International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

Detection of SNMD for Multiple Social Media Via Machine Learning

ShrutiKale[1], SimranPhade[2], ShivamYadav[3], SnehalMedhi[4], Prof. S.D.Deo[5]

[1,2,3,4]Student, Dept. of Information Technology, NBN SSOE, Maharashtra, India
[5] Guide-Professor, Dept. of Information Technology, NBN SSOE, Maharashtra, India

---

Abstract –The explosive growth in quality of social networking ends up in the problematic usage .An increasing vary of social network mental disorders (SNMDs), like Cyber-Relationship Addiction, info Overload, and internet Compulsion, are recently noted. Symptoms of the segmental disorders are usually observed passively today, resulting in delayed clinical intervention. In this paper, we argue that mining online social behaviour provides an opportunity to actively identify SNMDs at an early stage .It is difficult to find SNMDs as a result of the mental standing can't be directly discovered from on-line group action logs .Our approach, new and innovative to the observe of SNMD detection, doesn't admit self-revealing of these mental factors via questionnaires in psychological science solely. But also, we've got an inclination to propose a machine learning framework, namely, Social Network Mental Disorder Detection (SNMDD) , that exploits features extracted from social network data to accurately identify potential cases of SNMDs .We conjointly exploit multi-source learning in SNMDD and propose a brand new SNMD-based Tensor Mode l(STM) to boost the accuracy. To increase the quantifiability of memory, we further improve the efficiency with performance guarantee .Our framework is evaluated via a user study with online social network users. The results manifest that SNMDD is promising for distinctive on-line social network users with potential SNMDs.

Key Words

Online social network, Social network mental disorder detection, feature extraction, data logs, feature ranking, tensor decomposition, datasets.

1. INTRODUCTION

In today's world, with the explosive growth in quality of social network-ing and electronic messaging apps, online social networks and various (OSNs) became a neighborhood of the many people's daily lives. Most analysis on social network mining focuses on discovering the data behind the information for up people's life. whereas OSNs on the face of it expand their users' capability in increasing social contacts, they will truly decrease the face-to-face social interactions within the universe. Because of this large scale of these phenomena, new terms like Phubbing (Phone-Snubbing)and the another which is very important one is Nomophobia (No mobile Phobia) are introduced to explain these, that cannot stop victimisation mobile social networking apps. The symptoms are a unit the foremost vital parts of diagnostic criteria for SNMDs [1] e.g., excessive use of social networking apps – typically related to a loss of the sense of your time or a neglect of basic drives, and with-drawl – as well as feelings of anger, tension, and/or de- pression when the computer/apps area unit are inaccessible. SNMDs area unit social-oriented and have a tendancy to happen to users who usually typically move and interact with others via online social networks. Those with SNMD typically lack offline interactions, and as area unit cyber-relationships to compensate. Today, identification of potential mental disorders usually falls on the shoulders of supervisors (such as academics or parents) passively. How- ever, since there area unit only a few notable physical risk factors, the patients typically doesn't actively look for medical or psycho-logical services. Therefore, patients solely look for clinical interventions once their conditions become terribly severe. Although in previously the

psychology had identified several serious mental measures associated with SNMDs, they mostly examined as standard diagnostic criteria in normal survey questionnaires. To automatically detect potential SNMD cases of OSN users, extracting those factors to examine users online mental states are extremely challengable. For example, the affect of loneliness and the effect of disinhibition of OSN users are hard to observe. Therefore, there is a need to develop new approach for detecting different SNMD cases of OSN users. We argue that mining the social media data of people as a complementary alternative to the con- venational psychological approach which provides a superb opportunity to actively identify those cases at an early stage. In this paper, we develop a machine learning framework for detecting SNMDs, which we call as Social Network Mental Disorder Detection (SNMDD).With the help of machine learning techniques obtained with the current OSN users and the maximum. Diagnostic practices in psychology[1],we should extract and analyze the following serious categories of features from OSNs: 1) social comparison with people, 2) social organization, 3) social diversity in public, 4) para-social relationships, 5) online and offline interaction ratio, 6) social capital, and 7) bursting temporal behavior. The features are captured which are important factors or serve as proxies for SNMD detection. For example, studies manifest that users exposed to positive posts from others on Instagram with similar bio are inclined to feel malicious envy and depressed due to the social comparison [3]. The depression leads users to disorder behaviours, such as information overload or net compulsion .We here will not only give attention to the relationships caused by social media but also we will give attention on the intensity of those relationships to check the disorder state.

