

Prediction of Depression in Human Behaviour using Deep Learning Neural Network

Sandeep Kumar Singla, Rupinder Kaur

Guru Nanak Dev Engineering College, Ludhiana, Punjab, India.

sandeepkumar.singla@gmail.com, rupinderk1203@gmail.com

Anu Bala

B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

anu_singla13@yahoo.com

ABSTRACT

With the explosive growth of social media, the way of interaction with individuals has changed and it has also made interaction faster and more effective. Sentiment Analysis is the most fascinating and an alluring field of the research which involves linguistic processing, computational semantics or interpretation of content in order to identify the polarity of the text. The main aim of this method is to proficiently recognize the feelings conveyed in reviews of different domains like news articles, stock prices, e-learning field and product reviews etc. With the enormous expansion of web content, complexity of extracting sentiments from text is increasing. In this paper, sentiment analysis using k-Mean clustering and Back Propagation Neural Network was proposed to study the classification efficiency of product reviews. The performance results were evaluated in terms of false acceptance rate, false rejection rate and accuracy. The integration of clustering and classification gives much better results than the existing state of the art methods. For understanding the sentiments of customer reviews this approach proves to be an efficient method.

KEYWORDS

Sentiment analysis, Feature Extraction, k-Mean cluster, Back Propagation Neural Network

INTRODUCTION

Depression is the main cause of many diseases. It is a serious illness caused by changes in brain chemistry. Depression usually results from a combination of recent events and other longer term or personal factors, rather than one immediate issue or event. Based on the principle of “Chinta chita saman hai” . “Chinta is mainly known as worry that affect the peace of mind. Hypertension is also the main cause. No particular medical test is available for depression rather it can only be detected through symptoms which lead to depression . If the person is building up negative

thoughts then depression can be detected, on the other hand if the person’s approach is towards positive then there will be no depression in that case. Medication may help relieve some of the symptoms of depression, but it doesn’t cure the underlying problem. Our aim was to develop a model to represent user activity on social health networks. This enabled us to accurately represent user interactions by relying on the data’s semantic content.

The analysis in progressing countries declared that maternal depression might be a threat for low development of young children (Rahman et al, 2008). This threat explains that maternal psychological health in countries having low-income may have essential effect on development during childhood. The influence of depression not only affects this generation but also the next generation.

RELATED WORK

To get an idea about the prediction of depression in human behavior using the recent advancements in the field of computer science and IT, a number of research papers, notes, and books has been referred which are summarized below.

Getoor and Diehl, (2005) in their paper titled “Link mining: a survey”, Link mining is an emerging area within data mining that is focused on finding patterns in data by exploiting and explicitly modeling the links among the data instances. They have surveyed several of the more well studied link mining tasks: link-based object ranking, link-based object classification, group detection, entity resolution, link prediction, subgraph discovery, graph classification, and generative models for graphs. These represent some of the common

threads emerging from a variety of fields that are exploring this exciting and rapidly expanding field. They introduced probabilistic relational models, which that provide a unified generative model for objects and link structure. The survey described each link mining task in isolation. More generally, component link mining algorithms may be part of a larger knowledge discovery process. In some cases, all the links would not be observed and hence they predicted the existence of links between instances. In the domains, where the links are growing over time, their motive was to predict if a link would exist in future by observing the previous links. The patterns become more complex by considering links. This leads to other challenges focused on discovering substructures, such as communities, groups, or common subgraphs.

Lu. and Getoor, (2003) proposed a framework for modelling link distributions, a link based model that supports discriminative models describing both the link distributions and attributes of linked objects. The domains they examined include a combined logistic classifier built over object attribute and link statistics outperforms are- a simple content only classifier and a single flat classifier over both the content and link attribute. The mode of link statistics is not enough to capture the dependence. Actually modelling the distribution of the link categories at the finer grain is useful. In all cases link distribution improves the accuracy of the classification based on link and objects.

