REVIEW OF TECHNQIUES USED TO DETERMINE ABNORMALITIES WITHIN PRESENTED DATA

Punam Devi, Ankur Sharma, Anurag Rana

ABSTRACT

This paper presents comprehensive analysis of mechanisms used to determine the abnormality presents within the humans in terms of lack of concentration due to online social media along with usage of mobile phones. Main problem caused which may remains unidentified is no-mo-phobia. This literature also discusses the supervised and unsupervised learning mechanisms used through the use of tools and techniques associated with platforms like MATLAB, Netbeans etc. the dataset formation is critical in forming the result. The parameters required to check the result accuracy is also discussed through this literature. The comparative study is also presented to gives the desired future scope and selection of techniques for future enhancement.

Keywords: No-Mo-Phobia, Supervised, Unsupervised learning, dataset

INTRODUCTION

The use of technology leading to advancement in terms of execution time and causing better convergence rate. The technology although offers advantages but still over utilization is causing the problems of irritation and extra pressure. This problem if is caused by the application of mobile phones then it is initiation of No-Mo-Phobia. As this problem aggravates, users of the mobile phone cannot last much time without using the phone.[1]proposed a learning mechanism to check the affect of no-mo-phobia on the work which is being done within the organization. Total of 187 employees are tested for no-mo-phobia and there implication on organization work is checked using questionnaires. [2] proposed DSM-V mechanism to determine the anxiety level if the person with and without the application of mobile phone. The mechanism check the anxiety level which if exceeds the threshold value then No-Mo-Phobia is detected otherwise No-Mo-Phobia is not detected.

[3], [4]Techniques associated with supervised and unsupervised learning is available and can be used for determining the abnormality if any present within the humans. The abnormality detection procedure could be of distinct categories. The worth of the generated result is determined using parameters such as classification accuracy, specificity and sensitivity.

[5]Metrics evaluation is critical in the detection of success associated with the classifier. The classification accuracy is given by using the following equation

$$Accuracy = \frac{Correct_{Pre}}{Total_{Pred}}$$

Equation 1: Accuracy in terms of prediction

[6]Sensitivity is obtained by dividing number of positive predictions to the total true positive rate.

$$\text{Sensitivity} = \frac{Correc\ t_{Positiv\ e_{predictions}}}{Tota\ l_{Positives}}$$

Equation 2: Sensitivity evaluation formula

[7]Specificity is another parameter used to evaluate correctness of the proposed system. It is given as under

$$\text{Specificity}=\frac{True_{Negitives}}{TP+FN}$$

Equation 3: Specificity obtaining formula

Rest of the this paper is organised as follows: section 2 gives the literature survey, section 3 gives the comparative analysis of all the tools and techniques discussed in section 2, section 4 gives the conclusion and future scope and last section gives the references.

LITERATURE REVIEW

This section gives the in-depth of various techniques used to determine the problems within the data elements. The data elements which are abnormal can be detected by comparing it against the predefined threshold values. Most prominent techniques used to detect abnormality is discussed as under

[8] Masood & Al-Jumaily 2015 proposed support vector machine for the detection of skin cancer. Support vector machine is one of the commonly used learning mechanisms which is efficient enough in the detection of abnormal behaviour within text or image formats. The threshold values are also known as labelling information which is associated with the data. The data from which features are extracted and act as baseline are known as train data elements and the data which is compared against the train data is known as test data. Support vector machine has only two hyper planes associated with it. First hyperplane gives the normal values and second hyperplane gives the abnormal values. the prediction is on the basis of penetration. In case values penetrates the first hyperplane then first set of results are generated and if second hyperplane is intersected then second disease is predicted. To implement the svm, MATLAB is used in this case.

[9] Tams et al. 2018 proposed a novel strategy for the detection of No-Mo-Phobia which is a prime reason of stress in humans now days. This disease cannot be detected easily and hence causes road accidents, restlessness etc. this disease is uncommon and goes undetected and cause the harm even to the life of humans. This work conducts the case study of individuals to detect the phobia and also presents some mechanism to rectify the issues. The detection is on the basis of case study and no specific tool is used for this purpose.

[10] Chauhan & Vania 2013 proposed a J48 classifier in order to detect the problems present within the soil. This abnormality is presented using the metric classification accuracy. The classification accuracy if is high, can cause the faster detection of the problems within the concerned data elements. Wekka is used to implement J48 classifier.

[11] Wang & Kim 2016 proposed naïve bayes approach in order to detect the traffic on road. The mechanism produces accurate prediction of on road traffic. The validity of the result is judged using the parameter classification accuracy. The classification accuracy is achieved by subtracting the actual value from the approximate value. Higher the accuracy less will be the mean square error. Naïve bayes is implemented using wekka tools.

