Applications of Machine Learning in the Field of Depression

Authors: Dong Jianyu, Wei Wenqi, Wu Ke, Nina, Wang Canfei, Fu Ying, Peng Xin, Peng Xin

Date: 2019-09-19T00:00:00+00:00

Abstract

Insufficient disease awareness among patients with depression and the lack of early screening methods result in most patients having already progressed to major depressive disorder by the time of diagnosis. To address this situation, machine learning has been gradually applied in recent years to early prediction, early identification, auxiliary diagnosis, and treatment decision-making for depression. In practical applications, factors influencing the accuracy of machine learning models include dataset type and scale, feature engineering, algorithm type, among others. It is recommended that machine learning be further integrated into healthcare systems and mobile applications in the future, with continuous optimization of machine learning models, to improve issues related to the prevention, identification, diagnosis, and treatment of depression through comprehensive mining of patient health data.

Full Text

Preamble

The Application of Machine Learning in Depression Research

DONG Jianyu; WEI Wenqi; WU Ke; NI Na; WANG Canfei; FU Ying; PENG Xin
(School of Nursing, Jilin University, Changchun, 130012, China)

Abstract: The lack of disease awareness among patients with depression and the absence of early screening methods result in most patients having already progressed to major depressive disorder by the time of diagnosis. To improve this situation, machine learning has been gradually applied to early prediction, early identification, auxiliary diagnosis, and treatment decision-making for depression in recent years. In these applications, factors affecting the accuracy of

machine learning models include the type and scale of datasets, feature engineering, algorithm types, etc. It is recommended that future efforts further integrate machine learning into healthcare systems and mobile applications, continuously optimize machine learning models, and improve depression-related issues in prevention, identification, diagnosis, and treatment by fully mining patient health data.

Keywords: machine learning; depression; prediction; algorithm; model
Classification Number: B842

Depression, also known as depressive disorder (American Psychiatric Association, 2013), is a mood disorder characterized by significant and persistent low mood, with chronic, recurrent, protracted, and high suicide risk features. The World Health Organization (WHO) predicts that by 2020, depression will become the second leading disease worldwide, surpassing cancer (Brundtland, 2001). Approximately 8% of men and 15% of women will develop depressive disorders during their lifetimes, severely impacting quality of life, with nearly 15% of these individuals choosing suicide (Gold, Machado-Vieira, & Pavlatou, 2015). Therefore, early identification and diagnosis of depression patients followed by timely treatment is extremely important.

Currently, clinical identification and diagnosis of depression primarily rely on ICD-10 or DSM-V diagnostic criteria, combined with patient interviews, scales, and clinicians' diagnostic experience (Ren et al., 2016; Zhou, Lu, Wu, & Xu, 2017). This approach is only suitable for one-on-one assessment and is prone to misdiagnosis due to subjective factors such as patient cooperation and clinician proficiency. Furthermore, the lack of disease awareness and early screening tools means patients generally already meet criteria for major depressive disorder (MDD) by the time they seek treatment. Early detection of depressive tendencies and mild depression facilitates timely intervention and prevents worsening of the condition. Additionally, individual differences mean that identical treatment protocols produce inconsistent therapeutic effects across individuals. Consequently, finding rapid, objective, and accurate methods for depression identification and diagnosis, as well as determining optimal individualized treatments, is of great significance. In recent years, with rapid development of internet and scientific technology, machine learning, due to its powerful data processing and mining capabilities, has been widely applied in healthcare and gradually introduced into research on early prediction, identification, and auxiliary diagnosis of depression. However, domestic research reports in this area remain relatively scarce. This article reviews the application of machine learning in depression to provide reference for domestic research in this field.

2. Machine Learning Concepts and Principles

Machine learning (ML) is an interdisciplinary field involving multiple domains and represents the core of artificial intelligence. Its purpose is to enable computers to possess self-learning capabilities, thereby continuously improving and enhancing computer performance in data processing. Through learning from existing data and information, machine learning acquires potential patterns within data and applies these patterns to analyze and predict unknown data (Cabitza & Banfi, 2018; Senders et al., 2018). Based on whether output is manually labeled, it is divided into supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning. Based on model structure depth, it is divided into traditional machine learning and deep learning. The relationship between deep learning and machine learning is shown in Figure 1 [Figure 1: see original paper].