A.Ng, A. Zheng, and M. Jordan, (2001) examined the strength of HITS and PageRank to little activation of document gathering. They endeavor the automation of analysis and intimate how it gives acknowledgment into methods for designing stable link investigation techniques. The algorithm, Subspace HITS, is motivated by the observation that subspaces spanned by a few eigenvectors may sometimes be stable even when individual eigenvectors are not. They also composed the experimental performance of the four algorithms, and thus examined the inception of “diversity” of the output given by the algorithms. This leads into a discussion of the relationship between Latent Semantic Indexing (LSI) and HITS.

Taskar et.al , (2003) focused on predicting the existence and the type of links between entities in relational domains. They focused on the task of collective link classification, where they are trying to predict and classify an entire set of links correspondencely in a link graph. They have found two types of patterns that can be useful to improve the classification accuracy over flat models. The Similarity templates ,which relate the classification of links or objects that share a certain graph-based property (e.g., links that share a common endpoint) and the Transitivity templates ,that relate triples of objects and links organized in a triangle. These methods have been applied to the problem of predicting or classifying a single link at a time. They applied PRM and RMN framework for this task. The trained RMNs perform significantly better than generatively trained PRMs because the PRM framework cannot represent (in any natural way) the type of subgraph patterns that seem prevalent in link graph data.

Nowell et.al, (2004) formalized the inquiry as the connection expectation issue, and created techniques to interface forecast in view of measures for detecting the "vicinity" of nodes in a network. Experiments on substantial co-creation systems recommend that data about future associations can be extracted alone from network topology, and that genuinely perceptive measures for discovering node vicinity can outperform more straightforward measure. They utilized an irregular indicator as a norm which merely arbitrarily chooses sets of creators who did not work as a team in the training time. The objective is to enhance the efficiency of the nearness construct techniques with respect to extensive systems; quick calculations for approximating the dispersion of point to point separations might be one approach. The experiment treat all preparation period reconcile efforts similarly. Maybe one can enhance execution by regarding later coordinated efforts as more imperative than more established ones.

Akay et.al, (2016) build up a weighted network model to reflect users actions on social wellbeing systems which empowered them to precisely reflect user associations by depending on the information's semantic content. The three-stage strategy utilizes the weighted system model to show user interactions, and

system clustering and module detection to portray user communications and concentrate promote learning from users posts. The system displaying approach yielded an underlying inexactly associated arrange, connecting all clients inside the depository. The network's topological properties shows user activities such as posts' general topic as well as timing, whereas weighted system reflect the semantic content of post and relation among posts. A synthesis from word data frequency, statistical analysis of module content and the modeled health network's properties, has allowed them to gain insight into consumer sentiment of antidepressants. This approach will enable all gatherings to take part in enhancing future wellbeing arrangements of patients experiencing despondency. To finish up, they trust that the utilization of keen information mining instruments is a chance to extraordinarily enhance the nature of social health of buyers, medicinal services specialists, and the business while diminishing expenses.

Choudhury et.al, (2013) displayed work at utilizing a crowdsourcing philosophy to fabricate a vast corpus of postings on twitter that have been shared by people determined to have clinical dejection. Next, they built up a probabilistic model prepared on this corpus to decide whether posts could demonstrate depression. The model use signs of social activity, feeling and language showed on twitter to develop the SVM classifier. Utilizing this model they presented an online networking depression file that may serve to describe the levels of depression in people. The classifier anticipated with high exactness (73%) regardless of whether a post on twitter could be depression indicative.

Akay et.al (2013) proposed a two-stage analysis system that spotlights on positive and negative opinion and also the reactions of treatment. They utilized a self sorting out guide to analyze word recurrence data achieved from users gathering posts. They at that point presented a novel system based approach for demonstrating approach users gathering connections and utilized a network partitioning technique in view of enhancing a concord quality measure. This enabled them to decide consumer supposition and distinguish influential users inside the recovered modules utilizing

data carried from both word-recurrence information and system based properties.

Wu et.al (2012) introduced that the inter-sentential language patterns can catch relationship among numerous words inside and between sentences, therefore can give more exact data than word sets. To achieve inter sentential language designs, they built up a content mining structure by expanding the traditional association based rule mining process with the end goal that it can find as often as possible co-occurring designs over the sentence limit. Execution was predicted on a pile of texts. Experimental outcomes demonstrate that the utilization of inter sentential structures beat the utilization of word sets proposed in past studies.

