A Proposed Method Using Deep Learning from Unseen to Seen Anxieties of Children during COVID-19

[1]Prof. Samir Kumar Bandyopadhyay, [2]Prof Vishal Goyal and [3]Shawni Dutta

[1]GLA University, Mathura, India

[1]GLA University, Mathura, India

[3]Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India.

Address of the Corresponding Author: Prof. Samir Kumar Bandyopadhyay, GLA University, Mathura, India, Email: 1954samir@gmail.com

Abstract

COVID-19 and new concept, lockdown, change social life of all classes of humans. Children partially feel the changes of daily life and this situation has been children's free mind. Children are under a new type of restriction imposed on them by their parents. Normally they prefer play with their their friends than study and always waiting for holidays. They heard a new jargon i.e. lockdown where everything stands still. Very often they see peoples in the roads and few vehicles are moving in the roads. However, a peculiar thing happens now that they sit in front of computer to hear the virtual classes that are taken by the teachers. This also happens when there is no lockdown since COVID-19 still affects people. The environment is totally changed and they do not find any proper answers from the parents about the scenario. This study has been made an attempt to carry out the mental affairs of children in West Bengal, India. Several families are surveyed for collecting responses mostly from rural areas as well as urban areas for the time-period from April, 2020 to July, 2020. An effort has been given in this paper to predict the stress, depression and anxiety faced by children during the COVID-19. A Deep Learning Neural Network (DLNN) based method is applied to understand the stress level, depression level and anxiety level amongst the children. A hybrid DLNN has been presented in this research that combines both Convolutional Layer and Gated-Recurrent Unit (GRU) for obtaining the prediction of the mental health of children. The model obtains an accuracy of 89.57% for defeminizing mental anxiety of children.

Keywords: COVID-19; lockdown; CNN; DLNN; GRU; mental anxiety; hybrid approach

Introduction

All over world the mental health problems of children exhibits same scenario in most of the families. It is necessary to take care of the problem in the early life of children during COVID-19 so that it has no effect in their latter life. If it persists for a long time, it creates a variety of problems. This research does not consider the situation. It is true that the mental health of children and adolescents are neglected in lower-income and middle-income countries [1]. It was found that the Mental health problems in 2010 is the leading causes of duration of life of a human and disability of human all over the world due to depression and anxiety disorders [2].

During the COVID-19 crisis, both positive effect and negative effect happen over children and also on their families. It changes normal daily routines as well as environments since they are not freely met with their friends in the localities [3]. Positive effects of this situation are minimizing risk of health for elder persons as well as children. All family members now habituated with restricted facilities and with limited resources that in turn minimize their monthly expenses. Children are now provided on-line   schooling system at home. Guardians supervise their studies and they take care of their children efficiently than the care has been taken in the school. It is true since a teacher has to take cares a number of children in the school.

However, this scenario slightly differs when question of working parents came into play. Some parents are expected to put their efforts as much as possible with extended working hours at home. Child daycare routine changes for working parents. Assistance from grandparents and other family members can not be the same care from parents in most of the cases. These children get less care from parents than non working parents. The care by school is then important. It is not possible now since schools are closed. The restrictions of closures of schools, social and out-of-home activities for children and adolescents have put their lives in perplexing circumstances. They have no longer relish positive interactions with their sport coaches, music teachers, friends and peers. This will obviously increase their stress, anxiety, depression level in their daily lives. So it is an urgent need to treat basic mental health problems of children before the problem will become a complicated situation. It is important as mental health of children often have a negative effect on everyday functions. These have a long lasting effects and may be an instance of higher risk of impairment.