Currently, machine learning algorithms applied to depression are primarily traditional machine learning methods such as support vector machines, random forests, K-nearest neighbor algorithms, and shallow artificial neural networks (Zhao & So, 2019; Lee et al., 2018). However, with recent deep learning development, convolutional neural networks, autoencoders, deep belief networks, and others have gradually been applied to depression research (Li et al., 2019; Helbich et al., 2019; Lin et al., 2018). The basic principle of establishing depression prediction models using machine learning involves collecting depression risk factors, biomarkers, and other data, preprocessing these data to obtain normalized datasets, dividing them into training and testing sets according to a certain ratio, using the training set to train machine learning algorithms, and finally using the testing set to evaluate model performance while continuously optimizing the model during the validation and evaluation process. The conceptual framework for establishing depression prediction models based on ML is shown in Figure 2 [Figure 2: see original paper].

3. Data Types and Collection for Machine Learning-Based Depression Research

The advantage of machine learning lies in mining potential patterns within data. Therefore, the first step in establishing prediction models is data collection. Data collection methods have become diversified due to scientific and technological development, such as collecting survey data through questionnaires, acquiring health data through various sensors, and obtaining public data from online platforms (Fu et al., 2018; Kantoch, 2018; Enshaeifar et al., 2018; Marinelarena-Dondena et al., 2017). Data collection methods and data types for machine learning-based depression research also demonstrate diversity.

Depression prediction and identification using machine learning are based on factors reflecting depression tendencies and various disease markers, including depression influencing factors, symptom manifestations, and physiological characteristics. Currently, these primarily include sociodemographic data such as age, gender, and substance abuse; physical symptoms and psychological status data collected clinically; and physiological signal data acquired by medical instruments including electroencephalography (EEG), magnetic resonance imaging (MRI), eye movement data, and heart rate variability parameters (Patel, Khalaf, & Aizenstein, 2016; Wu et al., 2018; Garcia-Ceja et al., 2018; Lee et al., 2018; Lü, 2014). Among these, applications based on EEG and MRI are most widespread, but the collection cost of physiological signals such as MRI is relatively high. Therefore, scholars have gradually explored data with lower collection costs but good predictive power for depression, such as speech and facial expressions (He & Cao, 2018).

With the development of computers and information technology, large numbers of users express emotions and communicate daily through social network platforms such as Facebook and Twitter. This data provides opportunities for health departments to study users' mental health and mood disorders by combining natural language processing technology, sentiment analysis, and machine learning (Marinelarena-Dondena et al., 2017). Islam et al. (2018) used decision trees based on linguistic behavioral features in Facebook users' posts to establish a depression prediction model with classification accuracy reaching 99.0%. Simultaneously, the development of Internet of Things technology has enabled information exchange between objects, allowing various wearable devices to carry massive amounts of health information. Research has already applied these data to the depression field through machine learning (McGinnis et al., 2018). Scholars have also used smartphone sensor apps to obtain data on daily activity levels, sleep conditions, and social communication to predict individual depression status (Sarda et al., 2019).

4.1 Analyzing Influencing Factors of Depression Onset and Prognosis

Many factors influence depression onset and prognosis, such as genetic factors, dietary nutrition, and hormone levels (Jin & Ding, 2017; Yu & Niu, 2015), and methods for identifying these influencing factors continue to be updated. Predictive factors identified by traditional statistical methods have weak individual predictive power, necessitating the establishment of multiple linear regression models based on multiple predictive factors. However, to maintain stability and reproducibility of traditional linear models, the number of predictive factors must remain small relative to sample size, and subsets of factors cannot be highly correlated. These issues cause traditional multiple regression to inadequately utilize potential predictive factors and fail to fully mine data (Pearson, Pisner, Meyer, Shumake, & Beevers, 2018).

