

# *Construction method of college students' depression knowledge map based on education big data*

**Shanwen Zhang, Changqing Yu, Xuqi Wang, Dengwu Wang**

*School of Electronic Information, Xijing University, Xijing, 710123, China*

**Keywords:** College students' depression; Knowledge map (KG); Depression prediction; Depression intervention

**Abstract:** Depression prediction and intervention based on education big data, combined with real-time education big data and the influence of COVID-19, we collected college students' education big data, extracted entities related to depression and the relationship between entities, formed the relationship between entities into a knowledge triplet, and then constructed a knowledge map (KG) of depression, predicted depression, and provided intervention words for depression.

## **1. Introduction**

College students have become the main group suffering from depression and drug abuse because their physical and psychological development is not yet fully mature and they have been under the pressure of family, school and society for a long time. There are many reasons for college students' depression, such as psychological imbalance, increased life pressure, high learning pressure, lovelorn, high mental pressure, and high employment pressure. The depression of college students is mainly because they feel that they are ignored by others and their self-worth cannot be reflected. Now, more and more attention has been paid to the application of AI technology in depression, which provides a new method for clinical practice and has achieved preliminary results. Through machine learning methods, individual differences in depression can be predicted and treated. Depression is a complex and heterogeneous mental disease, which brings a heavy burden to the world. Although symptomatology based diagnostic methods have been widely used in various fields, this method is not conducive to the discussion of pathological mechanisms. The predictive effectiveness of many diagnostic methods is low, which makes it difficult to accurately evaluate and compare the efficacy of various treatment schemes. Qu [1] computational psychiatry method can solve the above problems through two complementary methods of theory driven and data driven, so as to improve the understanding, prevention and treatment of depression. The theory driven method is based on empirical knowledge or assumptions, and uses computational modeling methods to conduct multi-level analysis of data. Data driven methods are based on machine learning algorithms to analyze high-dimensional data, improve the accuracy of depression diagnosis and prediction, and thus improve the accuracy of treatment. Yang Bingbing et al. [2] summarized the construction of the prediction model, introduced the research status and existing problems of depression prediction, and realized the diagnosis of depression through the prediction model. Han et al. [3] reviewed the latest research progress of AI technology applied to depression prediction, early recognition, auxiliary diagnosis and

treatment, and discussed the potential problems and challenges that followed. Yuan et al. [4] reviewed the application of AI related technologies in depression, including AI based acoustic characteristics, EEG, physiological and biochemical information, brain imaging analysis and other depression identification and auxiliary diagnosis methods, as well as their application, shortcomings and prospects in depression early warning and intervention. Zha et al. [5] proposed a capsule network model combining local and global features, including local feature selection and global feature extraction. For local features, emotional dictionary is used to select, while for overall features, capsule network model is used to learn. These two features are fused to obtain the prediction probability of microblog users' depression. The results show that the model combines emotional dictionary and capsule network, and can effectively find potential depression patients among microblog users. Many researches on natural language processing have shown that there is a relationship between a person's personal discourse and his personality, demography and mental health. However, many machine learning models that predict human characteristics do not fully consider the role of pre training language models and context embedding. Kim et al. [6] took a person's degree of depression as a case study and conducted empirical analysis to determine which ready-made language model, single-layer and multi-layer combination seems to be the most promising when applied to human level NLP tasks. Dai et al. [7] developed a new two-stage feature selection algorithm, which is based on high-dimensional (more than 30000) features constructed through context aware analysis of DAIC-WOZ dataset (including audio, video and semantic features), compared the prediction performance with seven reference models, and analyzed the preferred topics and functional categories related to reserved functions. Safayari et al. [8] proposed the future work direction. The possibility of correct treatment based on patient specific biomarkers is an important step towards personalized psychiatry. Squarcina et al. [9] Machine learning methods are increasingly popular in the medical field. Training that considers the possibilities of individual patients rather than the whole group in the analysis makes them particularly interested in investigating treatment responses. Ji et al. urgently need an objective and effective evaluation method in the existing clinical research. From the perspective of combining voice features with AI technology represented by machine learning and deep learning, they summarized the commonly used voice features in the research of depression recognition based on AI technology, and summarized the relevant research methods for the auxiliary diagnosis and treatment of depression.

Knowledge map (KG) has strong description ability. It is mainly used to describe various entities, concepts and the relationship between entity concepts in the objective world. It is widely used in search result optimization, intelligent recommendation, intelligent customer service, financial risk control, safety and security, intelligent medical and other fields. Digital campus and educational information system have a large number of different structures and dynamic data, which are the basis of educational intelligent management. Knowledge mapping is an effective method to organize and manage these data. KG makes education informatization and medical information system closer to human thinking and cognitive mode. KG can be used for the prediction and intervention of college students' depression and medical question answering system. Medical question answering is another form of medical information retrieval. The answer returned is natural language, which is more professional. Clinical treatment decision support is based on the personal situation of depressed college students, combined with big data analysis in the medical field, to automatically generate treatment plans for each patient and provide them to doctors for reference.