Consider a typical situation. It is normally happened in India. Suppose a teacher is taking a online class and children is sincerely listen his/her lesson. Suddenly either power is off or

internet is not working for some time so children have a set back since he/she misses the lesson delivered by teacher. It creates anxiety in the children. Same things are also happened if teacher gave a task and children did not submit it on time due to the lockdown of internet for some time. Another example is for children whose parents are working people. If children are not cared by anyone in the family for some time then his/her minds may want to play with some playing kids. So children missed lesson of teacher which in turn may have an effect on weekly or monthly examination. If the result is not good, then parents will tell something unusual words to children and it may create depression to the children. Suppose in a family parents have two children of different ages. These children can not play with each other since their likings to play are not same. So both feel them as individual.

Machine Learning (ML) based techniques may be useful for learning and utilizing the patterns discovered from large database. It applies on a set of information in order to recognize underlying relationship patterns from the information set [4]. It is basically a learning stage. It can be tested with unknown incoming set of patterns. DL can process information with minimal processing time due to its self-adaptive structure. It is an expansion over conventional artificial neural networks since it facilitates the construction of networks by incorporating more than two layers [5]. A predictive model is proposed in this study that follows DL techniques for analyzing and assessing mental anxiety perceived amongst children. Such analysis may help to identify the suffering children from mental anxiety in advance. If children's mental health status is known in advance, parents or guardians can look into the matter to support them to get rid of such situation.

A hybrid model is proposed in this paper based on two dissimilar components. The first one is the convolutional layer of Convolutional Neural Network (CNN) [6] and the other one is GRU [7], a variant of Recurrent Neural Network (RNN) [8]. This model considers present life status of children and assesses the mental anxiety to be observed due to COVID-19 situation. The implemented model recommends mental anxiety of children after analyzing the the influential factors such as siblings present, playing time, time for online education, parental economic condition, reasoning capacity and many more. This model has been applied for assessing mental health status of children belonging to West Bengal, India. The data collection covers the children having age groups from 4 to 12.

Assessing mental health disorders is quite an interesting field and many researches have been carried out in this field. In this study, an approach to mental health anxiety prediction using

DL method is investigated. This study basically focuses on children mental health due to COVID-19. To best of our knowledge, the establishing relationship between the mental health of children due to COVID-19 and mental health of children are not studied yet. Hence, this research can give an indication for the future to analyze the current scenario which will be helpful to adjust the time to be spent by parents for primary care for children. It will also help to prepare a customized routine for the children.

The outlines of this research can be concisely presented as follows-

1.   Study the effects of mental health of children that deviate the concentration of children.

2.   The collected data from families who are residing in West Bengal, India. Collected data are based on questionnaire that covers several interfering aspects such as online education, economic condition, playing atmosphere, recent behavior, working parents, numbers of children and so on.

3.   Data are analyzed using DL technique for analyzing the status of children mental health. The learning outcomes obtained from this technique are utilized for prediction of mental anxiety amongst children.

4.   The proposed DL technique is exemplified by building a hybrid model based on the variant of RNN and CNN components. Convolutional layer of CNN and GRU-RNN is assembled for designing the predictive model.

Related Works

Mental health problem of children was detected by developing prediction models using multi-level logistic regression analysis. The main objective of this research is to assess the one-year risk of a first child mental health problem that can be used in primary care in clinical practice [9]. The comparisons have been made by implementing various machine learning algorithms such as support vector machines, decision trees, naïve bayes classifier, K-nearest neighbor classifier and logistic regression to identify state of mental health [10]. High school students, college students and working professionals were also considered as a target groups for assessing their mental health states [10]. For predicting the likelihood of developing psychological conditions such as anxiety, behavioural disorders, depression and post-traumatic stress disorders, neural network based technique has been exemplified with an average accuracy of 82.35% [11]. A combined approaches involving both depth first search and backward search strategy were implemented for diagnosing depression or dementia.

An expert system was developed by incorporating factors such as patient's behavioural, cognitive, emotional symptoms and results of neuropsychological assessments [12]. Screening of mental anxiety and depression amongst sea fearer of chidren are detected using ML algorithms such as CatBoost, Logistic Regression, Naïve Bayes, Random Forest, and Support Vector Machine. Comparative study confirmed in the research with an accuracy of 82.6% as the best possible result exhibited by CatBoost [13].

