DEPRESSION DETECTION USING MACHINE AND DEEP LEARNING MODELS TO ASSESS MENTAL HEALTH OF SOCIAL MEDIA USERS

Smita Ghosh[1], Sneha Ghosh[2], Diptaraj Sen[2] and Pramita Das[3]

[1]Department of Mathematics and Computer Science,
Santa Clara University, California, USA
[2]Department of Computer Science and Engineering,
University of Engineering and Management, Kolkata, India
[3]Department of Electronics and Communication Engineering,
National Institute of Technology Durgapur, Durgapur, India

ABSTRACT

During the COVID-19 pandemic millions of people were affected due to quarantine and restrictions. With more than half of the world's population active on social media, people resorted to these platforms as their outlet for emotions. This led to researchers analysing content on social media to detect depression by studying the patterns of content posting. This paper focuses on finding a data-driven metric called 'Happiness Factor' of a user to assess their mental health. Various models were trained to classify a post as 'depressed'. A user's 'Happiness Factor' was calculated based on the nature of their posts. This metric identifies degrees of depression of a user. The results show the effectiveness of the classifier in identifying the depression level. Also, a Mental Health Awareness Resource System is proposed which recommends mental health awareness resources to users on their social media interface based on their 'Happiness Factor'.

KEYWORDS

Depression Detection, Machine Learning, Deep Learning, Universal Sentence Encoder, Social Media.

1. INTRODUCTION

With the increasing accessibility of the internet and the availability of affordable electronic goods, Online Social Networks (OSNs) have seen an inflation in the number of users. Various researchers have concluded that users tend to express themselves through tweets, posts, and the usage of hashtags, by means of which an overview of the users' thought process can be obtained [1]. When the world went into a state of lockdown due to the COVID-19 outbreak and most people got confined to a single place, people started using online social networks all the more as a way of staying connected by posting about their daily activities and sharing their feelings with their online friends [2, 3]. It is studied that the ways in which a person expresses themselves are an insight into their sentiments [4]. Concerns about mental health issues have been spiking for quite some time and more than often people are reluctant to seek help face to face which makes it harder for friends and family to detect any anomaly in mental state, whereas users on social media can talk about their mental health state with their online friends with a choice of not revealing their real identities and thus being devoid of any judgements but at the same time

David C. Wyld et al. (Eds): MLDS, IBCOM, CSIA, NLAI - 2022
pp. 01-12, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.122101

hoping to receive help [5]. According to the World Health Organisation, depression is a common illness worldwide that affects around 5% of adults and often remains undiagnosed. When recurrent and with moderate or severe intensity, depression may become a serious health condition and hamper the usual way of living.

The persistent problem has been a lack of resources, lack of trained health-care providers [6] and social stigma associated with mental disorders. In countries of various income levels, people who experience depression are often not correctly diagnosed. A substantial gap is observed between the care-givers and care-takers.

Motivated by the urgency of spreading mental health awareness combined with the potential of utilising social media as a mental health resource, this paper focuses on assessing a user's depression level by analysing the content of what they are posting and recommending helpful resources to each user on their social media wall. A data-driven metric titled 'Happiness Factor'(HF) is proposed which helps assign a 'degree' of depression to a user rather than just a binary classification of depression. A **M**ental Health **A**wareness **R**esource **S**ystem (MARS) is proposed that recommends and suggests resources to the users based on their HF. The approach in this paper has been divided into three phases. Figure 1 describes the pipeline of the proposed idea. The first phase focuses on building a reliable classifier that classifies a given post as 'depressed' and 'not depressed'. Multiple models are experimented with and the one with best accuracy is selected. The posts of each user are sent as input to the chosen classifier and generates the corresponding labels for each post. Phase 2 focuses on using the labelled posts of each user to calculate their HF. The HF calculated in Phase 2, is passed onto the MARS Recommender system that recommends mental health resources based on the value of HF. For different values of HF, different resources are suggested. Section 4 describes this process in detail.

