International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

A STUDY ON FORECASTING ON DEPRESSED MOOD BASED SELF REPORTED HISTORIES USING RECURRENT NEUTRAL NETWORKS

B.VENKATA SRINIVASULU
Research Scholar, Department of Computer Science
JJT University, Rajasthan, India
Dr. S NAGAPRASAD
Faculty of Computer Science and applications
Tara government college, Sangareddy.
Dr. VINOD MORESHWAR VAZE
Associate Professor
Department of Computer, JJT University, Rajasthan, India

ABSTRACT:
Depression is a common illness worldwide with potentially severe implications. Early identification of depressive symptoms is a crucial first step towards assessment, intervention, and relapse prevention. With an increase in data sets with relevance for depression, and the advancement of machine learning, there is a potential to develop intelligent systems to detect symptoms of depression in written material. Depression is a prevailing issue and is an increasing problem in many people's lives. Without observable diagnostic criteria, the signs of depression may go unnoticed, resulting in high demand for detecting depression in advance automatically. This paper tackles the challenging problem of forecasting severely depressed moods based on self-reported histories. Despite the large amount of research on understanding individual moods including depression, anxiety, and stress based on behavioral logs collected by pervasive computing devices such as smart phones, forecasting depressed moods is still an open question. Experimental results show that our method forecast the severely depressed mood of a user based on self-reported histories, with higher accuracy than SVM. The results also showed that the long-term historical information of a user improves the accuracy of forecasting depressed mood.
Keywords: Depression, Smart Phones, fore casting, Mobile Applications

1.0 INTRODUCTION:

Depression, or depressive disorder, is a common disease. According to the World Health Organization (WHO), the number of people with depression was estimated at more than 300 million affected worldwide. Depression may severely impact well-being and functioning at work, school, and family, and can even lead to self-harm. Adolescent depression is associated with mood disorders and severe mental illness in adult life. Amongst the top major diseases causing disability or incapability, five are mental illnesses—depression being the most prominent of these. Hence, the disease burden due to depression is vast. The prevalence of depression in the adult population is approximately 5% across cultures, and 20% in its milder forms (i.e., partial symptoms, mild depression, and probable depression). Among adults, those most at risk are within the middle-aged population. Also, the world-wide occurrence of depression is increasing, with a rise of 18% between 2005 and 2015. However, early professional intervention can improve mental symptoms (e.g., absence of self-confidence and rumination) and resolve somatic problems (e.g., gastrointestinal problems and sleeping disorders) in most of the cases.

Depression is a highly prevalent disorder associated with a huge loss of quality of life, increased mortality rates high levels of service cost. Depression is currently the fourth disorder worldwide in terms of disease burden. A lot of

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

developments in treatments for depression can be seen in the last decade, where a shift is taking place from the more traditional face-to-face counseling to self-help therapies or blended care settings. These changes have been driven by advancements in technologies in society: Internet and mobile phones are widely available and enable more technologically supported forms of interventions.

2.0 LITERATURE REVIEW:

Jonas Busk et al (2020) Accurate forecasting of symptom scores can be used to improve disease monitoring, enable early intervention, and eventually help prevent costly hospitalizations. Although several studies have examined the use of smart phone data to detect mood, only few studies deal with forecasting mood for one or more days. Aimed to examine the feasibility of forecasting daily subjective mood scores based on daily self-assessments collected from patients with bipolar disorder via a smart phone-based system in a randomized clinical trial.

Cho et al (2019) conducted a prospective observational cohort study to evaluate the mood of 55 patients with major depressive disorder and bipolar disorder types 1 and 2. They collected light exposure data passively via mobile phones of patients and self-reported daily mood scores. Using activity trackers, they registered activity, sleep, and heart rate data. This information was then processed into 130 features based on circadian rhythms, and mood prediction was performed using the RF method. Their approach generally showed good sensitivity and specificity for mood state and episode prediction.

Umematsu et al (2019) compared nontemporal (SVM and LR) and temporal (long short-term memory [LSTM]) machine learning methods to forecast the stress level of the upcoming day using a predefined number of days of previous data (physiological signals, mobile phone use, location, and behavioral surveys).