## **2. KG of college students' depression**

The knowledge related to depression is extracted from education big data of different sources and structures, including named entity recognition and relationship extraction. Entity extraction is to

identify and extract named entities, such as drug names and symptoms, from text education big data. Relationship extraction is to extract the relationship between named entities. There are two methods to extract triples from text: one is to identify named entities first, and then extract relationships. The other is to extract entities and relationships at the same time. Because the summary in this research data is mainly around a disease entity, it is only necessary to extract another entity and the relationship R between the two entities. In this study, entities are divided into 7 categories, including disease, drug treatment, symptom, diagnosis, non-drug treatment, epidemiology and others. The relationships between entities formed by them include disease diagnosis, disease drug treatment, disease symptom, disease non drug treatment, disease epidemiology and disease other.

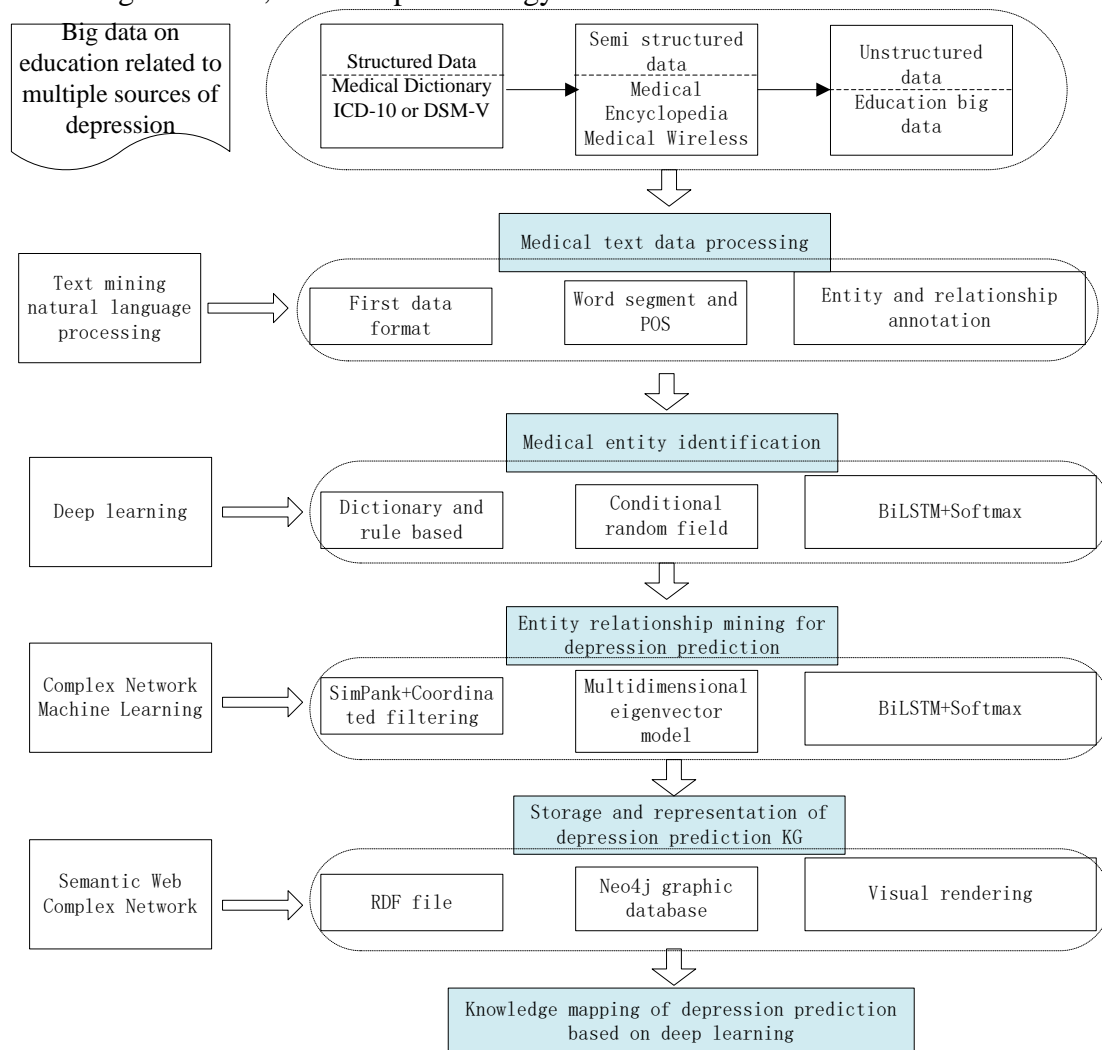


Figure 1: Flow chart for constructing knowledge map of depression prediction based on deep learning

