Topic: *Deep learning in social media*

I. Abstract:

Social media is a platform used by people for sharing their thoughts. In this society the mental health of a person is really important for surviving in this competitive world. After doing a deep and thorough research I found that using social media contents can be used to detect the mental state of an individual. This is done using some deep learning techniques where they constructed their own models for doing such a complicated task. This study is focused on to explain detail about how they implemented those techniques and how it helps an individual by detecting their mental state.

Keywords—social media, deep learning

II. Introduction:

Social media is a term which is used for calling different web applications which is also called as apps like *Instagram, Facebook, Linked In* etc., where people can interact with other peoples and share their thoughts for example in Instagram people share about their daily routine in their lives where as in Linked In, they share about their career.

The use of social media is that it allows people to be in touch with their families and friends. As I am an international student the social media allows me to keep in touch with my family from India in a regular basis that is day-to-day, which was not possible before few years. This shows the rapid growth of social media in recent years and lots have people have benefited from it.

But is it only useful for these specific tasks? Can we use social media for different purposes than we know? The answer is yes, there are a lot of different applications of social media but we only know a few of them.

In this paper I am going to explain about the different applications in social media by using deep learning techniques.

What is deep learning? Deep learning is a part of machine learning which is a combination of two or more layers which forms a neural network. There are different types of deep learning algorithms like, Convolution Neural Networks *(CNNs)*, Long Short-Term Memory Networks *(LSTMs)*, Recurrent Neural Networks *(RNNs)*, Generative Adversarial Networks *(GANs)* etc.,

The negative aspects of social media are inadequacy in people's life, fear of missing out, isolation, depression, anxiety etc., In this paper I am going to concentrate about how deep learning techniques can be useful in this and also how deep learning techniques used in social media for the use of people's life.

I am going to explain about some of the articles that has already illustrated about the deep learning techniques in the social media that was implemented for detecting mental illness, adapting methods for mental health prediction and characterisation of mental health conditions.

III. Literature Survey:

I have analysed the three different tasks that has been done in these papers where all the three methods used the content from the reddit posts from the people around the world. The different tasks are detecting mental illness from user content on social media, mental health prediction on social media and characterisation of mental health conditions in social media. The following would explain detail about the three different deep learning techniques that have used in their approach.

(A)Article 1- A deep learning model for detecting mental illness from user content on social media: [1]

Review:

This article's goal is to detect whether the user has mental illness such as anxiety and depression. For this they have collected data from reddit where they collected different types of mental health related data from social media. This is called as subreddits, there was a lot of subreddits but only six were identified as mental health related ones. The detailed overview of the subreddits is shown in this table. [1]

Channel	# of users	# of posts	Description
r/mentalhealth	27,177	39,373	The Mental Health subreddit is the central forum to discuss, vent, support and share information about mental health, illness and well-ness
r/depression	136,506	258,496	Peer support for anyone struggling with depression, the mental illness
r/Anxiety	49,735	86,243	Discussion and support for sufferers and loved ones of any anxiety disorder
r/bipolar	14,372	41,493	A safe haven for bipolar related issues. We are a community here not just a help page. Be a part of something that cares about who you are
r/BPD	13,913	38,216	A place for those who have BPD (Borderline Personality Disorder) (also known as EUPD [Emotionally Unstable Personality Disor-der])–, their family members and friends, and anyone else who is interested in learning more about the disorder
r/schizophrenia	5,392	17,506	Welcome! This is a community meant for a discussion of schizophrenia spectrum disorders, and related issues. Feel free to post, discuss, or just lurk. There is no judgement in this place: we are here for each other. Please refrain from self-diagnosis, diagnosing others, or advising specific medical treatments
r/autism	4,754	7,142	No description

Table 1. Summary of the data from reddit

After obtaining the data they pre-processed the data using two techniques natural language toolkit (NLTK) and the other technique is Porter Stemmer. The NLTK is used to filter out the frequently used words by the users. Porter Stemmer is used to change those words into its root meaning and decrease the repetition of those words. A more detailed view of their pre-process can be seen in this picture.[1]

Figure 1. Data pre-processing procedure

For classification they have used six different models because there was a total of six different subreddits. All these models were independent binary classification models, the reason to develop six different independent binary classification models is to be more accurate in the results for example if a user suffers a multiple mental health issue, then for each separate issue different models predict the results. The data was then labelled for their respective class ad also their opposite class like anxiety and non-anxiety class. They also used synthetic minority over-sampling technique (SMOTE) algorithm which is used overcome the class-imbalance issue.

After splitting the datasets into test and train XGBoost and convolution neural network (CNN) was done. In both these models the words need to be changed in the XGBoost they used TF-IDF vectorizer which is used to convert words into n-dimensional vectors and in CNN in the pre-processed text data they applied word-embedding procedure. Note that this was done using word2vec and it was trained in the training dataset.