Noessner et.al introduced a new confront to point collaboration that depends on a current semantic closeness measure for connected data. They assimilate the measure to object collaboration issue, introduce correct and estimated calculations that effectively execute the techniques, and give a deliberate experimental assessment in view of a target dataset. As our fundamental outcome, we demonstrate that the utilization of lightweight data structure and composition data altogether enhances question compromise with regards to connected open data. They proposed a system for object compromise on the basis of a semantic uniformity value between A-Boxes. They contended that most cutting edge approaches for bit coordinating concentrate exclusively on approaches to figure lexical likenesses. These methodologies are a few- times stretched by structural approval system where class organisation is utilized as a coordinating filter.

Rosenquest et.al (2010) analysed thickly interconnected informal community of 12 067 individuals surveyed several times over 32 years as a feature of the Framingham Heart Study. Longitudinal factual models were utilized to look at whether depressive manifestations in one individual were related with comparable scores in companions, associates, spouse, kin, and neighbors. Symptoms of depression were surveyed utilizing CES-D scores that were accessible for subjects in three waves measured in the vicinity of 1983 and 2001. Results indicated

both low and high CES-D scores in a given period were emphatically connected with such scores in one's companions and neighbors. This affiliation reached out up to three degrees of division (to one's companions). Female companions seems to be peculiarly prestigious in spreading depression from one individual to the next. The results also show that friends over time are known to influence more than spouses. At last, the FHS interpersonal organization data collection is surprising because of its longitudinal nature and relative consistency . The researchers expect that the further gathering of interpersonal organization information in different settings will give more chances to assess such speculations.

Choudhury (2013a) highlighted some current endeavors looking at the potential for utilizing online networking postings as another kind of main point in understanding dysfunctional behavior in individuals. Data gathered from online networking bears potential to compensate conventional overview methods in its capacity to give better grained estimations of conduct after some time while population test sizes grows in a radical manner after some time. The researcher aim was to spotlight how this research petition might be helpful in creating devices for enumerating the attack of depressive issue in people, for the benefit of people, or for the use by human safety organizations, sanctioning those to be more active about their intellectual wellness.

Tsugawa et.al (2015) examined the adequacy of utilizing a client's online networking exercises for assessing level of depression. They analysed the data of 209 users. SVM and LDA classifiers were used to estimate the existence of proactive depression among users. They used the consequences of online questionnaire as ground truth information for measuring level of depression of Twitter clients. They extricate a various components from the histories of Twitter users. The outcome demonstrated that the activity history of clients can be employed to perceive the existence of depression with a precision of roughly 69%. The future task can be to predict technique that could be assessed to estimate depression from user's social sites postings.

Nguyen et.al (2013) examined the qualities of online depression groups (CLINICAL) in comparison with those joining other online groups (CONTROL). Three perspectives have been inspected: influence, psycholinguistic procedures and points inside the content. Machine learning and factual strategies were utilized to separate online messages amongst gloom and control groups. All viewpoints - influence, the composed substance and composing style were observed to be fundamentally unique between these two gatherings. Furthermore, inert points were found to have greatest power of prediction than lingual features for predicting depression groups. This investigation recommends that information mining of online web journals has the ability to recognize significant information for depression studies. They found that clinically depressed people reflects more negative emotions. At the last they concluded that clinical group possess low strength then those in control group as detected from sentimental analysis.

Mok (2014) investigated the relation between social media and different age groups. The authors proposed the early depression detector to identify and control the risk of social media usage. They surveyed at Singapore and observed that among the total participants 48.9% used social media from 5-7 days per week. 9.8% used gaming platforms, 69.2% used video sites and furthermore 82.3% among them used social site. They observed that senior citizens above the age of 55 years were not using social media. The results showed that among all age groups in both males and females mostly the teenaged females were supersensitized to risk of depression.