Machine learning, as a new statistical method, can capture simultaneous effects of all relevant predictive factors, outperforming traditional stepwise analysis methods. Therefore, many scholars use machine learning to analyze influencing factors of depression onset and prognosis. Luo et al. (2017) used parameters provided by random forest models to clarify influencing factors for depression onset in post-stroke patients, with the top three being stroke lesion location, traditional Chinese medicine intervention methods, and whether stroke patients had a family history of depression. Through single-rule algorithms, they determined that the most important factor for post-stroke depression was stroke lesion location, with patients having lesions in frontal and temporal lobes showing higher probabilities of post-stroke depression. Pearson et al. (2018) used elastic net and random forest models to explore predictive factors for the effectiveness of self-guided internet interventions in depression patients, analyzing 120 candidate factors including demographic variables, clinical data, and system usage. The results identified 16 important predictive factors including pre-treatment depression total scores, psychiatric comorbidities, dysthymia, and system module usage. This machine learning-based analysis of influencing factors can better understand depression mechanisms and lay foundations for subsequent accurate prediction of depression risk and prognosis.

4.2 Screening for Depression Risk

Like most chronic diseases, depression is non-sudden and factor-induced. If potential depression risk and depressive tendencies in patients can be anticipated in advance, early intervention can be implemented to timely correct patients' dangerous behaviors and prevent depression onset and worsening. This early prediction function can be achieved by training machine learning prediction models. For example, Sau and Bhakta (2017a) used sociodemographic variables such as age, population, and substance abuse in elderly populations, along with comorbidity data including vision problems, hearing problems, and insomnia, to establish a depression risk prediction model through multilayer perceptron with 97.2% accuracy. Under this model, primary healthcare personnel can screen for elderly individuals at risk of depression by inputting corresponding socioeconomic and physical symptom data. Moreira et al. (2019) combined machine learning with Internet of Things and cloud computing technologies, using data from pregnant women during pregnancy to establish an emotion-aware intelligent system for predicting postpartum depression, which can effectively detect mothers with depressive tendencies and provide timely intervention to prevent postpartum depression.

4.3 Assisting in Depression Identification and Diagnosis

The growing number of depression patients has increasingly burdened doctors' time and energy, making large-scale identification of depression patients through complex psychological measurement tools nearly impossible. Therefore, machine learning has been used to assist clinical diagnosis and automated identification to classify depression patients and healthy populations. Zheng et al. (2017) conducted nuclear magnetic resonance spectroscopy analysis of plasma sample metabolites from 72 major depression patients and 54 normal individuals, establishing a classification model based on support vector machine with 96.0% accuracy. This model achieved auxiliary diagnosis of major depressive disorder while excluding subjective factors. Sau and Bhakta (2017b) trained multiple machine learning algorithms based on sociodemographic characteristics and health-related factors in elderly populations, achieving maximum prediction accuracy of 91.0%. Although prediction accuracy was slightly reduced, this approach had advantages of low data source cost and easy availability.

These studies primarily focused on diagnosis of major depressive disorder; however, early identification of mild depression is particularly important for preventing severe consequences such as major depression and suicide. Li et al. (2019) established a computer-aided detection (CAD) system for mild depression based on convolutional neural networks, achieving 85.6% prediction accuracy. This system can achieve objective and rapid identification of mild depression patients. Research has also demonstrated that classification models based on information such as user-generated pictures, posts, and comment likes on social media platforms including Instagram, Twitter, and Facebook can identify user depression status (Ricard, Marsch, Crosier, & Hassanpour, 2018; Fatima, Mukhtar, Ahmad, & Rajpoot, 2018). To protect user privacy, such studies sign anonymity agreements and informed consent before data extraction and mine social media user data using machine learning under ethics committee supervision. Depression prediction models based on social media can expand identification of depressive mental disorders at the population level, enhance disease awareness among undiagnosed depression social media users, guide them to seek timely medical care, obtain early diagnosis and receive corresponding psychological intervention or medication treatment, and prevent outcomes such as suicide.