2. RELATED WORK

Research on the mental disorder in online social network receives an increase amount of attention recently. Among them, content-based textual features are extracted from user- generated information (such as blog, social media, profile, YouTube history) for sentiment analysis and topic detection [1]. An NLP-based approach is there to collect and extract precise and content-based features from online social media to identify the Borderline Personality Disorder and the Bipolar Disorder patients. And then extract the topical and linguistic features from online social media for depression patients to analyze their patterns and state of mind in current situation.Analyse emotion and linguistic styles of social media data for Major Depressive Disorder (MDD) [1]. However, most previous research focuses on individual behaviours and their generated textual contents but do not carefully examine the structure of social networks and potential Psychological features. Moreover, the developed schemes are not designed to handle the sparse data from multiple OSNs. In contrast, we propose a new multi-source machine learning approach, i.e., STM, to extract proxy features in Psychology for different diseases that require careful examination of the OSN topologies, such as Cyber-Relationship Addiction and Net Compulsion. Our framework is constructed upon SVM (support vector machine), that has been wide want to analyze OSNs in several aspects and areas. However, the SNMD knowledge from totally different OSNs could also be incomplete because of the heterogeneousness However, the SNMD knowledge from totally different OSNs could also be incomplete because of the heterogeneousness [3] [5]. For instance, the profiles of users could also be empty because of the privacy issue, totally different functions on different OSNs(e.g., game ,check-in ,event), etc .We propose associate degree tensor-based approach to deal with the issues of victimization heterogeneous data and incorporate domain data in SNMD detection.

3. EXISTING SYSTEM OF SNMD FRAMEWORK

In this scenario, we aim to discuss various data mining framework to detect 2 types of SNMDs : 1) Cyber-Relations (CR) Addictions, which includes the

addiction to social networking, checking and messaging to the point where social relationships to virtual and online friends become more important than real-life ones with friends and families offline. 2) Information Overloading (IO), which includes addiction to surf user's status and news feeds, leading to lower work productivity and fewer social interactions with families and friends offline [3].Accordingly, we find the detection of SNMD cases as classification problem. We detect every type of the SNMDs with a binary SVM. In this study, we propose a two-phase framework, called Social Network Mental Disorder Detection (SNMDD), as below in Figure. [1]We tend to specialise in extraction of discriminative and informative options for style of SNMDD. This task could be a non trivial for the subsequent 3 reasons. 1. Lack of mental options. Psychological studies have shown that a lot of mental factors square measures associated with SNMDs, e.g., low shallowness [3] loneliness. Thus, questionnaires unit designed to reveal those factors for SNMD detection. The first half extracts varied discriminative choices of users, whereas the second part presents a replacement SNMD-based tensor model to derive latent factors for coaching and use of classifiers designed upon Transductive SVM (TSVM) [7].Two key challenges exist in fashion of SNMDD: i) we've a bent to not able to directly extract mental factors like what square measures done via questionnaires in scientific discipline and thus would like new choices for learning the classification models; [4],ii) we've a bent to aim to use user data logs from multiple OSNs and thus would have new techniq

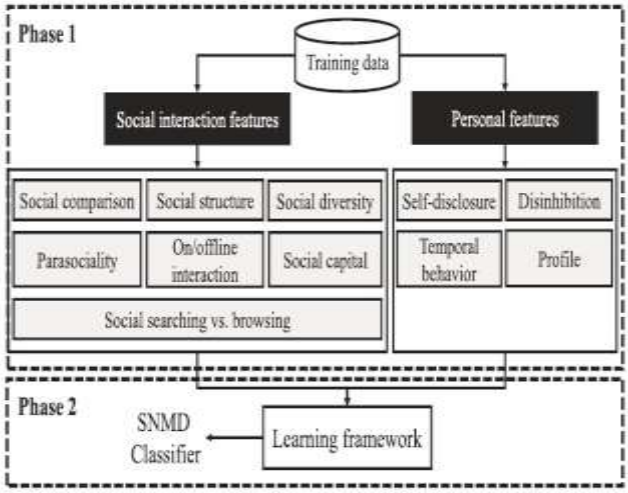

Fig. 1. The SNMDD framework.

ues for integration multi-source data primarily based on SNMD characteristics.

4. PROPOSED SYSTEM

The proposed system explained in this paper , we' re going to extract data from multiple social media sites like face book , instagram , twitter and history of various application like YouTube history. This will be very helpful in extracting the behaviour and the mental stress level or state of the user. In the first step we will take data of all users' accounts and side by side we will form a set of questionaries for the user for some more clarification. In the next phase we extract the preprocessed data and give various discriminative features of users. Many users are inclined to use different OSNs, and it is expected that data logs of these OSNs could provide enriched and complementary information about the user behaviour. Thus, we have a tendency to aim to explore multiple information sources (i.e., OSNs) in SNMDD, so as to derive a additional complete portrait of users' behavior and effectively cope with the information Scarcity problem.