Pau et.al (2014) depicted a statistical topic modeling process to recognize the health information from millions of health-related tweets. They also described the Ailment Topic Aspect Model (ATAM) to filter general data based on health keywords and supervised classification. They explored that how ATAM and other topic models on their own analysed health topics in 144 million tweets. Instead of focusing on a single health issue, their purpose was to identify different multiple health issues. Their main purpose was to directly evaluate the coherence of ATAM ailments. The results explored that it is possible to discover topics automatically that achieve statistically considerable interdependence with ground truth data,

without using minimum human supervision and no previous data to train the model. In addition to this, the results also showed that an individual general-purpose model alone can predict several health topics in social media.

Zhao et.al (2014) proposed a novel metric to identify influential users in online health communities. The author considered number of influential responding replies (IRRs) metric to directly measure an individual's capability to influence the sentiments of others. The research not only considered how much an online health community participant has contributed to community, but also how his/her contributions has affected others. The study does not distinguish between healthy negative sentiment influences (eg, sadness due to the death of a community member) and those that are not healthy for the community.

Culjak (2006) examined the Theory of Behavior Change how the Internet provides pathways to care through widespread access and awareness of depression and related psychological disorders than previously possible through traditional media. It further revealed how such access and awareness influences individual behavior and increases the motivation of depression sufferers to reach out and get the help they need. The results show that the use of Internet self help sites for depression does influence health behavior and that this influence is positive. It identifies how changes in behavior lead to a decrease in the incidence, severity and longevity of depression. Recommendations for further research are outlined.

iii. RESEARCH METHODOLOGY

Methodology means the flow of the work that has been already planned to develop. This section presents the steps in the methodology (Figure 1) that this research study will follow in order to achieve the objectives of the study. The process is divided into two sections i.e. preprocessing and processing section.

A. Pre-processing Stage: This stage includes data collection, extracting features, removing stop words, and assigning tokens.

Step1: Collect the data from online social media.

Step2: Extract the required information into database.

Step3: Load the database.

Step4: Extract features.

Step5: Tokenize the parameters.

Step6: Remove stop words.

B. Processing Stage: In this stage, eight parameters have been used to classify if the post is depression-indicative or not. Time stamp detail of the post is also been considered in order to describe whether it has been posted during day time or in night. To predict depression, the classifier is trained by implementing Deep neural network classifier.

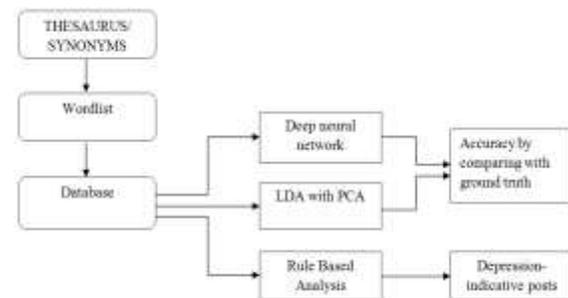


Figure 1: Proposed Methodology

Deep Neural Network Technique- This technique is applied to classify a huge amount of data with maximum accuracy. It is very efficient to recognize pattern problems, complicated decision bound problems over many variables. The network is designed using the traits of neighbors to train the system to produce the exact target class. First of all the classifier is trained, then validated and at last it is tested to give the accurate output. Usually it works on two-layers.

The NN Training Tool shows the network being trained and the algorithms used to train it. It also displays the training state during training and the criteria which stopped training will be highlighted in green. Performance is measured in terms of mean squared error, and shown in log scale. It rapidly decreased as the network was trained. Performance is

shown for each of the training, validation and test sets. The version of the network that did best on the validation set is was after training. The mean squared error of the trained neural network can now be measured with respect to the testing samples. This will give us a sense of how well the network will do when applied to data from the real world. Another measure of how well the neural network has fit the data is the confusion plot. Here the confusion matrix is plotted across all samples. The confusion matrix shows the percentages of correct and incorrect classifications. Correct classifications are the green squares on the matrices diagonal. Incorrect classifications form the red squares. If the network has learned to classify properly, the percentages in the red squares should be very small, indicating few misclassifications. If this is not the case then further training, or training a network with more hidden neurons, would be advisable.