Schizophrenia is another type of mental health disorder which is diagnosed by implementing a decision support system with an accuracy of 66-82% [14]. Another study also focused to handle Schizophrenia by implementing integrated structured methodologies in decision support such as Multi-Criteria Decision Analysis (MCDA) and structured representation of knowledge into production rules and probabilities [15].

The literature reviews show the existing research works for assessing different mental health status on different target population with the help of machine learning based methods. In our study, the effect of COVID-19 has been taken into consideration for mental health disorder analysis. The impact of current situations on children mental health are assessed using DL framework. In this case, the target populations are children for the age groups between 4 and 12. Early diagnosis of this disorder can help in providing an improved quality of lives, especially under the current scenario.

Backgrounds

3.1 Neural Network and Deep Learning

ML is super field of DL and sub field of Artificial Intelligence (AI). ML automates and develops the machines and it is the actual goal of AI. DL provides a multi-layered hierarchical data representation typically in the form of a neural network when the concepts of multi-layer have been incorporated in the model. Since it is self-adaptive in nature so it is not required to perform the manual feature engineering task. It is used to solve complex problem. Like neurons in human brain a large number of processing elements (nodes) are present in the proposed neural network model for acquiring best problem solving technique [5].

3.2. Recurrent Neural Network and GRU

The multiple neural networks are used in RNN. It is used for analysing sequential data. Here the output of previous step is fed into input of current step. The output obtained in step S_i affects the parameters of step S_{i+1}. So RNN accepts two types of input- one is present input and other one is the previous output for obtaining the final output. The signals are travelled both forward and backward if loops present in RNN. It suffers from the vanishing gradient problem [8]. Variants of RNN such as Gated Recurrent Units (GRU) is explored to resolve the problem.. This neural networks control the flow of information through the sequence chain. GRU introduces the concept of gates for short-term memory.

GRU networks perform well when it is learning long-term dependencies on data. GRU consists of two types of gates such as update gate and reset gate. It overcomes the problems of vanishing and explosion of gradients in sequence learning tasks. Update gate decides addition and elimination of information. The use of reset gate identifies how much information are to hold from the past. GRU uses update gate and reset gate for solving the vanishing gradient problem of a standard RNN. These vectors decide what information should be passed to the output. Without washing it through time or remove information, vectors can be trained to keep information from previous data. It is not dependent on the prediction of future data [7].

Given $x_t = (x_1, \ldots, x_T)$ as an input sequence, W is the weight matrices and σ states the sigmoid function for a GRU. At the time t, the activation function of GRU is $h_j{}^t$ which depends on previous activation $h_{t-1}{}^j$ and candidate activation function $h`_t{}^j$. This is given in equation (1). The update gate ($u_t{}^j$), and reset gate ($r_t{}^j$) can be formulated as in equation (2) and equation (3) respectively.

$$h_j^t = \left(1 - u_t^j\right)h_{t-1}^j + u_t^j\, h'^j_t \qquad\qquad\qquad (1)$$

$$u_t^j = \sigma\left(W_u \cdot [h_{t-1}^j, x_i]\right) \qquad\qquad\qquad (2)$$

$$r_t^j = \sigma(W_r \cdot [h_{t-1}^j, x_i]) \qquad\qquad\qquad (3)$$

3.3. Convolutional Layer

Convolutional Neural Networks (CNNs) are also an improved version over traditional neural network. CNN can extract underlying hierarchical features by discovering the local relationship between nodes. Convolution operation is exhibited by each neighbour node in

order to capture inherent relationship in adjacent nodes. Convolutional Layers are one of the components of CNN. The calculation of the scalar product between their weights and the region connected to the input volume will produce the output of neurons that are connected to local regions of the input. The layers parameters focus around the use of learnable kernels [6]. In other words, an input data and a convolution kernel are subjected to particular mathematical operation to generate a transformed feature map. Convolution is often interpreted as a filter, where the kernel filters are the feature map for information of a certain kind. The convolutional layer performs an operation called a "convolution". It is a linear operation that involves the multiplication of a set of weights with the input. An array of input data and a two-dimensional array of weights, called a filter or a kernel, are multiplied for obtaining results. ReLu activation function is popularly used in Convolutional layer and is efficient in most situations. Applications of non-linearity are applied after convolution to assist for successful simulation [16].

