Weakly Supervised Neural Representation Learning through Exploiting Expert Knowledge

Xu Guo[1], Han Yu[1], Yiqiang Chen[2] and Chunyan Miao[1]

[1]School of Computer Science and Engineering, NTU
[2]Institute of Computing Technology, CAS

{xu008, han.yu}@ntu.edu.sg, yqchen@ict.ac.cn, ascymiao}@ntu.edu.sg

Abstract

Chronic pain is a condition that negatively affects people's quality of life and incurs significant burden on public healthcare resources. Currently, the lack of systematic pain management training and support among primary care doctors limits their ability to provide quality care for chronic patients. With the advent of deep learning, the integration of Artificial Intelligence (AI) and Medicine has inspired a range of research on Electronic Health Records (EHRs). To empower non-specialist primary care doctors to provide better care for chronic pain patients, we develop a decision support system leveraging EHR data. The core AI engine of this system is a discriminator-based autoencoder neural network that can achieve good representation performance on small scale datasets to help primary care doctors search for similar chronic pain cases previously treated by specialists. The system has been beta-tested on a dataset collected from 440 chronic pain patients together with doctors from Tan Tock Seng Hospital, Singapore, and has shown promising results. In this way, we enable AI to help non-specialists triage patient conditions and formulate personalized treatment plans.

Keyword: Electronic Health Records, Artificial Intelligence.

I. Introduction

Chronic pain is affecting a growing number of people around the world. It is estimated that 20% of adults worldwide suffer from pain at any given moment, and 10% are newly diagnosed with chronic pain each year [4]. Statistics show that chronic pain is the leading cause of long-term disability in the United States, contributing to an estimated $560 billion each year in direct medical costs, lost productivity, and disability programs [3]. Inadequate pain management negatively affects the patients' welfare and incur high cost on public healthcare resources due to extended hospitalization and readmission [19]. However, chronic pain is hard to cope with due to the diversity of causes and symptoms. Ideally, multidisciplinary interventions are required in order to effectively manage chronic pain [13].

Nevertheless, in practice, chronic pain is often viewed in isolation and managed with a static approach. It is recommended by the National Institutes of Health (NIH) that the treatment of chronic pain should combine non-pharmacologic options including physical therapy [12]. Physical therapy, such as exercising, is a useful adjunctive modality of pain relief which enhances the effectiveness of other elements in the treatment geared towards resolution of movement impairments and restoration of physical functions [1,6]. Through tailored exercises, physical therapy can help patients reduce pain by building muscular-skeletal strength and increasing mobility.

Pain patients often seek help first from primary care doctors, who will then help them make appointments with pain specialists if necessary. However, as reported in the NIH Pathways to Prevention Workshop on chronic pain [12], primary care doctors are often ill-prepared to appropriately assess, treat, and monitor patients with chronic pain as a result of the lack of knowledge in the subject, limited availability of treatment options, and resource constraints. There is an urgent need for technology-empowered decision support to enable primary care doctors to provide high quality care for chronic pain patients.

Using electronic health records data, machine learning algorithms have been applied to analyze many important clinical parameters, ranging from Alzheimer's disease to death. As pointed out in a recent article in Nature Medicine [17], almost all clinicians, ranging from specialty doctor to paramedic, will be using AI technology in the future. Unlike the conventional decision support systems, which relied on the curation of medical knowledge base and on the explicit expressions of decision rules, recent AI technology can account for complex interactions, to identify patterns from data [22].

Pain specialists have left their thought processes in some form in the diagnosis and treatment plans in the Electronic Health Records (EHRs) of previous patients. In this work, we aim to use AI to extract this knowledge, learn from it, and advice non-specialist primary care doctors so that they can serve pain patients better and alleviate the burden on pain specialists. More specifically, we aim to design an AI system that can tap into the expertise of pain specialists to

1. help primary care doctors to triage chronic pain symptoms at an early stage; and
2. advise doctors to develop effective intervention strategies to manage chronic pain.

In order to achieve this goal, we have developed the Agent-based Clinical Decision Support System (ACDSS) [5]. It is designed to perform 1) risk prediction for estimating the probability that the patient may develop chronic pain; 2) general pain assessment based on patient classification and patient similarity measurement; and 3) recommendation of exercise therapy together with the types of drugs to be used based on the assessment results which have been checked by the doctor. These functions require an efficient patient representation methodology. However, research on AI-powered decision support using both structured and unstructured free text medical records remains scarce [2,21,11]. Thus, at the core of our AI Engine, we incorporated our proposed Discriminator-based Autoencoder Neural Network (DANN) for patient representation followed by heuristic search to leverage knowledge from related cases.

The system has been beta-tested on a dataset collected from 440 chronic pain patients together with doctors from Tan Tock Seng Hospital, Singapore, and has shown promising results. In this way, we enable AI to help non-specialists triage patient conditions and formulate treatment plans.

II. The Proposed DANN model for Pain Patients

The Discriminator-augmented Autoencoder Neural Network (DANN), is an extension of the autoencoder neural network, for improving discriminative representation learning. The architecture of DANN, as shown in Figure 1, comprises a multi-layer autoencoder and a discriminator, which is a binary classifier. While training the autoencoder to achieve dimensionality reduction, DANN simultaneously perform coarse-grained classification (i.e., recognizing chronic pain) to induce the feature learning process. The meaning of this architecture is to enable the representation model to be able to represent chronic pain patients differently with non-chronic pain patients, which often require different subsequent interventions.