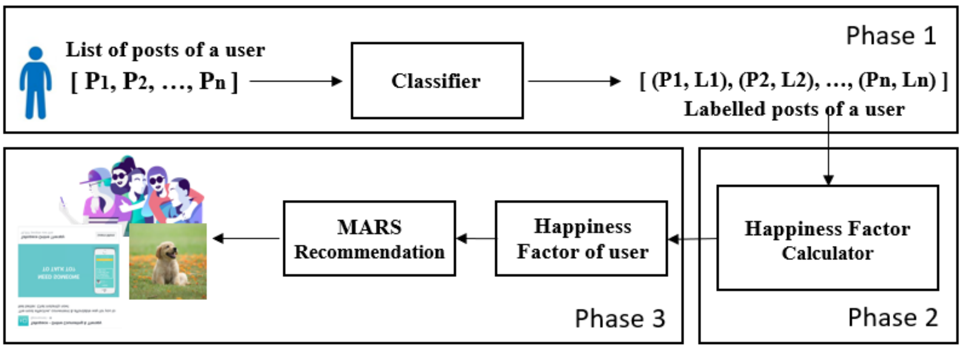

Figure 1. Proposed Approach Pipeline

Main contributions of this paper can be summarised as follows:

1) Building a robust classification model to categorise social posts as 'depressed' or 'not depressed'
2) Formulating a new data-driven metric called Happiness Factor to assess the user's level of depression on a scale of 0 to 1, instead of a binary value.
3) Propose the MARS (**M**ental Health **A**wareness **R**esource **S**ystem) Recommendation system that provides resources and help to those assessed with different levels of depression.

The paper can be divided into the following sections: Section 2 describes the related work and literature. Section 3 describes the proposed approach in this paper. Section 4 describes the proposed MARS recommendation system. Section 5 summarises the results. Section 6 concludes the paper and Section 7 provides description for future steps for the proposed idea.

2. RELATED WORK

In the past decade we have seen a steep rise in the number of people being diagnosed with depression. A rise in people using social media as an outlet for their emotions has also been observed [1,7]. Significant work has been done on detecting depression by accessing the posts made by different users on various online social network platforms. In [8], four machine learning techniques have been used to classify comments of facebook users with 60% - 80% accuracy, based on linguistic style, emotional process, temporal process as features.

The authors of [9] used Natural Language Processing algorithms in order to analyse social media data to enhance mental health care. The investigation was based on secondary analysis of a series of five focus groups with Twitter users, including three groups consisting of participants with a self-reported history of depression, and two groups consisting of participants without a self-reported history of depression. Even though there was scepticism regarding how well social media represents a users' mental health state, it was concluded that overall participants were interested in the opt-in utilisation of social media in the context of clinician-led mental healthcare.

In [10], four machine learning models- decision tree, random forest, support vector machine and naive bayes are first used, later a hybrid algorithm is developed that can recognize the emotions efficiently and utilise the Twitter dataset to measure the depression of people from their posts and other actions performed on online social networks.

[11] conducted a study to analyse whether machine learning (ML) methods could be effectively put to use to detect depression in people by analysing their social media texts, but without relying on specific keywords as people suffering from depression rarely leave such cues in their social media posts. It is of note that the approach presented in this paper is based on only supervised ML classifiers, thus limited to using labelled datasets for training the classifiers.

Existence of recommender systems can be traced back to 1979 in cognition science [12] and information retrieval [13], and the Usenet communication system created by Duke University [14]. In recent years, they have evolved into personalised recommendation systems [15,16], where user's of a platform are recommended resources based on their profile. Most common examples are 'You may also like' or 'Recommended for You' on Amazon, Netflix, etc. Inspired by the benefits of a personalised recommendation system, the MARS Recommendation system is proposed in this paper.

Motivated by the work done in [17] where the authors have analysed user posts to identify depression, this paper extends the problem of depression detection of a user to assess the 'degree' of depression of the user by analysing their post. This paper further extends the idea to provide helpful resources to the user based on the depression assessment.

3. PROPOSED APPROACH

In this paper, the proposed approach is divided into three phases. Phase 1 focuses on building a reliable and robust classifier that efficiently and accurately classifies a post as 'depressed' or 'not depressed'. Figure 2 describes the steps involved in this phase. The purpose of this phase is to be able to label the raw posts that are gathered for each user. Phase 2 focuses on calculating the data-driven HF of each user. Section 3.4 elaborates on the formulation of HF and Section 3.5 explains the interpretation of the HF value. Phase 3 focuses on the MARS recommender system which provides support and help to user's based on their HF value. The resources are displayed as 'links' or 'posts' on the user's social media platform.