4.4 Assisting in Depression Treatment Decision-Making

Depression treatment modalities primarily include medication, face-to-face psychotherapy, internet-based interventions, and repetitive transcranial magnetic stimulation (rTMS). However, individual responses to specific treatment modalities vary. If patient responses to treatment can be predicted in advance, optimal treatment plans can be selected based on prediction results, avoiding blind trial-and-error treatment (Foster, Mohler-Kuo, Tay, Hothorn, & Seiborn, 2019) and achieving cost savings and improved cure rates. In recent years, machine learning has been gradually used to guide depression treatment selection due to its ability to predict treatment efficacy based on patients' pre-treatment baseline characteristics (Hollon, Cohen, Singla, & Andrews, 2019).

Chekroud et al. (2016) used gradient boosting machines to predict major depression patients' responses to medication treatment. The established prediction model achieved 60.0% classification accuracy for whether major depression patients would respond to escitalopram, with slightly lower prediction effects for responses to other medications. Although overall accuracy was not very high, this study revealed the possibility of using machine learning to prospectively predict patient medication responses. Zilcha-Mano et al. (2018) used random forests to analyze predictive factors for placebo, psychotherapy, and medication treatment effects in depression patients. Compared to traditional statistical methods that consider single influencing factors, this approach conducted collaborative analysis of influences among multiple complex variables. Bailey et al. (2019) showed that using support vector machines based on resting EEG signals one week after depression patients started treatment could predict whether they would respond to rTMS treatment, thereby guiding non-responding patients to select other suitable treatment methods as early as possible and avoid blindly receiving ineffective, treatment-delaying, and expensive rTMS therapy.

Machine learning has also been applied to other aspects of depression research. Wallert et al. (2018) used random forests based on 34 predictive factors including demographic characteristics and clinical data to predict adherence to self-guided internet-based cognitive behavioral therapy among post-myocardial infarction depression and anxiety patients. This model can early identify patients with low adherence for timely adherence interventions or adoption of face-to-face cognitive behavioral therapy to improve symptoms. Dipnall et al. (2016) applied machine learning clustering algorithms to cross-sectional surveys of 3,922 community populations to analyze relationships between depression and other coexisting symptoms, finding that depression patients were more common in pain, gastrointestinal, and urinary clusters, providing direction for further exploration of depression pathophysiological mechanisms and treatment targets. Zhao et al. (2019) used machine learning for drug repositioning, predicting new indications based on drug expression profiles to find medications for treating mental illnesses such as depression. This drug repurposing approach can shorten drug development cycles and significantly reduce time and economic costs compared to developing new drugs.

In summary, machine learning application research in depression has primarily focused on analyzing influencing factors of depression onset and prognosis, screening for depression risk, assisting in depression identification and diagnosis, and assisting in depression treatment decision-making. Among these, machine learning-assisted identification and diagnosis of depression is mainly used to distinguish healthy populations from major depressive disorder. Future efforts should strengthen its auxiliary diagnosis of mild depression and research on depression grading. Additionally, machine learning has significant meaning in assisting depression treatment decision-making, avoiding previous blind sequential trial treatments, achieving symptomatic treatment, improving treatment efficacy, and reducing medical costs. Future research can explore machine learning' s value in other aspects such as guiding depression drug development while optimizing applications in these areas.

5. Factors Affecting Accuracy of Machine Learning-Based Depression Prediction Models

In depression research applications, machine learning generally establishes classification models to mine data patterns for depression screening, identification, and diagnosis. Practical application effects are determined by model performance quality. Therefore, clarifying factors affecting the accuracy of depression classification models is crucial as the foundation for accurately predicting and diagnosing depression and guiding treatment decisions using machine learning, representing a prerequisite for achieving better application effects.

5.1 Sample Set Types and Scale

Depression has many influencing factors. Collecting all influencing factors for prediction may increase prediction accuracy, but extensive data collection and massive computation greatly increase time, human, and material resource costs. Therefore, achieving good prediction effects using appropriate sample sets is key to depression prediction using machine learning. Shen (2015) conducted depression identification research based on EEG and eye movement signals, with results showing that under the same machine learning algorithm, prediction models trained using EEG signals generally had higher accuracy than eye movement signal prediction models. Acharya et al. (2018) confirmed that convolutional neural network models trained based on right hemisphere EEG data had superior sensitivity, specificity, and accuracy compared to left hemisphere data. Evidently, using machine learning to select different depression markers as predictive factors yields differences in prediction accuracy. We should continuously search for optimal predictive factors to construct depression prediction models while considering both accuracy and cost.