Jaques et al (2019) conducted a study using physiological signals, location, smartphone logs, and survey responses collected over a month from 206 college students to model students' happiness. They applied classical machine learning methods, such as support vector machines (SVMs), random forests (RFs), neural networks, logistic regression (LR), k-nearest neighbor, naive Bayes, and Adaboost, to perform the classification task and reported 70% accuracy.

3.0 Forecasting:

Forecasting is the task of predicting the future, given all available information from the past and present. For forecasting to be feasible, it should be reasonable to assume that the history of recorded information somehow relates to the predicted future events. A typical forecasting task is illustrated in Figure 1; w denotes the size of history used in the forecast and h denotes the horizon of how far into the future the target is predicted. In our case of using daily self-reports, both w and h are measured in days.

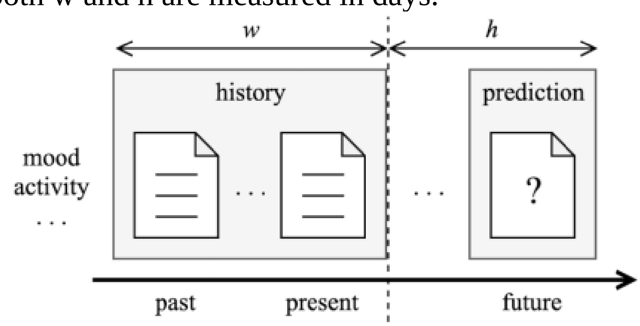

Figure 1. Forecasting is the task of predicting the future, given all relevant information from the past and the present. The window size, w, is the size of history defining the predictor variables and the horizon, h, is how far in the future the target variable is predicted.

Several methods for producing forecasts exist. The simplest forecasting methods use only historic information about the target variable and do not consider any other information but are designed to utilize time-dependent patterns, such as seasonality and trend, to extrapolate observed data into the future. Another approach is to apply standard regression or classification models to predict the variable of interest based on relevant information, such as prior (lagged) observations of the target variable along with additional predictor variables. This approach has the benefit of allowing the use of a variety of different methods from the machine learning and statistical inference literature but may not be as good at capturing long-term time-dependent patterns. For short-term forecasts or data without long-term time dependence, however, this might not be a problem. For these reasons, we chose to apply the latter approach in this work.

The corresponding training sets consist only of data observed before each test set. Thus, no future information is included when constructing the forecasts. The cross-validation error is then computed across all the test sets. As we considered data from multiple individuals, we applied two different time-series cross-validation in our experiments:

1. Leave-all-out time-series cross-validation: Each individual's data are partitioned into a sequence of T consecutive similar-sized test sets. Then the test sets are pooled across all individuals. The corresponding training sets consist of all data observed before each test set, resulting in $T-1$ test and training set pairs (the first test set has no prior data).
2. Leave-one-out time-series cross-validation: Each individual's data are partitioned into a training set and subsequent test set. The training set is then pooled with all data from all other individuals, resulting in a number of test/training set pairs equal to the number of individuals, J.

These two experiments correspond to two different scenarios: the leave-all-out time-series cross-validation simulates a situation where a group of patients starts monitoring at the same time without any additional historical data, whereas the leave-one-out time-series cross-validation simulates a situation where each participant starts monitoring when data are already available from a population of similar individuals.

3.1 Sparsity of mood reports: A quick inspection of the dataset revealed that users did not always report even if they were prompted to do so. Figure 1c shows the complementary cumulative distribution function (CCDF) of moods reported per participant, including those they were prompted to fill (expected), the ones they were prompted to fill but did not (missed) and the ones they filled (complete). This is also true for users that used the app for large periods. Indeed, those that used the app for 45 or more consecutive days (16, 8% of the users) reported, on average, less than half of the expected times. The absence of mood reports might be a symptom of boredom or dissatisfaction with the app, but could also be indicative of mental disorders, especially in cases where users have been reporting anger and depression related feelings.