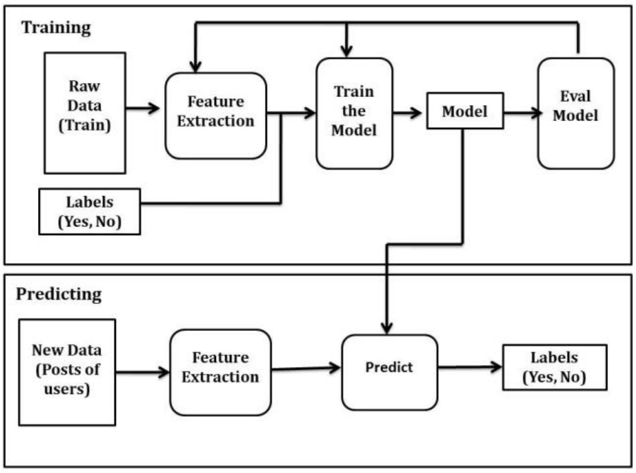

Figure 2. Classifier to detect 'Depressed' posts

3.1. Machine Learning and Deep Learning Models

Text classification and the detection of depression are very popular areas of research and as mentioned in the related work section, many papers describe the various Machine Learning and Deep Learning models that are used for depression detection. In this paper, Machine learning models such as Multinomial Naive Bayes(NB), Support Vector Machine(SVM), Decision Tree and K-Nearest Neighbours(KNN) and deep learning models such as Universal Sentence Encoder, Sequential Model and Long Short Term Memory(LSTM) are used. After evaluating their performance, the best one would be chosen to predict the label of the list of posts in order to calculate the Happiness Factor of each user. Figure 2 describes the training and testing process of the classifier.

3.2. Dataset Description for Phase 1

The dataset used for training the different classifiers was formed by combining posts from two other datasets. The first dataset (Dataset 1) [18] had 2314 posts labelled as 'depressed' and '8000' labels that were not depressed. The second dataset (Dataset 2) [19] had 494105 instances labelled as depressed and marked as '0' and 554470 tweets labelled as depressed and marked as '1'. A third dataset (Dataset 3) [20] of 291431 instances containing only tweets marked as depressed was found to make Dataset 1 balanced.The details of the datasets are in Table 1.

Table 1. Training Dataset Description

	#Depressed	#Not Depressed
Dataset 1	2314	8000
Dataset 2	494105	554470
Dataset 3	291431	0
Dataset T	8000	8000

3.2.1. Final Dataset for Training Classifiers

As most of the datasets obtained were either imbalanced or very large in size, a training dataset (Dataset T) was created by adding ‘depressed’ tweets from Dataset 3 into Dataset 1 to make it balanced. The final dataset has a total of 16,000 instances of which 8000 are of class 0 (depressed) and the rest 8000 are of class 1(not depressed).

3.2.2. User Post List Generation

The choice of the social media platform in this paper was a network similar to Facebook [21], as Facebook resonates more as a platform where user’s socialise and talk to friends in an informal format as compared to Twitter or Instagram.

Due to time constraints and Facebook’s data privacy policies, in this paper, a new approach was proposed for assigning the list of posts of each user on the network. The network dataset [21] selected to mimic a Facebook-like network had 899 nodes and 142760 edges where the nodes represented the users and the edges represented the relationship between them. In the proposed approach, as the focus is on posts of individual users, and not their connections, we ignored the edges.

A ‘posts’ dataset was created by combining 10,000 random posts from dataset 2 and 10,000 random posts from dataset 3. This dataset was then used to randomly assign posts to users. A random number between 0 to 20 was generated for each user in the network as the average number of posts a user makes on a social media platform. This number was based on a calculation following the statistics mentioned in [22] which states the total number of users on Facebook and total number of comments and statuses posted on Facebook daily. Figure 3 describes the process of the list generation for each user. This process helped to mimic the posts that the user would make on an average, during a given period of time.