Additionally, research has confirmed that overly small sample sizes affect prediction model accuracy due to excessively high specificity, resulting in poor model generalization ability (Gao, Calhoun, & Sui, 2018), and that single-type sample sets have lower prediction performance than comprehensive data (Lee et al., 2018; Hilbert, Lueken, Muehlhan, & Beesdo-Baum, 2017). Lee et al. (2018) conducted a systematic review and meta-analysis of literature on machine learning algorithms for predicting depression treatment outcomes, with overall prediction accuracy reaching 82.0%. Results showed that prediction models established using combined sample set types achieved 93.0% accuracy, higher than single-use phenomenological data, genetic data, neurotransmitter data, or imaging data (68.0%~85.0%). Therefore, future research should collect sufficient sample sizes as much as possible, or reduce specificity effects through cross-validation, leave-one-out, and other methods, and combine various sample set types to improve model generalization ability.

5.2 Feature Engineering

Raw collected data contains excessive redundant information that affects prediction accuracy, requiring extraction of useful information for algorithm training from the data. This process is called feature engineering, including missing value processing, feature selection, and feature extraction. Li (2018) used three different feature selection algorithms combined with the same machine learning algorithm based on the same EEG dataset, with results showing that the model established using features obtained by the correlation-based feature selection algorithm had the highest accuracy. Research has also proven that different features have different prediction abilities for results, and prediction model performance based on different feature subsets from the same dataset also differs (Kang, 2018), requiring continuous exploration to extract optimal predictive feature sets to improve prediction accuracy.

Evidently, feature engineering is very important for model prediction performance. However, implementing this data processing process requires professional knowledge background and skills (Mukherjee, Obaidullah, Santosh, Phadikar, & Roy, 2018), and represents the most time-consuming and labor-intensive part of the entire learning process. The emergence of deep learning has solved this problem, greatly simplifying feature engineering and achieving automated feature extraction (Acharya et al., 2018), with powerful ability to mine complex patterns among big data, gradually being applied to various fields including depression in recent years.

5.3 Algorithm Types

Algorithm types also affect model prediction accuracy. Li et al. (2019) established depression classification models based on EEG data using convolutional neural networks and five traditional machine learning methods. Among these, the highest classification accuracy among traditional machine learning models was 73.1%, while convolutional neural network classification accuracy reached 84.7%, representing an 11.6% improvement. Liu, Li, and Chen (2017) extracted features from EEG signals and established support vector machine and Logistic classification models with maximum accuracy of 90.0%; they then established a convolutional neural network classification model based only on filtered and denoised EEG signals, achieving 96.7% accuracy, demonstrating better depression classification effects than traditional machine learning models requiring feature extraction. Evidently, convolutional neural networks can better reveal potential relationships among data to achieve higher classification accuracy. Compared to traditional machine learning algorithms, convolutional neural networks do not require feature extraction before training and testing data, and can still guarantee classification accuracy while simplifying procedures, making the potential of deep learning worthy of further research. However, deep learning has strong data dependency, and when sample data is limited, its trained model accuracy may not necessarily be higher than traditional machine learning (Zhao & So, 2019). Therefore, appropriate training models should be selected based on actual data conditions while considering time and economic benefits. A comparison between traditional machine learning and deep learning is shown in Table 1 (Lundervold & Lundervold, 2019; Kolossvary, De Cecco, Feuchtner, & Maurovich-Horvat, 2019).

Table 1. Comparison Between Traditional Machine Learning and Deep Learning

Traditional Machine Learning	Deep Learning
Manual feature extraction	Automatic feature extraction, not required

Traditional Machine Learning	Deep Learning
Support vector machines, random forests, K-nearest neighbor algorithms, shallow artificial neural networks, etc.	Convolutional neural networks, autoencoders, recurrent neural networks, belief networks, etc.