A triple is used to describe facts and is expressed as  $(e_1, r, e_2)$ , where  $e_1$  and  $e_2$  represent entity 1 and entity 2 respectively, and  $r$  represents the relationship between entity 1 and entity 2. For each entity, select the entity themed triple for display. KG consists of multiple entities and relationships between entities. Each entity is identified by a globally unique and determined ID. Each relationship connects two entities. Manually annotate and automatically extract triples. The annotation method uses "relationship name+BIOES", where "R-B" marks the beginning character of the entity, "R-I" marks the middle character of the entity, "R-O" marks irrelevant characters, "R-E" marks the end character of the entity, and "R-S" marks a single character. As shown in Figure 1, when training data,

first manually mark the data in the training set according to the annotation method of "relationship name+BIOES", and then send it to the model for learning. When testing data, first use the BiLSTM+Softmax model to predict the label of each word, then match the label, and finally form a triple (entity, relationship, entity). In this study, 386 entities and 472 relationships were obtained from the paper abstracts using this method. The open source graph database Neo4j is used to store concave KG, because it has strong graph search capabilities and certain horizontal expansion capabilities.

Analyze college students' depression in detail from the education big data. On this basis, a prediction and intervention model of college students' depression is constructed. With the help of a prediction model based on deep learning, college students' depression is diagnosed and whether there is depression is judged. In order to compare the prediction effects of notification intervention and knowledge question answering, students in depression were divided into two groups to implement two intervention strategies. Notification intervention refers to informing students of existing problems by sending e-mail and providing effective auxiliary suggestions: the knowledge question answering support environment intervention group will receive a diagnosis report of depression corresponding to their life and learning performance, and can use the prediction support services provided by the prediction platform, Such as peer-to-peer interaction, self-assessment tools and educational scaffolding. In order to ensure the effective implementation of online intervention strategies, it is proposed to use credit scores and early warning indicators to predict depression and the intervention system to guarantee the effective implementation of prediction and intervention strategies.

### 3. Conclusion

Depression is one of the most common mental disorders, which has a high morbidity and mortality, and seriously affects the health and social functions of the population. In order to explore more efficient and objective diagnosis and treatment methods. College students' depression is a multifactorial disease. It is a kind of mood disorder characterized by significant and lasting depression, which lasts for more than two weeks. It has a significant impact on patients' social functions and is a major disease that seriously harms human mental and physical health. College students' depression has become a common problem faced by the medical system, the education system and even the whole society. The depression KG based on education big data can be applied to the depression question and answer system. In the future, we will continue to use more algorithms to improve the prediction accuracy of college students' depression, study how to improve the accuracy of answers, and more fully add attribute information to the depression KG.

### Acknowledgement

Fund Project: Shaanxi Higher Education Teaching Reform Research Project (21BG051), Shaanxi Education Science "Fourteenth Five Year Plan" Project (SGH21Y0283 and SGH21Y0279), Ministry of Education's Cooperative Education Project (Project No. 220603414290713), 2021 Undergraduate Teaching Reform Research Project of Xijing University (JGYB2117), Xijing University Teacher Education Reform and Teacher Development Research Project (JFGG202203).

### References

- [1] Qu Jianxin. *Computational Psychiatry: A New Perspective of Depression Research and Clinical Application* [J]. *Progress in Psychological Science*, 2020, 28 (1): 111-127.
- [2] Ji Guangjun, Wang Ning, Wei Dongshuai, et al. *Research progress of depression recognition based on artificial intelligence and speech features* [J]. *Journal of Clinical Psychiatry*, 2022,32 (5): 15-417.
- [3] Han Jiali, Feng Lei. *Research progress in the application of artificial intelligence in the field of depression* [J]. *Beijing*

*Medical Journal*, 2020, 42 (4): 317-322.

[4] Yuan Qinmei, Wang Xing, Shuai Jianwei, et al. Research progress of depression based on artificial intelligence technology [J]. *Chinese Journal of Clinical Psychology*, 2020, 28 (1): 82+87.

[5] Zha Meng, Ye Ning, Wang Ruchuan, et al. Research on depression prediction based on capsule network model [J]. *Computer Technology and Development*, 2021, 31 (11): 28-34.

[6] Kim J, Hong J, Choi Y. Automatic Depression Prediction using Screen Lock/Unlock Data on the Smartphone[C]. *18th International Conference on Ubiquitous Robots (UR)*, 2021.

[7] Dai Z, Zhou H, Ba Q, et al. Improving depression prediction using a novel feature selection algorithm coupled with context-aware analysis [J]. *Journal of affective disorders*, 2021, 295:1040-1048.

[8] Safayari A, Bolhasani H. Depression diagnosis by deep learning using EEG signals: A Systematic Review [J]. *New technologies and equipment in medicine (English)*, 2021 (4): 16. DOI: 10.20944/reprints202107.0028.v1.

[9] Squarcina L, Villa F M, Nobile M, et al. Deep learning for the prediction of treatment response in depression[J]. *Journal of affective disorders*, 2021, 281: 618-622.