3.4.Hyper-parameters used in neural network training

Some pre-stage fine-tuning of hyper-parameters is necessary to perform before training this neural network. It contains number of layers, number of nodes, learning rate, epoch size, batch size and drop-out rate. Adjusting hyper-parameters before providing the training, the neural network model assists in obtaining the maximized performance. Hyper-parameters that need to be taken into consideration includes number of layers, number of nodes for each layer, learning rate, epoch size, batch size and drop-out rate.

These values should be adjusted to help the network to learn successfully. Activation function is one of the necessary tasks which can maximize the training procedure. These functions allow neural networks to learn non-linear relationship among data and to produce meaningful output signal. Sigmoid activation function may be used for activating output nodes for predicting binary class probabilities. This activation function accepts the input data and transforms it in the range of 0 to 1 and it is shown in equation (4) [16]. Tangent hyperbolic (tanh) is also non-linear activation function and it is a smoother and zero-centered function [16]. The function range in between -1 to 1 and the output of the Tanh function are given in equation (5).

$$f(x) = 1/(1 + \exp^{-x}) \tag{4}$$

$$f(x) = (e^x - e^{-x})/(e^x + e^{-x}) \tag{5}$$

Hyper-parameters such as epoch and batch size are also used in neural network training. These hyper-parameters receive integer values which need to be chosen wisely to make the best use of the model's performance. The size of the epoch is defined to be number of passes to complete through training dataset. The dataset is passed to and fro through the neural network exactly one time within each epoch. During passing the entire dataset into the algorithm, it must be partitioned into fixed size of batches. The size of the batch keeps track of number of processed instances before the model updates its internal parameters. It needs to be ensured that batch size should not be too small or too large. If the size of the batch is too small then it will present high variance. It means that it does not represent the entire dataset. On the other hand, large batch size may not fit in memory to compute samples used for training and may lead to over-fitting problem [17]. Over fitting is a problem that appears when the model understands the details as well as noise presents in the data during training process to the extent that impacts negatively for new unknown patterns of data. In other words, over-fitting exhibits higher accuracy for the training data and minimised accuracy for testing data. In order to eliminate such problems, use of drop-out technique is used. This technique ensures the deactivation of nodes along with their incoming and outgoing connections present in the neural network. Incorporation of this technique into neural network based architecture often ensures the acquaintance of benchmark result as in supervised classification task [18].

The use of optimizer is mandatory in order to stack multiple neural layers under a single framework. Adam is computationally efficient with optimised memory requirement and is also easy to implement. The proposed method can be applied to optimize for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. It is quite well accepted due to its applicability on non-stationary objectives and problems with very noisy and/or sparse gradients [19].

Dataset Used

The study collects data from 736 families in West Bengal, India for four months from April 2020 to July 2020. Total 263, 157, 124, 192 family details are collected during the month of April, May, June, and July respectively. Table 1 to table 4 show distribution of families participating in the survey. Each of these families is provided with survey form and they entered information in the form for analyzing the task for prediction of mental anxieties of

children during COVID-19. The data collection mainly emphasizes on the children age groups from 4 to 12. Table 5 gives summarization of retrieved dataset in terms of attributes.