Linear discriminant analysis (LDA) classification-

It is a system used to search the simple collection of attributes or characteristics that divides two or more than two classes of entities. The reduced collection normally be used as a simple or linear classifier or for dimensionality reduction . LDA is used to recognize patterns, in statistics and machine learning. LDA is very much related to PCA (Principal Components Analysis), in case both search for linear combinations which best describes the data. PCA works with LDA to produce linear functions in order to provide reduction of data by rearranging the information. The main aim of LDA is to create the maximum class discrimination and that of PCA is to compress the variables into least few components possible. LDA produces equal number of linear functions as that of classes, on the other hand PCA produces equal number of linear functions to the original variables.

iv. EXPERIMENTAL RESULTS

Data gathered from various resources is used to examine the system. The accuracy of the system may vary for different databases. Good and bad days are the part of one’s life like the two sides of coins. Parents, school, friends, relationships, surroundings can be frustrating or irritating. Things can be right at one minute and horrific at the other. In everyday life sadness can be caused due to loss of someone special, accident, breaking of relations. It becomes more serious problem when this sorrow lasts for over more than 2 weeks and interfere one’s life. When someone

talk about feeling tensed or depressed, they might mean they have a sad day or they might talk about clinical depression. No particular medical test is available for depression rather it could only be detected through symptoms which lead to depression. Facebook, twitter and other social sites provide in built mechanisms to instantly share captured multimedia things. Generally, such shared media questions that happen over the span of individuals' daily lives gives high signs to catch their conduct than content alone. Possibly, such shared interactive media can be utilized as a part of conjunction with content to portray the difficulties in the lives of individuals experiencing emotional instability such as depression. The proposed system comes out with the finding that how the behavior of person can be detected to predict the depression-indicative posts. The experiment can be used to detect depression beforehand and it suggests to take proper treatment in case of depression.

The existing system developed a weighted network model to sought out consumer interactions on social health networks. The author developed three-step method first represents consumer’s activity using weighted-network model, secondly uses network clustering and analysis of module to categorize consumer activities and also extract further information from consumers posts. This method sought out only two features positive and negative to predict depression. The proposed system works on 6 features and uses 3 techniques rule based prediction, DNN and LDA method. Deep neural network plots a confusion matrix to represent depression class, non-depression class and misclassifications. The overall accuracy of the system given by LDA and DNN varies. The rule based prediction only predicts if the person is suffering from depression or not.

Table 1 Comparison of two Techniques

Sr. No.	Techniques	Accuracy (%)
1	Deep neural network learning	86%
2	LDA	67%

Deep neural network using softmax layer converts 8 features into 2 features giving 86% accuracy whereas Linear discriminant analysis using principal component analysis takes 8 features as input and changes them to 4 features giving 67% accuracy. The DNN gives higher accuracy than LDA which shows

that DNN is more efficient and effective to explore depression behavior.

We observed that 2 classes are formed :first class shows the persons in depression and the second class shows the persons not in depression. The red class indicates inaccurate classification and the class which is grey in colour shows misclassifications. The system gives 86% efficiency.

It is observed that the actions of users could be utilized to identify the appearance of depression by DNN technique with an efficiency of near about 86% where as approximately 67% accuracy with LDA (Table 1). The study explored the features with which the past activities of users are helpful in identifying the levels of depression.

V. CONCLUSION AND FUTURE SCOPE

Through our experimental findings, we have demonstrated how the social media posts could be extracted to reflect the behavior of person, and also to predict whether the posts are depression-indicative or not. We come to the conclusion that mostly the people suffering from depression are active during night. This system automatically detects the signs of depression from postings and estimates the depression. The user's action on social media are effective to determine the level of depression. The estimates of depression prediction could be used to enable early detection and treatment of depression at right time.

We look at the potential for utilizing web-based social networking postings as another sort of focal point in understanding causes of depression in humans. Although the system predicts depression from various features but in future the features can be improved such as gender, friends list and likes can be added and the advanced predictive model can be developed to automatically gather online data from social sites. Automatically gathering posts would detect signs of depression which will result in better results before hand. The advanced factors can be considered to predict depression with better efficiency.