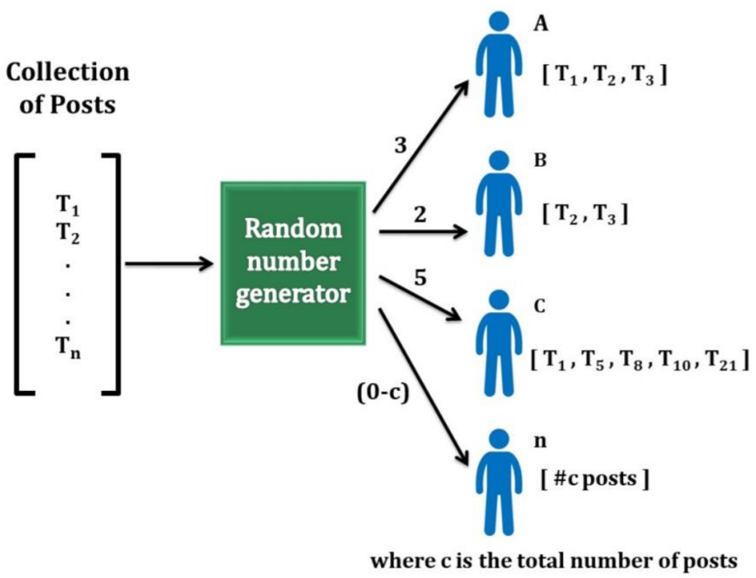

Figure 3. User Post List Generation

3.3. Happiness Factor (HF) Formulation

In this paper a new data-driven metric called 'Happiness Factor' (HF) was proposed to assess the level of depression of users. This would be represented on a scale of 0 to 1, with 0 indicating no depression and 1 indicating high levels of depression. This metric goes beyond a binary classification of depression, to depression levels between 0 and 1. The formulation of HF comprises of the following steps:

1) Using the best classifier built in section 3.1 to label the posts of each user.
2) Calculating the ratio of the number of depressed posts to the total number of posts the user makes during a given period of time.

HF(A) = |Depressed Posts| / |List of Posts|

where HF(A) denotes the Happiness factor of user A, **|Depressed Posts|** denotes the total number of tweets that were classified as depressed by the trained classifier and **|List of Posts|** refers to the total number posts that the user makes in a given time period.

3.4. Happiness Factor Interpretation

The HF value of a user would range between 0 to 1. A value of '0' indicates that the user made no depressed tweets in the given time frame. A value of 1 indicates that all the tweets of the user were classified as depressed. A value between 0.5 to 1 would indicate that more than half the posts of the user were classified as depressed and a value between 0 to 0.5 would indicate at most half the posts were classified as depressed. This scale allows HF to be interpreted as a level of depression of a user, instead of just a binary value of 'depressed' or 'not depressed'. Based on the values of the HF, the levels of depression are categorised as

1) **Level 1: (a HF value of 0 to 0.4) -** a user belonging to this level is considered depressed but on the lower scale of seriousness as at most 40% of their posts were classified as 'depressed'.
2) **Level 2: (a value of 0.5 to 0.7) -** a user belonging to this level is flagged as more depressed than Level 1 but not as serious as Level 3.
3) **Level 3: (a value of 0.8 to 1) -** a user belonging to this level is assessed as being extremely depressed and must be recommended resources to seek help in the form of therapy or other resources.

4. MARS RECOMMENDER

The MARS recommender stands for **M**ental Health **A**wareness **R**esource **S**ystem. Motivated by the benefits of personalised recommendation systems [15, 16], the MARS system proposed would take in as input the value of HF and provide resources based on the HF value. Section 3.4 describes the levels of depression interpreted for the HF value. These levels signify the 'degree' of depression and help identify the kind of help that can be provided to a user. One 'degree' of depression could be stress of not being able to cope with specific demands and events. If not managed properly, it could lead to chronic illness. A lot of research papers talk about stress management [23,24] for cases of extreme depression, but in the past years, with the advancement of YouTube, Instagram, Facebook etc, some papers describe watching funny videos [25,26], happy memories that release 'happy hormones' [27] becoming popular for managing anxiety and stress.

In this paper, we propose the 'social media methods' to help with depression. Figure 4 illustrates the working of the MARS Recommendation system. Depending on the various levels of the HF value, MARS would recommend the following:

1) **A HF value of 0 to 0.4**: This value indicates that at most half of the posts made by the user were depressed. This is considered as Level 1 of depression. To help these users, the MARS system would recommend motivational quotes, funny videos, meditation videos and other motivational articles which would appear on the user's home page. As they scroll through their homepage, the recommended resources would be the posts they see.
2) **A HF value of 0.5 to 0.7**: A user with this value of HF is considered depressed but at an intermediate level between Levels 1 and 3. For this user, the MARS would recommend in addition to motivational quotes, funny videos and other motivational articles, memories and pictures with friends and family.
3) **A HF value of 0.8 to 1**: A user is considered to be extremely depressed. For a user at Level 3, the MARS would recommend (in addition to resources from Level 1 and Level 2) resources for therapy, recommendations to contact friends and family, meditation resources, articles for dealing with stress, anxiety and depression.

