in child pornography. International criminal investigation have been trying to handle this situation but with a very small margin.

Piracy: In early 2002, International Planning and Research Corporation (IPR) completed an analysis of software piracy for the year 2001 as part of an ongoing study for the Business Software Alliance (BSA) and its member companies. The world piracy rate increased in two consecutive years, 2000 and 2001. The 2001 piracy rate of 40% is a marked increase from 37% in 2000. Both years were up from the low set in

1999 at 36%. Since then, the rate of piracy has been increasing tremendously over the years.

Neumann & Stevens [10] also mentioned online radicalization as the one of the biggest challenge that security institutions have not been able to handle. They define radicalization as "the process (or processes) whereby individuals or groups come to approve of and (ultimately) participate in the use of violence for political aims". Most policymakers and scholars have only a general idea that internet is the main breeding ground for radicalization but only the most cursory idea of how it works.

Therefore, online platform forms an ideal cheap ground for extremist organizations to sell their ideals. According to Neumann & Stevens [10], internet can be used in three main ways by the extremists;

- To illustrate, communicate their ideological messages and/or narratives
- Internet provides a risk-free platform for potential recruits to find like-minded people. It is also used to for a network of these like-minded
- It creates a new social platform in which unacceptable views and behavior are normalized which would have otherwise been rejected. "Surrounded by other radicals, the internet becomes a virtual 'echo chamber' in which the most extreme ideas and suggestions receive the most encouragement and support."

According to Olumoye & Y [11], he suggested the following measures that could curb negative effects of technology:

- The society must ensure each person is accountable for everything he or she does, no matter how inexplicable his or her action may appear.
- Since there are growing complexities of ethical and social issues that revolve around multiple breaches, it becomes imperative for the educators and computer
  professional bodies to develop curriculum on ethical and professionalcodes of conduct in the information society.
- There is need to lay emphases on information systems security controls
- The government should develop a comprehensive laws and legislations to create a sense or awareness of compliance requirements that affects information systems professionals.
- Our law enforcement agents should be more sophisticated in their computer crime investigation. This can been enhanced with the use of computer forensics, which is a formal investigative technique used in evaluating digital information for judicial review

The use of intelligence techniques have of late taken a centre stage especially in the use of web based applications in identifying and categorizing ideas based on users preferences. According to Mukhopadhyay, Vij, & Wanaskar [8] recommender systems are tools used to filter, sort and order items and data based on opinions of users or community to assist users to determine content of interest from a doubtless overwhelming set of decisions. Two algorithms became very popular: collaborative filtering and content-based filtering

Content-based recommender systems work with individual profiling of users from the very beginning. A profile has biographic data about a user and preferences based on items rated by the user. In the entire process of recommendation, the system makes a comparison of items that were already positively rated by user with items that were not rated and looks for similarities. This is carried out mainly using tags or keywords. In this case the profiles of other users are not essential.

Collaborative filtering method involves searching and locating users in a community that share item preferences based on their common online browsing habits. If two or more users have similar or closely similar rated items, they are considered to have similar preferences. A user gets recommendations to choose items that he/she has not rated before, but that were already positively rated by users in his/her neighbourhood [8].

Mukhopadhyay, Vij, & Wanaskar [8] identified the following challenges affecting recommender systems;

1. Cold-start: This occurs when providing recommendations to new users. The system is not able to recommend preferences to new users as their profiles are almost empty. Cold-start also affects new items that have not been captured by the system and haven't been rated before.
2. Trust: Consideration of people with short history cannot be compared with those with rich history in their profiles. The issue of trust arises towards evaluations of a certain customer.
3. Scalability: The number of users continues to grow tremendously and therefore, the system needs more resources for processing information and forming recommendations. Majority of resources is utilized during determination of users' preferences.
4. Sparsity: This refers to the problem of lack of information. In online shops that have a huge amount of users and items there are some users who rate just a few items.
5. Privacy: In order to receive the most accurate and correct recommendation, the system must acquire the most amount of information possible about the user, including demographic data, and data about the location of a particular user. Naturally, the question of reliability, security and confidentiality of the given information arises. Many online shops offer effective protection of privacy of the users by utilizing specialized algorithms and programs.

This idea could be extended in profiling users interest and thereby be in apposition to determine and predict the kind of a person the user is based on his/her online activities. These systems are developed using intelligence techniques such as association rule mining, Bayesian theory, cluster analysis, reinforcement learning among other techniques.