3.2 Variability of mood reports: A longitudinal exploration of the mood reported shows large differences between users in the way they report, in terms of both specific positions on the grid and area covered. Figure shows moods reported by two different individuals who have self-reported for, at least, 300 days, and who are representative of two different behavioral patterns we identified. The first user reports consistently over time, both in the short and long term, and her reports are concentrated on the positive and calm area of the grid. As time goes, her reports progressively become more negative (but still in the positive area) and active. The second user has quite the opposite behavior. That is, at the beginning (purple dots), she reported mixed affect states during consecutive days (purple dots are almost all over the grid), but, as time goes, her reports concentrate mainly in the negative and active area.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

3.3 Deep feed forward neural network:

Artificial neural network (ANN) is proposed with the intention of mimicking how human brain works, where the basic element is an artificial neuron depicted Mathematically, an artificial neuron is a nonlinear transformation unit, which takes the weighted summation of all inputs and feeds the result to an activation function, such as sigmoid, rectifier (i.e., rectified linear unit [ReLU]), or hyperbolic tangent. An ANN is composed of multiple artificial neurons with different connection architectures. The simplest ANN architecture is the feed forward neural network (FNN), which stacks the neurons layer by layer in a feed forward manner where the neurons across adjacent layers are fully connected to each other. The first layer of the FNN is the input layer that each unit receives one dimension of the data vector. The last layer is the output layer that outputs the probabilities that a subject belonging to different classes (in classification). The layers between the input and output layers are the hidden layers. A DFNN usually contains multiple hidden layers.

3.4 Recurrent neural network:

RNNs were designed to analyze sequential data such as natural language, speech, and video. Given an input sequence, the RNN processes one element of the sequence at a time by feeding to a recurrent neuron. To encode the historical information along the sequence, each recurrent neuron receives the input element at the corresponding time point and the output of the neuron at previous time stamp, and the output will also be provided to the neuron at next time stamp (this is also where the term "recurrent" comes from). The recurrence link (i.e., the edge linking different neurons) enables RNN to capture the latent semantic dependencies among words and the syntax of the sentence. In recent years, different variants of RNN, such as long short-term memory (LSTM) and gated recurrent unit have been proposed, and the main difference among these models is how the input is mapped to the output for the recurrent neuron. RNN models have demonstrated state-of-the-art performance in various applications, especially natural language processing (NLP; e.g., machine translation and text-based classification); hence, they hold great premise in processing clinical notes and social media posts to detect mental health conditions.

3.5 Stacked RNNs: In our examples thus far, the inputs to our RNNs have consisted of sequences of word or character embeddings (vectors) and the outputs have been vectors useful for predicting words, tags or sequence labels. However, nothing prevents us from using the entire sequence of outputs from one RNN as an input sequence to another one. Stacked RNNs Stacked RNNs consist of multiple networks where the output of one layer serves as the input to a subsequent layer, as shown in Fig.

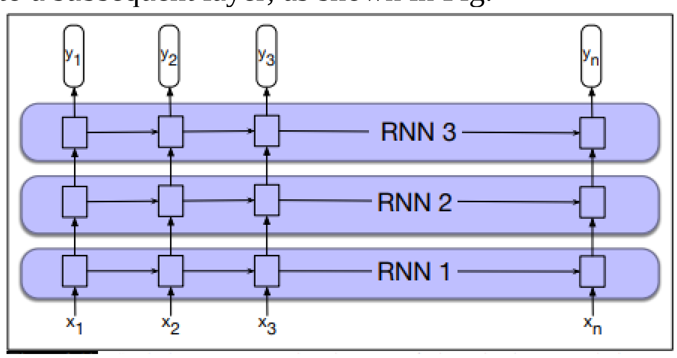

Fig 2: Stacked recurrent networks. The output of a lower level serves as the input to higher levels with the output of the last network serving as the final output.

The optimal number of stacked RNNs is specific to each application and to each training set. However, as the number of stacks is increased the training costs rise quickly.