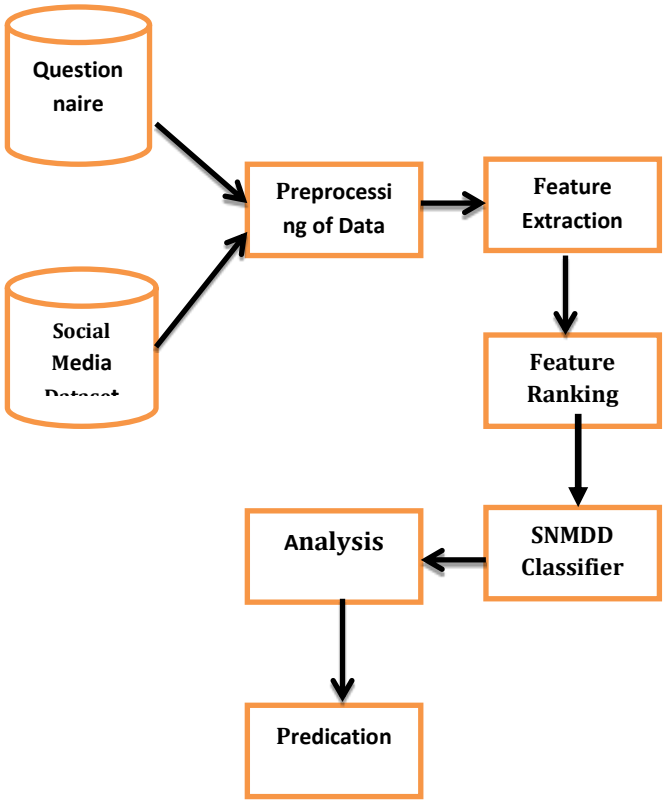

Fig - 2: System Architecture diagram

To exploit multi-source learning in SNMDD, one straightforward means is to directly concatenate the options of every person derived from totally different OSNs as an enormous vector. According to the ranking set in the model, the extracted data is placed in the table according to their rankings. The ranked data is saved in the database and the dataset is then classified in SNMDD classifier and at the end the data is analyzed in the analysis phase. After the analysis phase the predicted data is carried out from the dataset and the predicted analysis of users mental health is predicted and given to the user or if required to the doctor if the users want to give it.

5. Model Buiding

To find out the simplest acting models, the subsequent machine learning and deep learning algorithms were thought of and enforced on the s4 dataset:

I) KNN Naïve-Bayes

Within the multivariate event model, options area unit freelance Booleans (binary variables) describing inputs. Just like the multinomial model, this model is standard for document classification tasks, wherever binary term prevalence options area unit used instead of term frequencies. If x_i maybe a Boolean expressing the prevalence or absence of the ith term from the vocabulary, then the probability of a document given a category C_k is given by

$$(x|C_k) = \prod_{i=1}^{n} p_{ki}x_i (1-p_{ki})^{(1-x_i)}$$

Where as p_{ki} is the probability of category C_k performing the term x_i. This event model is specially popular for classifying little texts. It has the benefit of explicitly modelling the absence of terms. When implemented on the dataset, it had an accuracy of 0.801786%.

II) Multinomial Naïve-Bayes

With a multinomial event model, sample (feature vectors) represent the frequencies with that sure events are generated by a multinomial (p1...,pn) is that the chance that event i happens (or K such that multinomials within the multiclass case). A feature vector x = (x1,.....,xn) is then a bargraph, with xi tally the quantity of times event i used to be discovered during a explicit instance. This can be the event model usually used for document classification, with events representing the incidence of a word during a single document. The probability of observant a bargraph x is given by

$$P(x|C_k) = \frac{(\sum_i x_i)!}{\prod_i x_i! \prod_i p_{ki}^{xi}}$$

If a given category and have worth never along within the training data, then the frequency-based

chance estimate are going to be be zero. This can be problematic as a result it'll wipe out all the data with the different probabilities once there square measure are increased. Therefore, it's typically fascinating to include a small-sample correction, referred to as pseudo count, in all chances estimates specified no chance is ever set to be precisely zero. This manner of regularizing Naïve Bayes is named Laplace smoothing. Implementing this algorithmic rule on the dataset gives an accuracy of 1.0000%

III) Random Forest Classifier

Random Forest learning is that the construction of a decision or call tree from class-labelled coaching tuples. A random forest could be a flow- chart-like structure, wherever every internal (non-leaf) node associates n a nursing attribute, every branch represents the end result of a test, and every leaf (or terminal) node holds a class or category label. The uppermost node in an exceedingly tree is that the root node. Classification and Regression Tree (CART), repetitious Iterative Dichotomiser 3(ID3) and Chi-squared Automatic Interaction Detector (CHAID) area unit few styles of call tree learning algorithms.