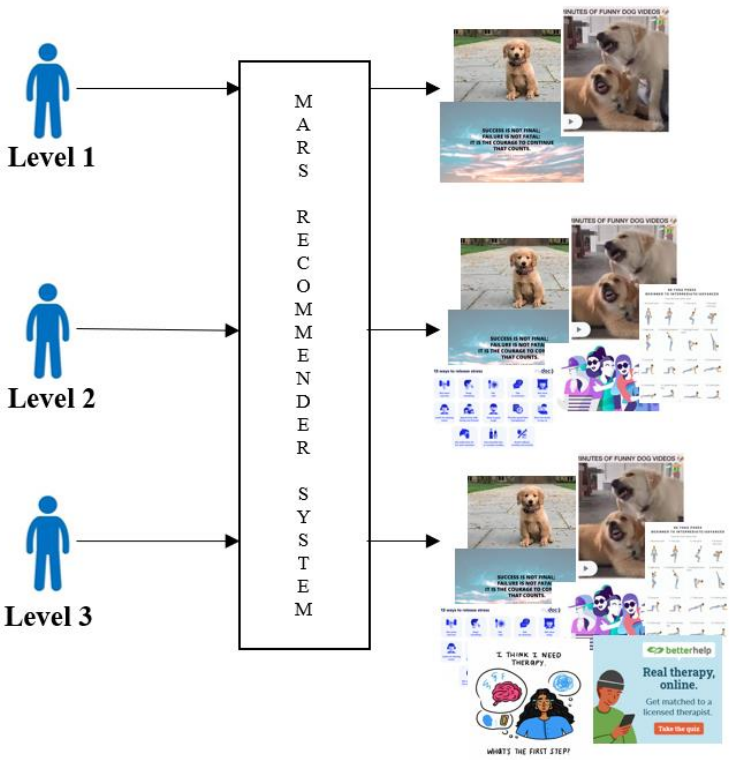

Figure 4. MARS Recommendation System

5. RESULTS

5.1. Building the Classifier

Phase 1 focuses on building the classifier. After extensive literature review on depression detection text classification models (as mentioned in the related work), some of the most popular machine and deep learning models were selected to experiment with as the 'Depression Detection' Classifier. They were Multinomial Naive Bayes, Support Vector Machines, Decision

Tree, K-Nearest Neighbours, Universal Sentence Encoder, LSTM and Sequential Model. Additionally it was observed that Universal Sentence Encoder (USE) was not used in works related to depression detection, thus USE was used as one of the prime models in this approach.

Experiments were conducted using Dataset 1 and Dataset T described in Table 1, Section 3.2 and Subsection 3.2.1. Table 2 describes the F1 score obtained for each algorithm.

Table 2. Performance of Different Models

Model	Dataset 1	Dataset T
Multinomial Naive Bayes	0.98	0.83
Support Vector Machine	0.98	0.96
Decision Tree (max_depth = 5)	0.99	0.91
K-Nearest Neighbours (K = 7)	0.96	0.67
Long Short-Term Memory(LSTM)	0.93	0.92
Universal Sentence Encoder(USE)	**0.99**	**0.98**
Sequential Model	0.94	0.87

For Dataset 1, it was observed that all the algorithms performed really well. A reason for this could be because the size of the dataset was very small.

For Dataset T, It was observed that K- Nearest Neighbours gave the lowest accuracy of 67% but all the other results were very similar to each other, with the USE performing best with an accuracy of 98%. One reason for that could be that the dataset used has balanced instances.

Overall, the Universal Sentence Encoder performed significantly better for both the datasets, so it was chosen as the classifier to predict the labels of the post list that was generated using the approach in subsection 3.2.2.

5.2. Calculating the Happiness Factor

A Facebook-like network [21] was used to represent the users in a network. For each user, a list of posts were generated. As described in subsection 3.2.2, due to restrictions of Facebook data policies, a new approach was proposed in this paper to assign a list of posts to each user. Once their list was generated, each list was passed through the USE classifier and labels of ‘depressed’ and ‘not depressed’ were predicted for each post. After the labels were predicted, the Happiness Factor (HF) for each user was calculated using the formula mentioned in Section 3.3. Figure 5 shows the HF for a subset of users on the network. The x-axis represents the USER ID and y-axis represents the HF values of those corresponding users.