4.0 METHODOLOGY:

With our large dataset, we used a supervised machine learning technique to build a predictive model for forecasting severe depression. To use individual histories as time series data, we

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

needed a technique that was capable of incorporating dependencies from previous states. Among the techniques that are capable of handling sequential features, RNNs with hidden LSTM units are known to be powerful models for learning from sequential data. They effectively model varying length sequences and capture long range dependencies. Following the application of RNNs to clinical diagnoses classification, we present the first study to evaluate the ability of LSTM to forecast severe depression. The method also recognizes patterns in time series of self-reported user histories by visualizing representative patterns of nodes in a hidden layer.

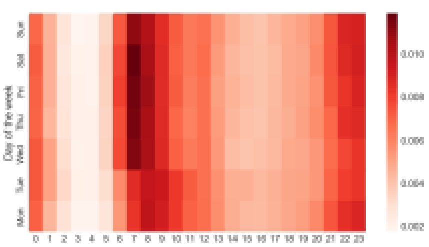

Figure 3: Distribution of user input hours for each day of the week

Our model prepares the same number of nodes as the number of categories in each embedding layer. Features are extracted from collected logs shown in Table 1. We extract bedtime and wake-up time into sleeping hour and sleeping irregularity features. Sleeping hour is divided into 24 bins. We create a sleeping irregularity feature to capture the irregularity in sleeping hours, following previous studies that showed the relationship between sleeping irregularity and depressive moods. The sleeping irregularity feature takes 1 if the sleeping hour during day t is more different than 3 hours compared to the sleeping hour during day t − 1, and 0 otherwise.

Long Short Term Memory Networks (LSTM): LSTMs have the ability to learn long-term dynamics while avoiding vanishing and exploding gradient problems and have recently gained great success in sequence learning tasks such as speech recognition and machine translation. We designed our LSTM with a single LSTM layer with 32 nodes and a dropout of 0.2. Drop-outs were used between the LSTM and dense layers. The output of the last cell of the LSTM layer was connected to a dense layer. Finally, a sigmoid activation layer predicted the high/low stress levels. We trained our LSTM using RMSprop with binary cross-entropy loss and an iteration number of 1000.

5.0 RESULTS:

The dataset consists of 15,975 daily self-assessments and 280 clinical evaluations from 84 participants. This corresponds to an average of 190.2 self-assessments per individual and an average self-assessment adherence of 82.8% between the first and last submitted self-assessment. The population ranged from the ages of 21 to 71 years (mean 43.1, SD 12.4) and consisted of 62% (52/84) women. Figure 3 presents the distribution of self-reported mood scores across all individuals in the dataset (mean −0.14, SD 0.48). The majority of observed mood scores, y, are centered around zero, indicating euthymia (−0.75<y0.75=1.68%). As expected, the self-reported mood scores and HDRS scores were negatively correlated (r=−0.40; P<.001).

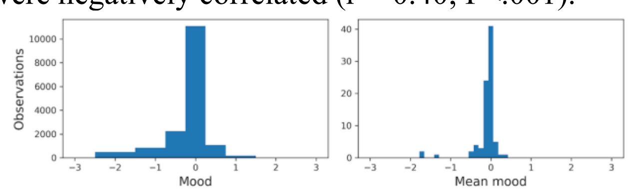

Figure 4. Distribution of all self-reported mood scores (left) and individual mean mood scores (right). The mood scores are generally close to zero indicating neutral mood with only a few exceptions indicating depressed or elevated mood

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

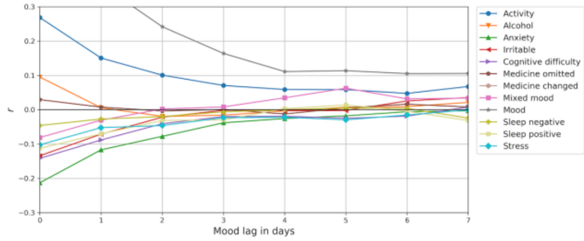

Figure 5. The mean of individual correlations of self-assessment items and mood lagged up to 7 days. Nonzero correlations indicate that items have some relation to mood on subsequent days that can be utilized for mood forecasting

Figure shows correlations of self-assessment items with self-reported mood lagged for up to seven days. Self-reported mood has a positive autocorrelation for the entire duration of 1 to 7 days. Additionally, activity has a positive correlation with mood for a few days, indicating that high activity levels coincides with elevated mood, and anxiety has a small negative correlation with mood, indicating that anxiety often coincides with negative mood scores. The remaining self-assessment items show small, diminishing correlations with lagged mood. A seasonality analysis of self-reported mood revealed no significant monthly or daily seasonality in the data and was left out for brevity.