April	Cities	Urban.Areas	Rural.Areas	
5	10	5	4	
12	12	18	20	
18	30	10	15	
23	25	20	30	
28	11	17	19	
30	4	14	9	
	82	84	97	263

Table 1: Day-wise distribution for the participating families during April, 2020

May	Cities	Urban-Areas	Rural-Areas	
10	6	10	8	
15	10	5	5	
20	9	4	7	
27	18	10	7	
31	11	15	6	
31	7	14	11	
	55	58	44	157

Table 2: Day-wise distribution for the participating families during May, 2020

June	Cities	Urban-Areas	Rural-Areas	
15	6	10	8	
10	10	5	5	
25	9	4	7	
27	18	10	7	
31	37	29	27	124

Table 3: Day-wise distribution for the participating families during June, 2020

July	Cities	Urban/Areas	Rural/Areas	
15	6	10	8	
10	10	5	5	
25	9	4	7	
27	18	10	7	
31	37	29	27	
	80	58	54	192

Table 4: Day-wise distribution for the participating families during July, 2020

Attributes	Detail of Attributes	Values Present
Single Child	Whether the child has siblings or not	0-No, 1-Yes

Age	Age of the child (in years)	4-12
Online_Education	Child is facilitated with online education or not	0-No, 1-Yes
Think	Whether the child takes more time to think or not	0-No, 1-Yes
Online_hours	Amount of hours spend during online education (in hours)	3-7
Panic_Online	Measures the level of child panic during online class disruption	0-5
Covid-19	Whether any of family members has suffered from Covid-19 or not	0-No, 1-Yes, 2-Prefer not to answer
Economy	Current economic condition of family	'Lost-Jobs', 'Closed Business', 'Work from home', 'Normal'
Loneliness	Whether the child is lonely at home or not	0-No, 1-Yes
Play	Whether the child plays for enough time or not.	0-No, 1-Yes
Behaviour	Perceived behaviour of child	'short-tempered', 'quite', 'annoyed', 'normal'
Sleep	Sleeping style of children	'normal', 'irregular sleep'
Stress	Whether the child is over-stressed or not	0-No, 1-Yes
Mental Anxiety	Whether the child faces mental anxiety or not	0-No, 1-Yes
Depression	Whether the child is depressed or not	0-No, 1-Yes

Table 5: Dataset summary

Once the dataset is collected, pre-processing techniques are applied to make cleaned dataset. All the categorical attributes present in the dataset are encoded into numeric data. This will be followed by scaling values of every feature with large set of data points. Feature scaling will assist the classifier to work using normalized data with an enhanced efficiency. Large set of data points are scaled down within the range of 0 to 1 using feature scaling operation. Once this feature scaling operation is performed, feature vector is fitted to classifier model as training purpose. The pre-processed dataset is bi-furcated into training and testing dataset with the ratio of 7:3. The training and testing dataset is mainly distinguished by the presence of dependent variable. The target variables are kept in training dataset whereas these are eliminated from the testing dataset. The classifier learns by extracting patterns from the training dataset during training phase. Later, prediction is made on the testing dataset.

Methodology

Classification procedure is applied in this framework on the collected dataset in order to obtain mental health status amongst children in advance. The proposed methodology uses deep neural network as classification strategy. The target of this research is to design a system that will accept attributes present in the dataset and utilise the underlying relationship among those features in order to fulfil the objective of this study. The system is implemented by designing hybrid neural network model that assembles Convolutional and RNN layers under a single platform. While designing this model it is necessary to fine-tune hyper-parameters so that maximised performance can be attained. The next section describes specification of the model along with its hyper-parameters.