5.3. Happiness Factor Value Analysis

As seen in Figure 5, different HF values are observed from as low as 0 to as high as 0.8. For the entire set of users, Table 3 shows the number of users in each level based on their HF values. Approximately 48% of the users were categorised into Level 1, approximately 44% of the users were categorised into Level 2 and the remaining were categorised into Level 3. Figure 6 represents the frequencies of each level in a chart.

Table 3. Number of Instances in Each Level of HF

Happiness Factor	Level	Number of Users
0 - 0.4	Level 1	433
0.5 - 0.7	Level 2	398
0.8 - 1	Level 3	75

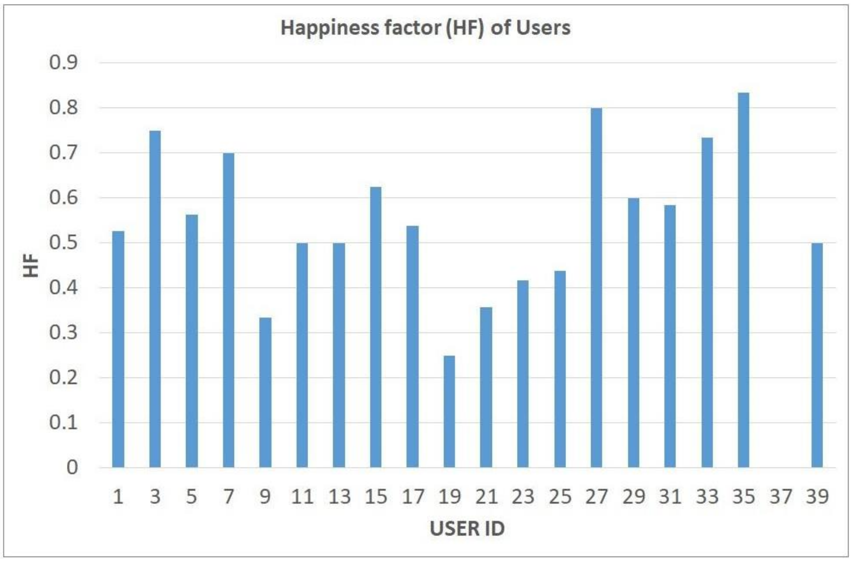

Figure 5. Happiness Factor of Subset of Users in the Network

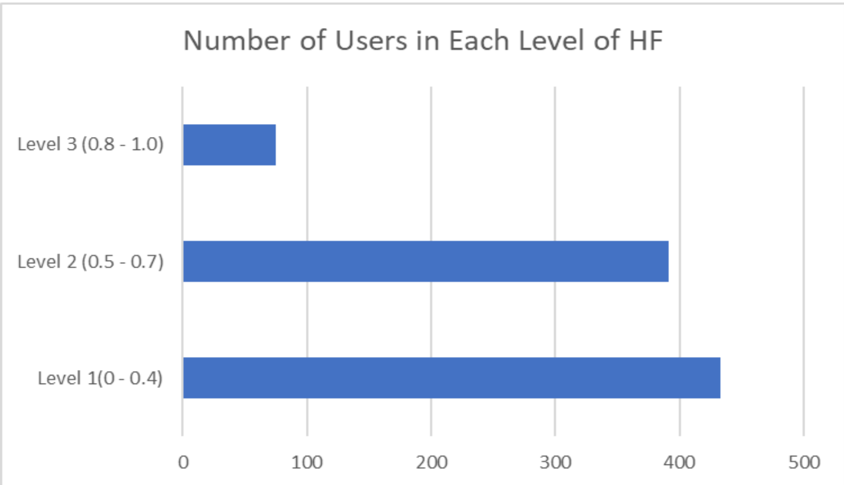

Figure 6. Number of Users in Each Level of HF (x-axis = #users, y-axis = Levels of HF)

5.4. MARS Recommendations

Once the HF are calculated for each user, this value will be passed on to the MARS Recommender system which would provide resources to the users based upon the conditions described in Section 4. These resources would appear as ‘posts’ on the user’s social media interface.