[12] Kureshi et al. 2016 proposed support vector machine based approach in order to detect the abnormality in terms of lung cancer. The support vector machine used in this work has only two hyper planes associated with it. The abnormality detection is presented through the parameter classification accuracy. The support vector machine is implemented by the use of MATLAB integrated with wekka for classification.

Next section gives the comparison of literature discussed above.

COMPARISON OF LITERATURES

The comparative study gives the mechanism to determine the best possible approach for future enhancement. The comparative study is given as follows

Author	Technique	Tool	Advantage	Disadvantage
Masood & Al-Jumaily 2015[8]	Support vector machine	MATLAB integrated with machine learning SVM	Classification accuracy is high	Complexity is high which can be reduced in case NETBEANS with WEKA is used for classification
Tams et al. 2018 [9]	Case study	Case Study	Complexity is low and result is derived using real time data elements	No algorithm or standardised formula is used and hence result could be different in different situations.
Chauhan & Vania 2013[10]	J48	Weka	Multiple metrics such as specificity , sensitivity etc are used to make the result multivariate	Classification accuracy can be further improved.
Wang & Kim 2016[11]	Naïve Bayes	Weka	This approach is efficient in case of nominal data	In efficient in the handling of large and string type data.
Kureshi et al. 2016[12]	Support vector machine	MATLAB	Matlab integration with Weka produces least classification error	Integration of Weka with MATLAB is complex in nature.

Table 1: Comparative analysis of tools and techniques used for the detection of diseases.

CONCLUSION AND FUTURE SCOPE

This work presents comprehensive analysis of tools and techniques used to detect the abnormality within the data presented. The testing and training data used within the existing literature decide the usefulness of study being conducted. The result obtained with the help of parameters such as classification accuracy and specificity. Difference in testing and training data also gives the reliability of information used. Most efficient tool used in the existing literature appears to be weka and netbeans which can be used in future for performance enhancement.

REFERENCES

[1] G. Wang, "Disorder or Driver ?: The Effects of Nomophobia on Work-Related Outcomes in Organizations," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI'18*, pp. 1–12, 2018.

[2] N. L. Bragazzi and G. Del Puente, "A proposal for including nomophobia in the new DSM-V," *Psychol. Res. Behav. Manag.*, vol. 7, pp. 155–160, 2014.

[3] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," *Procedia Technol.*, vol. 10, pp. 85–94, 2013.

[4] D. Mahapatra, P. Schueffler, J. A. W. Tielbeek, J. M. Buhmann, and F. M. Vos, "A supervised learning approach for Crohn's disease detection using higher-order image statistics and a novel shape asymmetry measure," *J. Digit. Imaging*, vol. 26, no. 5, pp. 920–931, 2013.

[5] B. S. Kumar and R. Anbuselvi, "Image Mining Techniques to Enhance the Classification Accuracy on Brain Glioma," *IEEE Access*, vol. 5, no. 3, pp. 525–539, 2016.

[6] T. Zhang, Y. Zhou, and C. L. P. Chen, "A new combined chaotic system for image encryption," *CSAE 2012 - Proceedings, 2012 IEEE Int. Conf. Comput. Sci. Autom. Eng.*, vol. 2, no. 2, pp. 331–335, 2012.

[7] S. Suman, F. A. Hussin, A. S. Malik, S. H. Ho, I. Hilmi, A. H.-R. Leow, and K.-L. Goh, "Feature selection and classification of ulcerated lesions using statistical analysis for WCE images," *Appl. Sci.*, vol. 7, no. 10, 2017.

[8] A. Masood and A. Al-Jumaily, "SA-SVM based automated diagnostic system for skin cancer," *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 9443, no. Icgip 2014, p. 94432L, 2015.

[9] S. Tams, R. Legoux, and P. M. Léger, "Smartphone withdrawal creates stress: A moderated mediation model of nomophobia, social threat, and phone withdrawal context," *Comput. Human Behav.*, vol. 81, pp. 1–9, 2018.

[10] Y. Chauhan and J. Vania, "J48 Classifier Approach to Detect Characteristic of Bt Cotton base on Soil Micro Nutrient," *IEEE Access*, vol. 5, no. 6, pp. 305–309, 2013.

[11] G. Wang and J. Kim, "The Prediction of Traffic Congestion and Incident on Urban Road Networks Using Naive Bayes Classifier," *IEEE Access*, no. November, pp. 1–14, 2016.

[12] N. Kureshi, S. S. R. Abidi, and C. Blouin, "A predictive model for personalized therapeutic interventions in non-small cell lung cancer," *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 1, pp. 424–431, 2016.