In addition to being affected by sample set types, size, feature engineering, and algorithms, machine learning prediction model accuracy is also influenced by parameter tuning and ensemble methods (Mumtaz & Malik, 2018). Therefore, every influencing factor should be considered during algorithm training to continuously optimize and obtain optimal prediction models.

6.1 Insufficient Sample Size

Establishing prediction models with high accuracy and strong generalization ability using machine learning requires sufficient sample sizes. However, in current domestic research, due to insufficient project funding and human resources, imperfect hospital and community health systems, and lack of unified medical information management systems among hospitals, adequate sample sizes cannot be obtained during data collection. Compared to China, health systems in the United States such as the Omaha System based on standardized nursing language are more complete, and patient health information can be shared among hospitals, facilitating collection of larger datasets for data analysis using machine learning (Gao, Calhoun, & Sui, 2018). Therefore, China should establish unified and standardized patient health information databases in the future to enable multi-site data sharing among hospitals and between hospitals and communities, better mine these health data, improve the generalization ability of established models, and make them applicable to broad populations.

6.2 Need for Further Model Optimization

Model prediction accuracy is affected by multiple aspects including features, algorithms, and parameters, requiring continuous optimization to achieve optimal models. Among these, training algorithms is the core of model establishment, making algorithm selection for training critical. For the same problem and sample set, prediction models based on different algorithms present different prediction accuracies (Karhade et al., 2019; Hasanzadeh, Mohebbi, & Rostami, 2019), and each algorithm has different applicable scopes and characteristics. Lee et al. (2018)' s systematic review showed that for the same problem to be solved, different sample set types also affect prediction results, and predictive factor selection is also an important factor affecting model accuracy. Therefore, future research needs to consider various aspects including applicable data types for each algorithm, algorithm functions, and time and economic factors, make accurate judgments on which algorithm is suitable for one' s own research, conduct training based on appropriate algorithms, continuously explore optimal predictive factors, and perform parameter tuning iterations to obtain optimal prediction models.

6.3 Lack of Medical Informatics Talent

In today' s information age, massive amounts of health data are generated daily. Clinical and community hospital medical staff input patient data into electronic medical records during each shift, while medical instruments, biosensors, wearable devices, and mobile devices transmit patients' life information in real time. This data contains rich information about patients, healthcare trends, and nursing phenomena, including much depression-related information awaiting mining. Mining these data requires certain informatics technologies, but China currently has very scarce medical informatics talent (Ma & Huang, 2018). There is an urgent need to focus on cultivating medical informatics talent and improving their ability to process large amounts of complex data using machine learning and other technologies. Simultaneously, multidisciplinary cooperation models should be strengthened to promote collaboration between computer information discipline and medical discipline talents. This interdisciplinary research model will further promote the development of psychology and medicine.

6.4 Low Clinical Implementation Rate

Machine learning can reveal hidden trends in patient conditions and treatment outcomes within health data, assist in guiding medical personnel to make data-driven decisions, and improve patient treatment and care. However, most current research remains at the theoretical stage of model establishment without integrating with other computer technologies and hardware to generate operational platforms or applications. Easy-to-operate applications and platforms can facilitate individuals or primary health departments to know disease risks in advance through simple data input (Jimenez-Serrano, Tortajada, & Miguel Garcia-Gomez, 2015), perform disease prevention work, reduce medical costs, assist doctors in depression diagnosis and targeted selection of personalized treatment plans, and help reduce doctors' burdens. Therefore, future efforts can combine machine learning with Internet of Things technology, cloud computing, big data, computer software, and hardware to develop complete clinical decision support systems.

In summary, machine learning has been gradually applied to early screening, auxiliary diagnosis, and treatment decision-making in the depression field. However, obstacles and deficiencies including insufficient sample sizes, lack of medical informatics talent, and low clinical implementation rates remain in applications. Therefore, China should focus on unified storage of medical health information, strengthen cultivation of medical informatics talent, promote machine learning research in the depression field, continuously optimize constructed depression prediction models, and gradually apply them to clinical practice.