In association rule for example as stated by Mukhopadhyay, Vij, & Wanaskar [8] an upgraded association rule mining method cab be used in recommender systems because of its scalability and gives high precision in results determination. Using this method, the weight of the web page is given in binary form to pages that are visited to find whether the page is present or not. This method assumes that if the web page is visited by the user, that page is considered important specifically to that user. However, not all the pages visited by the user are of interest. Some users may visit a page looking for useful information but it may not have what the user is looking for. Therefore, factors like time spent by the user and visiting frequency of the page are considered during web page

calculation. This idea could also be used in identifying web pages of interest to the user in order to understand users browsing behaviour thereby predicting their personalities. This therefore could be used to curb vices like cybercrime, radicalization among others.

Another technique highlighted by different scholars in recommender systems and data mining, especially in probability theory and statistics is Bayes' theorem also referred to as Bayes' rule. It is a result that is of importance in the mathematical manipulation of conditional probabilities. Bayes rule can be derived from more basic axioms of probability, specifically conditional probability. This has also been extended to machine learning.

According to Tipping [13] when applied, the probabilities involved in Bayes' theorem may have any of a number of probability interpretations. In one of these interpretations, the theorem is used directly as part of a particular approach to statistical inference. ln particular, with the Bayesian interpretation of probability, the theorem expresses how a subjective degree of belief should rationally change to account for evidence: this is Bayesian inference, which is fundamental to Bayesian statistics. However, Bayes' theorem has applications in a wide range of calculations involving probabilities, not just in Bayesian inference.

Mathematically, Bayes' theorem gives the relationship between the probabilities of A and B,P(A) and P(B), and the conditional
probabilities of A given B and B given A, P(A|B) and P(B|A). In its most common form, it is:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

The meaning of this statement depends on the interpretation of probability ascribed to the Bayesian interpretation [13].

Neal [9] provides the following steps in implementation of Bayesian method;

1. Formulation of knowledge about the situation probabilistically
2. Gathering of data
3. Computing the posterior probability distribution for the parameters
4. Given the observed data
5. Use of the identified posterior distribution to:
   i. Reach scientific conclusions, properly accounting for uncertainty.
   ii. Make predictions by averaging over the posterior distribution.
   iii. Make decisions so as to minimize posterior expected loss.

The posterior distribution for the model parameters given the observed data is found by combining the prior distribution with the likelihood for the parameters given the data. This is done using Bayes' Rule:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

Another form of Bayes's Theorem that is generally encountered when looking at two competing statements or hypotheses is:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}.$$

For proposition A and evidence or background B,

- P(A),the prior probability, is the initial degree of belief in A.
- P(-A), is the corresponding probability of the initial degree of belief against A: 1-P(A)=P(-A)
  - P(B|A), the conditional probability or likelihood, is the degree of belief in B, given that the proposition A is true.
  - P(B|-A), the conditional probability or likelihood, is the degree of belief in B, given that the proposition A is false.
  - P(A|B), the posterior probability, is the probability for A after taking into account B for and against A.

Tipping [13], has pointed out the main advantage and disadvantage of the Bayesian method. The greatest advantage of a Bayesian approach is that, there is an automatic preference for simple models that sufficiently explain the data without unnecessary complexity. This property holds even if the prior P(W) is completely uninformative. The practical disadvantage of the Bayesian approach is that it requires the modeller to perform integrations over variables, and many of these computations are analytically intractable. As a result, much contemporary research in Bayesian approaches to machine learning relies on, or is directly concerned with, approximation techniques.

The following are other artificial intelligence techniques as identified by Wu et al. [14]

a) Decision tree learning

This method uses a decision tree as a predictive technique mapping observations about an item to conclusions about the item's target value. This techniques is among the most predictive modelling technique used in statistics, data mining and machine learning. This approach depicts a tree like structure with leaves representing class labels and branches conjunctions of features that lead to class labels.

In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. Decision tree learning is a method commonly used in data mining. The goal is to create a model that predicts the value of a target variable based on several input variables.

Data comes in records of the form:

$$(\mathbf{x}, Y) = (x_1, x_2, x_3, ..., x_k, Y)$$

The dependent variable, Y, is the target variable that we are trying to understand, classify or generalize. The vector x is composed of the input variables, x1, x2, x3 etc., that are used for that task.

This method however has got the following limitations;

- Practical decision-tree learning algorithms are based on heuristics such as the greedy algorithm where locally-optimal decisions are made at each node. Such

algorithms cannot guarantee to return the globally-optimal decision tree.