A problem arises once exploitation ancient RNNs for information science NLP tasks as a result of the gradients from the target operate will vanish or explode once many iterations of multiplying the weights of the network. For such reasons, easy RNNs have seldom been used for information science NLP tasks like text classification. In such a situation one will inter-communicate another model within the RNN family like as the LSTM model. LSTMs are higher suited to the current task because of the presence of input gates, forget gates, and output gates, that management the flow of knowledge through the network. The dataset when used to train this algorithmic rule gave an outputs of an accuracy of 0.985714% .

IV) Support Vector Machine

The goal of the SVM algorithmic rule is to make the simplest line or decision or call boundary that may segregate n-dimensional house into classes or categories so we will simply place the new information within the correct class within the future. This best decision or call boundary is termed a hyperplane. SVM chooses the intense points/vectors that facilitate in making the hyperplane. These extreme cases square measures are referred to as support vectors, and therefore algorithmic rule is termed as Support Vector Machine. Contemplate the below diagram within which there are square measures 2 totally different classes that square measures classified employing a decision or call boundary or hyperplane. The dataset when used to train this algorithmic rule, outputs an accuracy of 0.789286%.

Table 1. Results

Algorithm Used	Accuracy
KNN	0.801786
Multinomial Naïve Bayes	1.000000
Random Forest	0.985714
SVM	0.789286

5.1. RESULTS

For the classification approach, Multinomial naïve bayes is used because they generally have a superior performance than other algorithms for learning relationships.

6. CONCLUSION

In this paper, we tend to build an endeavour to mechanically establish potential in online users with SNMDs and questionaries from multimedia system contents. We propose associate degree SNMDD framework that explores varied options from knowledge logs of OSNs and a innovative tensor technique for derivation latent choices from multiple

OSNs for SNMD detection. Simultaneously we'll type a collection of questionaries for the user for a lot of accuracy. This work represents a cooperative effort between laptop scientists and mental aid researchers to handle rising problems in SNMDs. We plan to extract data from multimedia social sites and detect the mental health of the users by using SNMDD framework .We also plan to further explore new issues from the perspective of a social network service provider to improve the well-bings of OSN users without compromising the user engagement.

REFERENCES

1)Shuai, H. H., Shen, C. Y., Yang, D. N., Lan, Y. F. C., Lee, W. C., Philip, S. Y., & Chen, M. S. (2017). A comprehensive study on social network mental disorders detection via online social media mining. IEEE Transactions on Knowledge and Data Engineering, 30(7), 1212-1225.

2)King, D. L., Delfabbro, P. H., Kaptsis, D., &Zwaans, T. (2014). Adolescent simulated gambling via digital and social media: An emerging problem. Computers in Human Behavior, 31, 305-313.

3)Baek, Y. M., Bae, Y., & Jang, H. (2013). Social and parasocial relationships on social network sites and their differential relationships with users psychologicalwellbeing. Cyberpsychology, Behavior, and Social Networking, 16(7), 512-517.

4)Ferrara, E., Interdonato, R., &Tagarelli, A. (2014, September). Online popularity and topical interests through the lens of instagram. In Proceedings of the 25th ACM conference on Hypertext and social media (pp. 24-34). ACM.

5)Guellil, I., &Boukhalfa, K. (2015, April). Social big data mining: A survey focused on opinion mining and sentiments analysis. In 2015 12th International Symposium on Programming and Systems (ISPS) (pp. 1-10). IEEE.

6)Shuai, H. H., Shen, C. Y., Yang, D. N., Lan, Y. F., Lee, W. C., Yu, P. S., & Chen, M. S. (2016, April). Mining online social data for detecting social network mental disorders. In Proceedings of the 25th International Conference on World Wide Web (pp. 275-285). International World Wide Web Conferences Steering Committee.

7)Rahman, R. A., Omar, K., Noah, S. A. M., &Danuri, M. S. N. M. (2018). A Survey on Mental Health Detection in Online Social Network. International Journal on Advanced Science, Engineering and Information Technology, 8(4-2), 1431-1436.

8)Nasution, M. K., Sitompul, O. S., Sinulingga, E. P., & Noah, S. A. (2016, July). An extracted social network mining. In 2016 SAI Computing Conference (SAI) (pp. 1168-1172). IEEE.