Figure shows the distribution of moods at each part of the day. Higher values denote more positive feelings during the periods. We confirm that the general trend of "afternoon > evening > morning" and "weekend > weekday," except for Sunday evening. The drop in the positive feeling ratio on Sunday evening shows so-called Sunday night blues. This coincides with the results estimated from millions of Tweets in. The figure also supports the necessity of collecting user records for each part of the day separately because individual mood strongly depends on a part of the day.

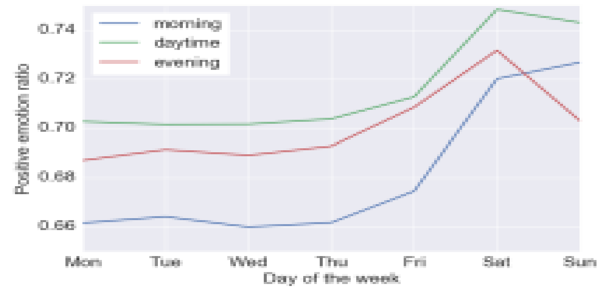

Figure 6: Positive feeling ratio for each day of the week. Each line corresponds to a time period in a day. The positive feeling ratio denotes the ratio of the count of positive feelings (fine or fair) to the total count of any feeling by all users

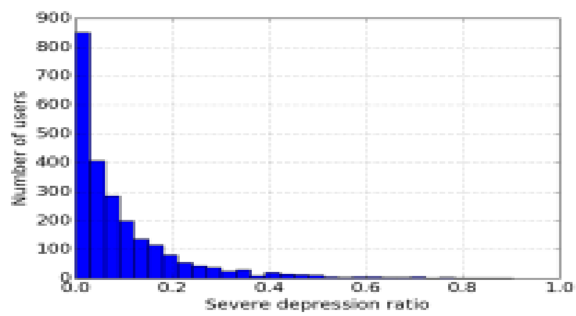

Figure 7: Distribution of the severe depression ratio of users. A total of 84.6% of users experienced at least one severe depression day in the dataset

Table 3: Experimental results. The mean value of Test AUC-ROC is listed for each method. The standard deviations are in parentheses. ∗∗∗ denotes p < 0.01. The p-values were calculated based on paired t-test after making adjustments using the Holm-Bonferroni method for multiple comparison.

	AUC-ROC		
Method (feature set)	h = 1	h = 3	h = 7
SVM (all	837 (.031	822 (.045)	822 (.047)
LSTM-RNN (severe	846 (.032	821 (.042)	800 (.053)

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

only)			
LSTM-RNN (all)	886∗∗∗ (.020)	860∗∗∗(.031)	842∗∗∗(.044)

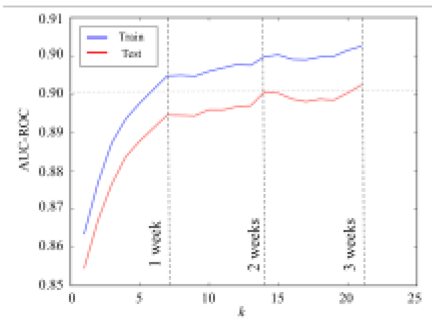

Figure 8: Forecasting performance over different days used for training. The x-axis denotes the number of days used for the model and the y-axis denotes the AUC-ROC values.