- Decision-tree learners can create over-complex trees that do not generalize well from the training data (over fitting).
- There are concepts that are hard to learn because decision trees do not express them easily, such as parity (the evenness or oddness of the number of bits with value one within a given set of bits, and is thus determined by the value of all the bits. It can be calculated via an XOR sum of the bits, yielding 0 for even parity and 1 for odd parity. This property of being dependent upon all the bits and changing value if any one bit changes allow for its use in error detection schemes.) or multiplexer (a device that selects one of several analog or digital input signals and forwards the selected input into a single line) problems. In such cases, the decision tree becomes prohibitively large.
- For data including categorical variables with different numbers of levels, information gain in decision trees is biased in favour of those attributes with more levels.

b) Association rule mining

This is a popular and well researched method for discovering interesting relations between variables in large databases. It is intended to identify strong rules discovered in databases using different measures of interestingness. Based on the concept of strong rules, Agrawal & Srikant [2], introduced association rules for discovering regularities between products in large-scale transaction data recorded by point-of-sale (POS) systems in supermarkets. For example, the rule $\{onions, potatoes\} \Rightarrow \{burger\}$ found in the sales data of a supermarket would indicate that if a customer buys onions and potatoes together, he or she is likely to also buy hamburger meat. Such information can be used as the basis for decisions about marketing activities such as, e.g., promotional pricing or product placements.

As stated by Mukhopadhyay, Vij, & Wanaskar [8], association rule mining technique can be easily used in recommendation systems and it is scalable, gives high precision , and only gives binary weight to the pages that are visited i.e. to find whether the page is present or not. User may visit a page but it may not have useful information for him. So factors like time spent by the user and visiting frequency of the page should be considered for the page consideration. So in association rule mining method the weight of the page is also included.

One limitation of the standard approach to discovering associations is that by searching massive numbers of possible associations to look for collections of items that appear to be associated, there is a large risk of finding many false associations. These are collections of items that co-occur with unexpected frequency in the data, but only do so by chance.

c) Cluster analysis

Also known as clustering is the task of grouping a set of objects so that the objects in the same cluster are more similar to each other than to those in other clusters. It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning.

Popular notions of clusters include groups with small distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. It will often be necessary to modify data pre-processing and model parameters until the result achieves the desired properties.

d) Reinforcement learning

As emphasized by Peters et al. [6] Reinforcement learning is an area of machine learning inspired by behaviorist, concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. Reinforcement learning is mostly applied in dynamic programming and machine learning.

Reinforcement learning is suitable to problems which include a long-term versus short-term reward trade-off. Reinforcement learning has been applied successfully to problems like, robot control, elevator, scheduling, telecommunications, backgammon and checkers.

Reinforcement learning is considered powerful because of two components: The use of samples to optimize performance and the use of function approximation to deal with large environments. Peters et al. [6] also identify situations where reinforcement learning can also be used successfully. These situations are:

- A model of the environment is known, but an analytic solution is not available;
- Only a simulation model of the environment is given (the subject of simulation-based optimization);
- The only way to collect information about the environment is by interacting with it.

The main limitation in this technique is that, although finite-time performance bounds appeared for many algorithms in the recent years, these bounds are expected to be rather loose and thus more work is needed to better understand the relative advantages, as well as the limitations of these algorithms.

The researcher intends to use the Bayesian theory combined with association rule learning to generate the proposed algorithm. The aspects of recommender systems, using content-based approach will also be considered in predicting what sites are most likely to be visited by the user. This will ensure that the limitations experienced from both methods are covered thereby increasing dependability percentage of the algorithm. In Bayesian technique, much contemporary research to machine learning relies on, or is directly concerned with approximation techniques and is based on

approximation. Since both methods are based on probabilities, combining both increases the chances of acquiring the correct results.

3. DISCUSSION

Machine learning/artificial intelligence is a relatively new area in information technology. A variety of techniques have been introduced each with its benefits and shortcomings based on different applications. More research has been carried out on integration and enhancement of these techniques to form hybrids in a bid to make them much better and effective.

Inspired by Forsati & Rahbar [3] using weighted association mining method; the following criteria are considered;

1. Page duration
2. Frequency of access
3. Indegree – in this case it refers to the average number of objects in a page

These three criteria are used to generate the average weight of the web page. The more the weight, the more important the page is considered by the user.