The deep model comprises of two 1-dimensional Convolutional layers with filter size of 256 and 128 respectively. Each of these layers is adjusted with kernel size of 1. Next, two GRU layers are stacked into this model with 64 and 32 nodes respectively. All these four layers are receives dropout layers having dropout rate of 0.2. Finally, four fully-connected dense layers are amalgamated into the deep model with 8,4,2,1 numbers of nodes respectively. Finally these aforementioned layers are compiled using adam solver by means of 30 epochs and with a batch size of 64. Layers in this model are activated by sigmoid, relu or tanh functions. Fine-tuning of these hyper-parameters is mandatory since it assists in achieving the best predictive result. The deep neural network receives a total of 90,089 parameters and trains those parameters for achieving prediction. Components of the model in terms of layers, shape of output data from each layers, and number of parameters received in each layers are defined in Table 6. The employed hyper-parameters while designing the proposed deep model are summarised in Table 7. The experiment has been carried out in Windows 10 Home with Intel Core i5-9300H (9th Gen), 8GB memory, and an NVIDIA GeForce GTX 1650 GPU.

Layer Type	Number of Nodes/ filter size/ Rate	Output Shape	Number of received parameters	Activation function Used
conv1d_1 (Conv1D)	256	None, 16, 256	512	ReLu
dropout_1 (Dropout)	Dropout Rate- 20%	None, 16, 256	0	None
conv1d_2 (Conv1D)	128	None, 16, 128	32896	ReLu
dropout_2	Dropout Rate-	None, 16, 128	0	None

(Dropout)	20%			
gru_1 (GRU)	64	None, 16, 64	37056	ReLu
dropout_3 (Dropout)	Dropout Rate-20%	None, 16, 64	0	None
gru_2 (GRU)	32	None, 32	9312	ReLu
dropout_4 (Dropout)	Dropout Rate-20%	None, 32	0	Tanh
dense_1 (Dense)	8	None,8	264	Tanh
dense_2 (Dense)	4	None,4	36	Tanh
dense_3 (Dense)	2	None,2	10	Tanh
dense_4 (Dense)	1	None,1	3	Sigmoid

Table6: Summary of the Stacked Convolutional-GRU model

Number of Epochs	30
Batch Size	64
Loss function	Binary Cross-entropy
Optimizer Used	Adam

Table 7: Specification of Hyper-Parameters

Experimental Results

During training, while fitting the training data into the stacked Convolutional-GRU classifier, the training process is evaluated in terms of accuracy as well as loss. For each epoch, data loss and accuracy is calculated. The best performing model will show accuracy to be increased as the number of epochs is increased. Similarly, the best model will show loss to be decreased when the number of epochs is increased. Accuracy and loss obtained for each epoch during training process of the classifier is shown in Fig 1.

After completion of training process, testing data is used for acquiring predictions. The prediction result is evaluated in terms of accuracy, f1-score and MSE. The evaluated results are shown in Table 8. Children mental health status is predicted from this model with an accuracy of 89.59%.
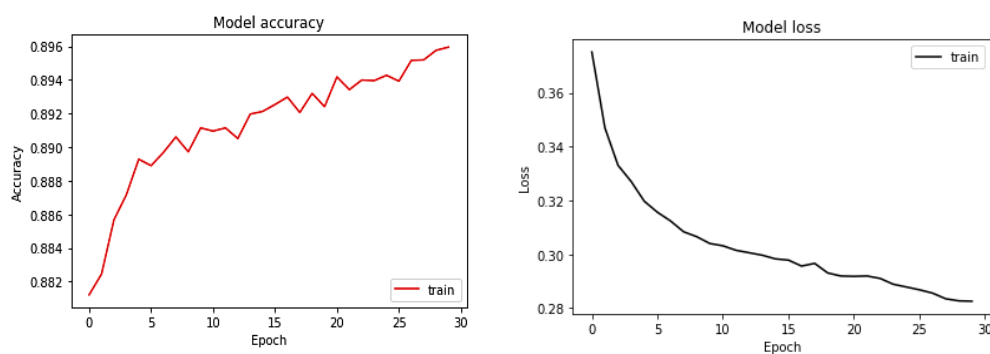
Fig.1. Accuracy and loss acquired for each epoch of Stacked Convolutional-GRU model

Performance Measure Metrics	Accuracy	MSE	F1-Score
Stacked Convolutional-GRU Model	89.57%	0.1041	0.896