6. CONCLUSIONS

Mental health awareness has risen as a vital concern in recent times. Social media has been used as an outlet for many to express their emotions. But social media can also be used as a resource for mental health. In this paper, a new data-driven metric called 'Happiness factor' is proposed to assess the levels of depression of a social media user by analysing their posts. Classifiers were built with 98% accuracy for predicting posts as 'depressed'. These labelled posts were used to calculate every user's HF. It was observed that approximately 44% of users in the network used in the experimentation were categorised into Level 2 and approximately 8% were categorised into Level 3. Based on the interpretation of the HF described in section 4.4, necessary steps can be taken by the Online Social Network(OSN) provider to provide help and resources to those in need through the MARS Recommendation system. The proposed idea in this paper helps in identifying the users on the social network who are expressing themselves through their posts but may be hesitant to reach out for help. The Happiness factor and the MARS Recommendation System allows for the OSN provider to assist the users indirectly through posts and suggestions to help them cope with their depression.

7. FUTURE WORK

7.1. MARS Implementation

Implementation of the MARS System is the most important future work for the proposed idea. First approach would be to get user information from a social network using an API. The second approach would be to build a dummy social network using Object-Oriented Programming concepts to deploy the MARS recommendation system.

7.2. Acquiring Real User Data

In this paper, the posts of each user were assigned based on a proposed approach. In the future, real data describing user profiles and their posts is intended to be scraped and explored. Bright Data [28] is a paid service that allows users to scrape data from social media sites. From data collection infrastructure to ready-made datasets, Bright Data allows retrieving public available web data. Also Python has the Facebook-Scraper [29] module that helps in scraping facebook public pages without an API key.

7.3. Improving the Classifier

In this paper Deep Learning models namely the Universal Sentence Encoder, LSTM and Sequential Model and Machine Learning models namely Multinomial Naive Bayes Model, Decision Tree, K-Nearest Neighbour and Support Vector Machines (SVM) were trained. As text classification using more advanced models like Transformers are becoming popular, the future step for this work would be to explore more state-of-the-art models to build a more reliable and robust depression detection classifier. Also due to unavailability of more labelled datasets, only two datasets were used in this paper to train the classifiers. A future step would be to acquire more training datasets to evaluate the models better.

ACKNOWLEDGEMENTS

Special thanks are due to the department of Mathematics & Computer Science of the College of Arts & Sciences at Santa Clara University, as well as to the department of Computer Science Engineering at the University of Engineering & Management, Kolkata and the department of

Electronics and Communication Engineering at National Institute of Technology Durgapur, for all their continued support of the project.

REFERENCES

[1] E. M. Lachmar, A. K. Wittenborn, K. W. Bogen, and H. L. McCauley, "# mydepressionlook-slike: Examining public discourse about depression on twitter," JMIR Mental Health.

[2] J. Van Agteren, J. Bartholomaeus, D. B. Fassnacht, M. Iasiello, K. Ali, L. Lo, M. Kyrios et al.,"Using internet-based psychological measurement to capture the deteriorating communitymental health profile during covid-19: Observational study," JMIR mental health.

[3] A. G. da Silva, D. M. Miranda, A. P. Diaz, A. L. S. Teles, L. F. Malloy-Diniz, and A. P.Palha, "Mental health: why it still matters in the midst of a pandemic," Brazilian journal ofpsychiatry.

[4] N. Berry, F. Lobban, M. Belousov, R. Emsley, G. Nenadic, and S. Bucci, "# whywetweetmh:understanding why people use twitter to discuss mental health problems," Journal of medicalInternet research.

[5] M. Park, C. Cha, and M. Cha, "Depressive moods of users portrayed in twitter," 2012.

[6] Kendrick, Tony, "Depression in primary care: what more do we need to know?" The CanadianThe Canadian Journal of Psychiatry.

[7] De Choudhury, Munmun, Michael Gamon, Scott Counts, and Eric Horvitz, "Predicting depression via social media," Seventh international AAAI conference on weblogs and social media.

[8] Islam, Md and Kabir, Muhammad Ashad and Ahmed, Ashir and Kamal, Abu Raihan M and Wang, Hua and Ulhaq, Anwaar and others, "Depression detection from social network data using machine learning techniques," Health information science and systems.

[9] Mikal, Jude and Hurst, Samantha and Conway, Mike, "Investigating patient attitudes towards the use of social media data to augment depression diagnosis and treatment: a qualitative study," Proceedings of the fourth workshop on computational linguistics and clinical psychology—from linguistic signal to clinical reality.