The associated method is used in capturing the time and webpage access frequency which forms the basis of importance estimation. Generally, a user is assumed to spend more time on more useful web pages than on less useful ones. However, this alone cannot be sufficient as other factors like size of the page and number of objects in a page might equally be important. The formula of duration is given in *Equation (1)* below. Frequency is the number of times that a web page is accessed by a user. The general assumption is that web pages with a higher frequency are of stronger interest to users. However, it is important to consider in the calculating the frequency of a web page the in-degree of that page, in this research the number of objects in a page e.g. images, videos, flashy objects etc. The frequency is calculated using the formula in *Equation (2)* below [8].

Bayesian aspect can then be used to capture and predict user's behaviour based on the usage of the social media like Twitter and Facebook. This therefore concentrates on the probabilistic assumptions specifically on social media. These could be measured through comments posted, likes or tags identifying the trends and already known facts in regard to social media usage and radicalization statistics. In particular, this is used to generate the probability Maximum a Posteriori (MAP) as to whether the user is radicalized or not based on his usage statistics of the social media sites. This is also forms part of the equation generated for association rule mining as *Equation (3)*. Therefore, time spent by a user on a web page, frequency of visiting and probability estimation based on social media usage and information, are then used as three crucial aspects in measuring the user's interest on the web page and relationships among different web pages using *Equation (3)* as indicated below.

$$\text{Duration} = \frac{\text{Time spent}}{\text{Size (bits)}} \qquad Equation~(1)$$

$$\text{Frequency} = \text{Number of visits} * \text{Indegree} \qquad Equation~(2)$$

$$\text{Probability} = \text{MAP} \qquad Equation~(3)$$

$$\text{Weight} = \text{Duration} * \text{Frequency} * \text{Probability} \qquad Equation~(4)$$

4. CONCLUSION

Internet can no longer be ignored and therefore, monitoring measures efficient in controlling users' activities online need to be put in place that are. As much as more research has been carried out in enhancing artificial intelligence techniques that are used in profiling and recommender systems, more is still required. In this survey, association rule mining has been enhanced to weighted association rule mining considering the average weight of the web page. However, not all web pages should be treated the same way. Social media web pages have different kind of information and should be treated as such, differently from other web pages. Therefore, as much as weighted association rule mining method would be appropriate for other informational web pages, Bayesian theory would be appropriate for social media sites. The combination of these two algorithms would ultimately form an enhanced algorithm with accurate and reliable results.

5. REFERENCES

[1]. Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. San Jose, CA 95120: IBM Almaden Research Center 650 Harry Road.

[2]. Forsati, R. M., & Rahbar, A. (2009). An efficient algorithm for web recommendation systems. AICCSA 2009. IEEE/ACS International Conference on Computer Systems and Applications, 579-586.

[3]. Gordon, L., & Johnson, D. G. (2004). Ethical psychological and societal problems of the application of technologys in education. Unesco Institute For Information Technologies in Education.

[4]. Hilty, L. M. (2002, December). The role of ethics in the field of information and communication technologies. Discussion paper prepared for the technology working group of the Swiss Academy of technical (SATW).

[5]. Peters, J., S, V., & S, S. (2003). Reinforcement learning for humanoid robotics. Third IEEE-RAS interantional conference on Humanoid Robots, 29-30.

[6]. Mohd, T., & Sembok, T. (2003). Ethics of information communication technology (technology). University Kebangsaan Malasia for the regional meeting on ethics of science and technology.

[7]. Mukhopadhyay, D., Vij, S. R., & Wanaskar, U. H. (2013). A hybrid web recommendation system

based on the improved association rule mining algorithm.

[8]. Neal, R. M. (2004). Bayesian methods for machine learning. University of Toronto.

[9]. Neumann, R. P., & Stevens, T. (2009). Countering online radicalization. A strategy for action; the international centre for the study of radicalization and political violence (ICSR).

[10]. Olumoye, & Y, M. (2013, November 11). Ethics and social impact of information systems in our society: analysis and recommendations. International journal of science and research (IJSR), 2, 2319-7064.

[11]. Rikowski, R. (2009). Using information technology, teaching ethical issues in IT (7th ed.). London south Bank University: Mc. Graw Hill.

[12]. Tipping, M. E. (2006). Bayesian Inference: an introduction to principles and practice in machine learning. Cambridge, UK: Microsoft Research.

[13]. Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., . . . Zhou, Z.-H. (2007). Top 10 algorithms in data mining. London: Springer-Verlag Limited.