REFERENCES

- [1] Getoor, L. and Diehl, C.P., 2005, "Link mining: A survey," *SIGKDD Explor. Newslett.*, vol. 7, pp. 3- 12.
- [2] Getoor, Q.L.L., 2003, "Link Based Classification", International Conference of Machine Learning.
- [3] Ng, A.Y., Zheng, A.X. and Jordan, M.I., 2001, "Stable algorithms for link analysis," in *Proc. SIGIR Conf. Inf. Retrieval*, New Orleans, LA, USA, pp. 258–266.
- [4] Taskar, B. Wong, M. F., Abbeel, P. and Koller, D., 2003, "Link prediction in relational data," *Advances Neural Information Processing Systems.*, Vancouver, BC, USA.
- [5] Nowell, D. L. and Kleinberg, J. M., 2004, "The link prediction problem for social networks," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 57, pp. 556–559.
- [6]. Paul, M. J. and Dredze, M., 2014, "Discovering health topics in social media using topic models," *PLOS One*, vol. 9, Issue 8, e103408.
- [7] Zhao, K., Yen, J., Greer, G., Qiu, B., Mitra, P. and Portier, K., 2014, "Finding influential users of online health communities: A new metric based on sentiment influence," *J. Amer. Med. Inform. Assoc.*, vol. 21, pp. 212–218.
- [8] Choudhury, M. D., Gamon, M., Counts, S. and Horvitz, S., 2013, "Predicting depression via social media," presented at the *Int. Conf. Weblogs Social Media*, Cambridge, MA, USA.
- [9] Frommholz and Lechtenfeld, M., "Determining the Polarity of Postings for Discussion Search," in *LWA 2008-Workshop-Woche*.
- [10] Akay, A., Dragomir, A. and Erlandsson, B.E., "A Novel Data-Mining Approach Leveraging Social Media to Monitor Consumer Opinion of Sitagliptin," *J. of Biomedical and Health Informatics* in publication
- [11] Bremner, J. D., "Structural changes in the brain in depression and relationship to symptom recurrence," *CNS Spectrums*, vol. 7, no. 2, pp. 129–139, 2002.
- [12] Tsugawa, S., Kikuchi, Y., Kishino, Kosuke, F., Nakajima, K., Itoh, Y. and Ohsaki, H., "Recognizing Depression from Twitter Activity", 2015.
- [13] Nguyen, T., Phung, D., Dao, B., Venkatesh, S. and Berk, M. "Affective and Content Analysis of Online Depression Communities", IEEE, 2013.
- [14] Akay, A., Dragomir, A. and Erlandsson, B.E., "Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care", 2013.
- [15] Shokri, R. and Kocabas, I. "Big Data Analytics in Healthcare: Social & Biological Data", International Journal of Engineering Inventions, 2016.
- [16] Choudhury, M.D., "Role of Social Media in Tackling Challenges in Mental Health", 2013.

[17] Culjak, G. and Spranca, M., "Internet Improves Health Outcomes in Depression", International Conference on System Sciences-2006.

[18] Wu, J.L., Yu, L.C. and Chang, P.C., "Detecting causality from online psychiatric texts using inter-sentential language patterns", BMC Medical Informatics and Decision Making 2012.

[19] Newman, M.E.J., "Detecting community structure in networks", Department of Physics and Center for the Study of Complex Systems, University of Michigan.

[20] Mok, W.T., Sing, R., Jiang, X. and See, S.L., "Investigation of Social media on depression", 2014.

[21] Noessner, J., Niepert, M., Meilicke, C. and Stuckenschmidt, H., "Leveraging Terminological Structure for Object Reconciliation".

[22] Choudhury, M. D., Counts, S. and Horvitz, E., "Social media as a measurement tool of depression in populations", 2013.

[23] Rosenquist, J.N., Fowler, J.H. and Christakis, N.A., "Social network determinants of depression", 2010.

[24] Akay, A., Dragomir, A. and Erlandsson, B.J., "Assessing anti-depressants using intelligent data monitoring and mining of online fora", IEEE Journal of Biomedical and Health Informatics, 2016.