The architecture of the CNN is represented in the following image [1]

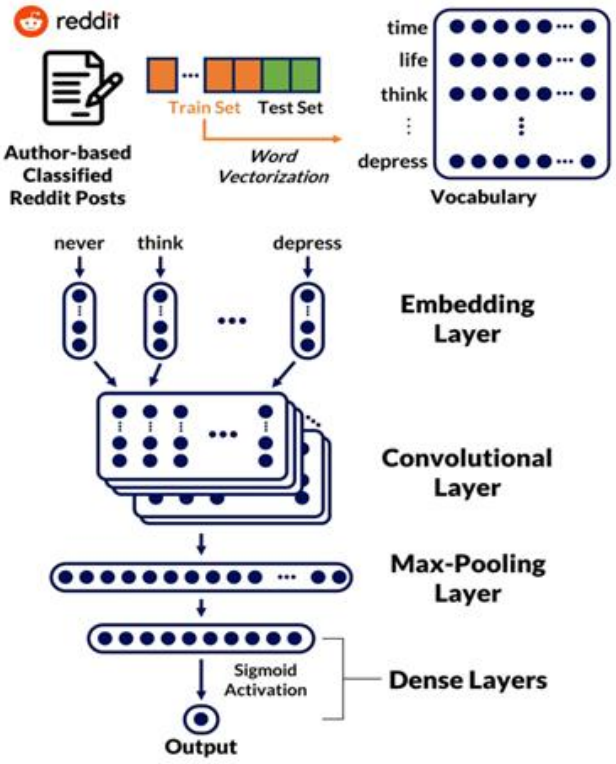

Figure 2. Architecture of the CNN-based classification model

Embedded layer: - pre-processed post with 20 dimensions and weights are pretrained by word2vec

Convolution layer: - 128 filters of size 5

Dropout layer: - 0.25 dropout rate

Max-Pooling layer: - maximum values within CNN filters with dimension of 128

Loss function: - binary cross-entropy

Optimizer: - Adam

Learning rate: - 0.001

Epochs: - 50

Batch size: - 64

Their model was able to detect users who might have psychological disorders. But their model could have a better performance if they have more data. The performance overview is represented in the below diagram. [1]

		XGBoost		CNN			
Channel	**Class**	**F1-Score**	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Accuracy (%)**
r/depression	Depression	58.02	71.69	89.10	71.75	79.49	75.13
	Non-depression	78.65		58.66	82.04	68.41	
r/Anxiety	Anxiety	55.92	70.41	87.54	41.44	56.25	77.81
	Non-anxiety	77.73		75.92	96.91	85.14	
r/bipolar	Bipolar	53.59	85.53	87.22	38.02	52.95	90.20
	Non-bipolar	91.43		90.40	99.05	94.53	
r/BPD	BPD	46.43	85.14	91.84	32.69	48.21	90.49
	Non-BPD	91.37		90.42	99.54	94.76	
r/schizophrenia	Schizophrenia	40.97	86.72	81.16	24.87	38.07	94.33
	Non-schizophrenia	92.52		94.62	99.56	97.03	
r/autism	Autism	38.31	94.91	48.08	49.39	48.73	96.96
	Non-autism	97.35		98.48	98.40	98.44	

Table 2. Model evaluation of XGBoost and Convolution Neural Network

In conclusion the article states that “Detecting mental illness problems in early stages and providing appropriate solutions can help potential mental disorder sufferers” [1]. But there were some drawbacks to this method where it did not involve “socio-demographic” and “regional differences” [1] this could affect the classification models according to the article.

In future using a recurrent neural network a “time series user level analysis” [1] can be implemented in order to develop a user-level detection model.

(B) Article 2- Adapting Deep Learning methods for mental health predictions on social media: [2]

Review:

The aim of this article is to predict if a user suffers from any one of the nine disorders such as depression, bipolar, autism etc. The dataset used in this article is the Self-reported Mental Health Diagnoses (SMHD) dataset. This dataset is based on the reddit posts of the user (diagnosed user) that is if the user posts something related to the disorders, then it is recorded in this dataset. After classifying according to the disorders, the dataset is classified as represented in the table shown below. [2]

	# users	# posts per user
Depression	14,139	162.2 (84.2)
ADHD	10,098	164.7 (83.6)
Anxiety	8,783	159.7 (83.0)
Bipolar	6,434	157.6 (82.4)
PTSD	2,894	160.7 (84.7)
Autism	2,911	168.3 (84.5)
OCD	2,336	158.8 (81.4)
Schizophrenia	1,331	157.3 (80.5)
Eating	598	161.4 (81.0)

Table 3. Number of users in SMHD dataset per condition and the average number of posts per user

To compare the texts of general discussion to the texts of the diagnosed user, the texts of the diagnosed user is normalized by removing the posts they have posted about the mental health status. Each disorder has their own respective control groups and for every disorder they are analysed the texts between the diagnosed user and their respective control groups. To choose the control group they were doing a selection process multiple time.