Table 8: Prediction performance of Stacked Convolutional-GRU model

Conclusion

This study has contributed for reducing the high risk of impairment happened among the children. Well-being of children is majorly affected by the present lockdown scenario. An extensive data collection has been carried out on people residing on West Bengal, India. Mental health of children is assessed from cities, urban, rural area perspectives. The children age group from 4 to 12 is majorly concentrated in this context. The collected is further analyzed using DL framework that hybrids convolutional layer and GRU as a single entity. This predictive model automatically finds the underlying relationships among the collected features and takes an informed decision about whether a child has mental health anxiety or not. Prediction of children mental status in advance using the proposed method may help psychiatrists for treatment of anxiety. This study can even be extended for assessing the mental health disorders among adolescences which provides insight of the future work of this research.

References

[1] Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., ... & Rahman, A. (2011). Child and adolescent mental health worldwide: evidence for action. *The Lancet*, *378*(9801), 1515-1525.

[2] Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., ... & Burstein, R. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The lancet*, *382*(9904), 1575-1586.

[3] World Health Organization(WHO). (2020, March 18). *Mental health and psychosocial considerations during the COVID-19 outbreak.* Retrieved from https://www.who.int/publications/i/item/WHO-2019-nCoV-MentalHealth-2020.1

[4] Carbonell, J. G., Michalski, R. S., & Mitchell, T. M. (1983). An overview of machine learning. In *Machine learning* (pp. 3-23). Morgan Kaufmann.

[5] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). Cambridge: MIT press.

[6] Murphy, J. (2016). An overview of convolutional neural network architectures for deep learning. Microway Inc.

[7] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

[8] Tsoi, A. C. (1997, September). Recurrent neural network architectures: an overview. In *International School on Neural Networks, Initiated by IIASS and EMFCSC* (pp. 1-26). Springer, Berlin, Heidelberg.

[9] Koning, N. R., Büchner, F. L., Vermeiren, R. R., Crone, M. R., & Numans, M. E. (2019). Identification of children at risk for mental health problems in primary care—Development of a prediction model with routine health care data. *EClinicalMedicine*, *15*, 89-97.

[10] Srividya, M., Mohanavalli, S., & Bhalaji, N. (2018). Behavioral modeling for mental health using machine learning algorithms. *Journal of medical systems*, *42*(5), 88.

[11] Dabek, F., & Caban, J. J. (2015, August). A neural network based model for predicting psychological conditions. In *International conference on brain informatics and health* (pp. 252-261). Springer, Cham.

[12] Basavappa, S. R., Rao, S. L., Harish, B., Basavappa, S. R., & Rao, S. L. (1996). Expert system for dementia/depression diagnosis. *Nimhans Journal*, *14*, 99-106.

[13] Sau, A., & Bhakta, I. (2019). Screening of anxiety and depression among seafarers using machine learning technology. *Informatics in Medicine Unlocked*, *16*, 100228.

[14] Razzouk, D., Mari, J. D. J., Shirakawa, I., Wainer, J., & Sigulem, D. (2006). Decision support system for the diagnosis of schizophrenia disorders. *Brazilian journal of medical and biological research*, *39*(1), 119-128.

[15] Comin Nunes, L., Pinheiro, P. R., Cavalcante, T. P., & Pinheiro, M. C. D. (2015). Handling diagnosis of schizophrenia by a hybrid method. *Computational and Mathematical Methods in Medicine, 2015*.

[16] Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*.

[17] Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1--learning rate, batch size, momentum, and weight decay. *arXiv preprint arXiv:1803.09820*.

[18] Baldi, P., & Sadowski, P. J. (2013). Understanding dropout. In *Advances in neural information processing systems* (pp. 2814-2822).

[19] Jais, I. K. M., Ismail, A. R., & Nisa, S. Q. (2019). Adam optimization algorithm for wide and deep neural network. *Knowl. Eng. Data Sci.*, *2*(1), 41-46.