[10] Smys, S and Raj, Jennifer S, "Analysis of deep learning techniques for early detection of depression on social media network-a comparative study," Journal of trends in Computer Science and Smart technology (TCSST).

[11] Chiong, Raymond and Budhi, Gregorius Satia and Dhakal, Sandeep and Chiong, Fabian, "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts," Computers in Biology and Medicine.

[12] E. Rich (1979): User modeling via stereotypes, Cognitive Science.

[13] M. Sanderson – W. B. Croft (2012): The History of Information Retrieval Research, Proceedings of the IEEE.

[14] P. Resnick – N. Iacovou – M. Sushak – P. Bergstrom – J. Riedl (1994): GroupLens: An open architecture for collaborative filtering of netnews, In Proceedings of the ACM Conf. Computer Support Cooperative Work (CSC).

[15] Hui, Bei, Lizong Zhang, Xue Zhou, Xiao Wen, and Yuhui Nian. "Personalized recommendation system based on knowledge embedding and historical behavior," Applied Intelligence.

[16] Sharma, Sunny, Vijay Rana, and Vivek Kumar. "Deep learning based semantic personalized recommendation system." International Journal of Information Management Data Insights.

[17] Al Asad, Nafiz, Md Appel Mahmud Pranto, Sadia Afreen, and Md Maynul Islam. "Depression detection by analyzing social media posts of user." In 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON).

[18] https://www.kaggle.com/datasets/krishbaisoya/tweets-sentiment-analysis

[19] https://www.kaggle.com/datasets/gargmanas/sentimental-analysis-for-tweets

[20] https://www.kaggle.com/datasets/vaibhavshahds/twitter-dataset

[21] T. Opsahl, "Triadic closure in two-mode networks: Redefining the global and local clustering coefficients," Social Networks.

[22] https://bernardmarr.com/how-much-data-do-we-create-every-day/

[23] Van den Bergh, Omer. Principles and practice of stress management. Guilford Publications, 2021.

[24] Sarkar, Sankalan, Bidyadhar Sa, Keerti Singh, Uma Gaur, Ambadasu Bharatha, Virginia Victor, Sayeeda Rahman, and Md Anwarul Azim Majumder. "Psychophysiological effects of yoga on stress

management among medical and allied health professional students during COVID-19 pandemic: A narrative review." Advances in Human Biology 11, no. 4 (2021): 3.
[25] Kelso, Lynn. "Decreasing Test Anxiety in Nursing Students Using Cute/Funny Animal Videos." (2021).
[26] Liu, Yang, Qin Gao, and Liang Ma. "Taking Micro-breaks at Work: Effects of Watching Funny Short-Form Videos on Subjective Experience, Physiological Stress, and Task Performance." In International Conference on Human-Computer Interaction, pp. 183-200. Springer, Cham, 2021.
[27] Jamal, Shehla. "Happy Hormones and Doctors."
[28] https://brightdata.com/products/datasets
[29] https://pypi.org/project/facebook-scraper/

AUTHORS

Dr. Smita Ghosh graduated with her PhD in Computer Science from The University of Texas at Dallas in May 2020. Prior to that she also has her Master's in Computer Science from the same university. Her research interests are Social Network Analysis (composed influence maximisation. information diffusion, hypergraphs, data analysis, machine learning in social networks analysis), Data Science (deep learning, data mining, trustworthy artificial intelligence, sentiment analysis, ethics in artificial intelligence).

Sneha Ghost' is pursuing her bachelor's degree in Computer Science and Engineering from the University of Engineering and Management, Kolkata, India. Her primary areas of interest which also constitutes her area of research are machine learning, natural language processing and social network analysis.

Diptaraj Sen is currently pursuing the bachelor's degree in Computer Science and Engineering with the University of Engineering & Management, Kolkata, India. He has been doing research on topics such as Association rule mining, imbalance classification and data-driven challenges. He has few conference publications in his name as an undergraduate researcher. His primary research interests arc machine learning. data mining, and social network analysis.

Pramita Das is an undergraduate student of National Institute of Technology Durgapur, India. pursuing bachelor's degree in Electronics and Communication Engineering. Having keen interest in AI. ML. DL and signal processing. digital communications and embedded systems, she has worked with Samsung RnD (India) and has jointly presented a poster in IEEE Rising Stars Conference 2022.