The classifier they have chosen is Hierarchical Attention Network (HAN) this is originally used for document classification. The HAN architecture has a word sequence encoder, a word-level attention layer, a sentence encoder and a sentence-level attention layer. In this article they replaced the document part of the HAN to social media user and the posts by the users are considered as sentences.

The architecture of the HAN is as follows, two layers of bidirectional GRU units with 150 as hidden size and after the first bidirectional layer there is an attention mechanism with a dimension of 100 and also the same for the second bidirectional layer. The first is to encode posts and the second is to encode a user as a sequence of encoded post. An output layer with a dimension of 50 and the input layer with Glo Ve word embeddings with a dimension of 300.

Loss function: - binary cross-entropy

Optimizer: - Adam

Learning rate: - 10^{-4}

Epochs: - 50

Batch size: - 32

They also performed other models like Logistic regression, Linear SVM, Supervised FastText and compared the results with HAN model. The comparison can be seen in the following diagram. [2]

	Depression	ADHD	Anxiety	Bipolar	PTSD	Autism	OCD	Schizo	Eating
Logistic Regression	59.00	51.02	62.34	61.87	69.34	**55.57**	**59.49**	56.31	70.71
Linear SVM	58.64	50.08	61.69	61.30	**69.91**	55.35	58.56	**57.43**	**70.91**
Supervised FastText	58.38	48.80	60.17	56.53	61.08	49.52	54.16	46.73	63.73
HAN	**68.28**	**64.27**	**69.24**	**67.42**	68.59	53.09	58.51	53.68	63.94

Table 4. F_1 measure averaged over five runs with different control groups

The table has the F_1 measures according to the disorder. From the table the HAN model beats the other models in the first four disorders. This is because they had a sufficient data in the respective groups. The F_1 scores of the HAN model can be seen in the below diagram. [2]

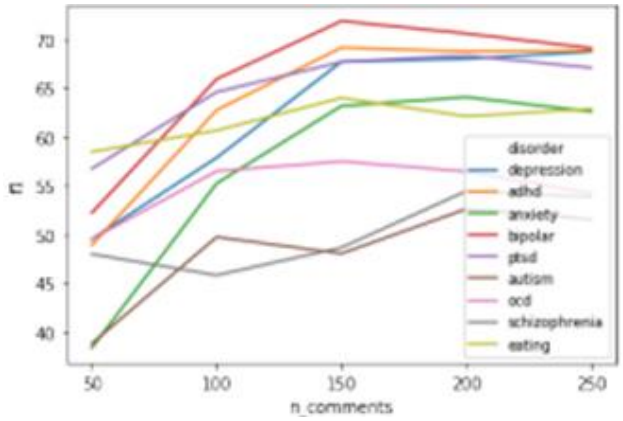

Figure 3. F_1 scores for different number of posts per users mad available to HAN, averaged over three runs for different control groups

In summary this article has some limitations that is the model needs more data to give its best performance than the standard models. For overcoming the limitation, they are going to use a transformer-based model in order to improve the performance accuracy.

IV. Discussion:

In both the articles they were using deep learning techniques to detect or predict the mental state of a person using the user data in the social media.

The first article compared the models XGBoost and CNN. The deep learning model performed better in all the six cases. This might be because they segregated the data as mental disorders and non-mental disorders (ex: bipolar and non-bipolar) which was beneficial for them. So, the model was trained based on a specific mental state because of the data handling in order to classify the disorder and predict the probabilities for that disorder. [1]

The second article used HAN model to predict whether the user has a mental disorder according to the nine groups they had. They compared the F_1 scores with some other models in that the HAN model was successful for only four out of the nine groups. In my opinion if the data was handled like in the first model, they might have given better results than they got now. [2]

V. Conclusion:

Mental state is very important to every person and with deep learning we can predict whether a person has a mental disorder or not by the social media usage of a person. In today's world people who are facing these types of situations are having a difficulty in their lives. So, by these techniques predicting the mental disorder of a person in early stages might help a lot of people. Some of the techniques needs to be improvised for a better result and hopefully they work better. In the future we can expect that the deep learning techniques can be more accurate in these tasks.

Around the globe lots of researchers are focusing on achieving the best model to predict the mental disorder of a person and hopefully those models can save lives of the people.

VI. References:

1. Kim, J., Lee, J., Park, E. *et al.* "A deep learning model for detecting mental illness from user content on social media". *Sci Rep* 10, 11846 (2020).

2. Sekulić, Ivan, and Michael Strube. "Adapting deep learning methods for mental health prediction on social media." *arXiv preprint arXiv:2003.07634* (2020).