The performance curve shows a rapid increase in the first week and a slower increase in the second and third weeks. The improvement apparently saturates as k increases. After k is higher than 14, the performance curve drops until k reaches 21. The results indicate that the history of the previous two weeks is sufficient to forecast future severe depression. This finding coincides with the design of commonly used mental health assessment such as the PHQ-9 and GAD-7, whose first questions are, "Over the last two weeks, how often have you been bothered by any of the following problems?" This consistency empirically ensures that the assessments' questions are reasonable for capturing individual mental health status.

6.0 CONCLUSION:

Mood prediction and forecasting can be used as early warning signs in clinical treatment. Furthermore, accurate symptom forecasting could be extended to detect risk of relapse of major affective episodes specifically, eg, by detecting if values exceed predefined thresholds over consecutive days. This could be useful in, eg, a telemedicine setup in which trained nurses or other clinical personnel supervise patients in outpatient treatment. Experimental results confirmed that our framework was able to forecast severe depression based on individual histories with high accuracy. The results showed that fine-grained information such as reporting moods in different parts of the day improved forecasting performance. The capability of LSTM-RNN to incorporate long-range dependencies of time series helped us determine the contribution of distant past histories up to the previous two weeks. The representative patterns of the model further showed that having a depressed mood only in the afternoon is not always a sign of future severe depression.

REFERENCES:

1. Jonas Busk, Maria Faurholt-Jepsen, Mads Frost, Jakob E Bardram, Lars Vedel Kessing, Ole Winther (2020), "Forecasting Mood in Bipolar Disorder From Smartphone Self-assessments: Hierarchical Bayesian Approach", JMIR Mhealth And Uhealth, Vol 8, No 4, PP: 1-14
2. Cho CH, Lee T, Kim MG, In HP, Kim L, Lee HJ. Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: prospective observational cohort study. J Med Internet Res. 2019 Apr 17;21(4)
3. Umematsu T, Sano A, Taylor S, Picard R. Improving students' daily life stress forecasting using LSTM neural networks. Proceedings of the IEEE EMBS International Conference on Biomedical & Health Informatics (BHI); IEEE EMBS International Conference on Biomedical & Health Informatics (BHI); May 19-22, 2019; Chicago, IL, USA. 2019.
4. Jaques N, Taylor S, Azaria A, Ghandeharioun A, Sano A, Picard R. Predicting students' happiness from physiology, phone, mobility, and behavioral data. Int Conf Affect Comput Intell Interact Workshops. 2019 Sep;2015:222–8
5. Pollak JP, Adams P, Gay G. PAM: a photographic affect meter for frequent, in situ

measurement of affect. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Proceedings of the annual conference on Human factors in computing systems - CHI ?11; May, 2011; Vancouver BC, Canada. 2011. pp. 725–34.
6. A. Sano et al., “Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: Observational study,” J Med Internet Res, vol. 20, no. 6, p. e210, Jun 2018.
7. S. A. Taylor, N. Jaques, E. Nosakhare, A. Sano, and R. Picard, “Personalized Multitask Learning for Predicting Tomorrow’s Mood, Stress, and Health,” IEEE Transactions on Affective Computing, no. 99, pp. 1–14, 2017
8. T. W. Colligan and E. M. Higgins, “Workplace stress: Etiology and consequences,” Journal of Workplace Behavioral Health, vol. 21, no. 2, pp. 89–97, 2006.
9. S. Saeb et al., “Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study,” Journal of Medical Internet Research, vol. 17, no. 7, p. e175, jul 2015.
10. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2163-2196
11. Sanchez-Moreno J, Martinez-Aran A, Colom F, Scott J, Tabares-Seisdedos R, Sugranyes G, et al. Neurocognitive dysfunctions in euthymic bipolar patients with and without prior history of alcohol use. J Clin Psychiatry 2009 Aug;70(8):1120-1127
12. Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. Annu. Rev. Clin. Psychol. 2018;14:91–118
13. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J. Med. Internet Res. 2017;19:e228.
14. Hamilton M. Development of a rating scale for primary depressive illness. Br. J. Soc. Clin. Psychol. 1967;6:278–296
15. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief. Bioinformatics. 2017;19:1236–1246.