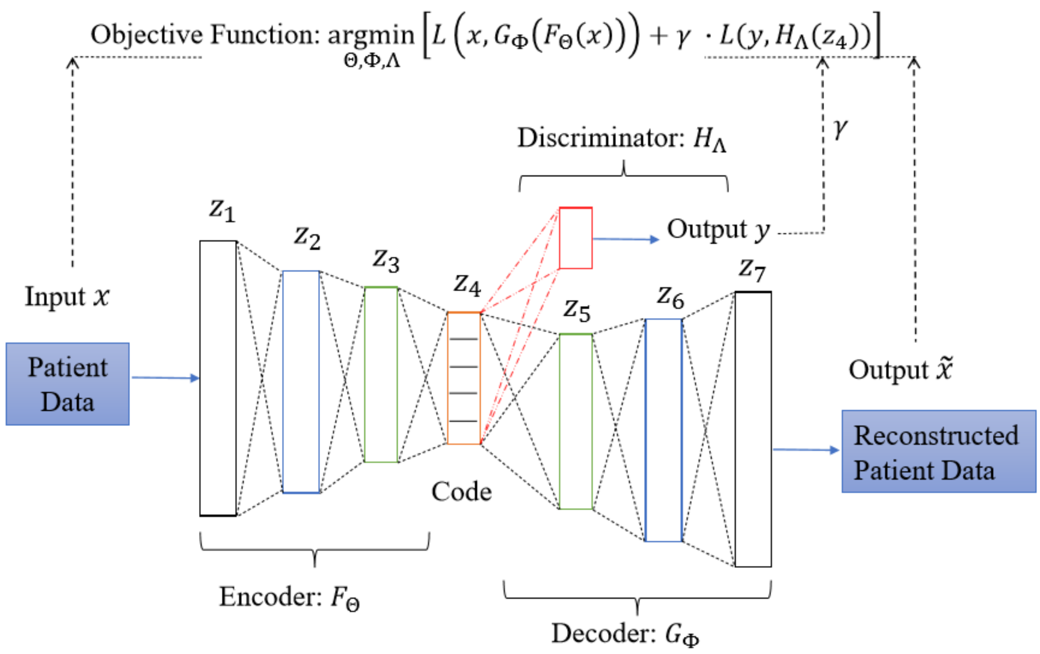

Figure 1. Weakly supervised neural representation learning model for pain patients.

Let $D = \{X_j | X_j \in \mathbb{R}^n, 1 < j < N\}$ denote the entire dataset, where X_j represents the j-th sample and N is the total number of samples. We split D into $L(L > 1)$ independent subsets according to domain specific heuristics, denoted as $\{D_k\}_{k=1}^{L}$, where $D_k \subseteq D$, $\cap_{k=1}^{L} D_k = \emptyset$ and $\cup_{k=1}^{L} D_k = D$. As our aim is to predict a fine-grained target such as certain therapy as accurately as possible, we exploit expert knowledge to annotate those samples to be either chronic or non-chronic samples for the discriminator to learn to discriminate among them. In this case, we can split the dataset into two parts (i.e., $L = 2$) for DANN to learn the discriminative feature representation.

We denote the hypothesis function of the encoder as F_Θ and that of the decoder as G_Φ. We define $\sigma(W_l \cdot z_{l-1} + b_l) = z_l$, where $\sigma(\cdot)$ is the hyperbolic tangent activation function, $\Theta = [W_l, b_l]_{l=1}^{m}$ and $\Phi = [W_l, b_l]_{l=m+1}^{2m}$ are the parameter matrices of encoder and decoder, respectively. $z_0 = x$, the input sample, $z_{2m} = \tilde{x}$, the reconstructed sample, and z_m refer to the *code* or middle-most latent representation in the autoencoder.

The k-th discriminator neural network is defined as $H_{\Lambda_k}: \phi(W_k \cdot z_m + b_k) = r_k$, where r_k represents the likelihood of predicting the label, $y_k \in \{0,1\}$. Here, $y_k = 1$ indicates that the latent vector z_m is generated from the k-th subset D_k; otherwise, $y_k = 0$. $\phi(\cdot)$ is a sigmoid activation function. $\Lambda_k = [W_k, b_k]$ is the parameter matrix of the k-th discriminator. Figure 1 shows a special case of $k = 1$.

We adopt the mean squared error to measure the reconstruction loss, L_1, of autoencoder and the classification loss, L_2, of discriminator. The objective function of DANN is:

$$\arg\min_{\Theta,\Phi,\Lambda} \left[L_1(x, G_\Phi(x)) + \sum_k \gamma_k \cdot L_2(y_k, H_{\Lambda_k}(z_m)) \right]$$

where γ_k is the weight parameter to control how much contribution of the k-th discriminator classification error is used to train the autoencoder.

III. Federated DANN for Heterogeneously Fused Pain Data

DANN is proposed to solve the problem that data instances belonging to different classes may appear similar, which is common in the scenario of pain management analytics. While in the scenario when patients data are fused from different sources and the data types are difficult to align to the same space, we propose to solve it based on a federated DANN model. We assume that a patient's heterogenous data sources are preserved in their own feature space, and feed to separate DANNs for representation learning. To align them to the same prediction task (e.g., pain sites prediction and therapy recommendation), additional neural layers are applied to map them into the same feature space. For each DANN, its discriminator could focus on different expert knowledge depending on the subdomain. As a result, the federated version is supposed to be able to transform the high-dimensional heterogeneous patient data to an aligned task-aware feature space, which preserve the differential characteristics from data itself and is portable for downstream applications such as searching for similar cases for personalized recommendations.

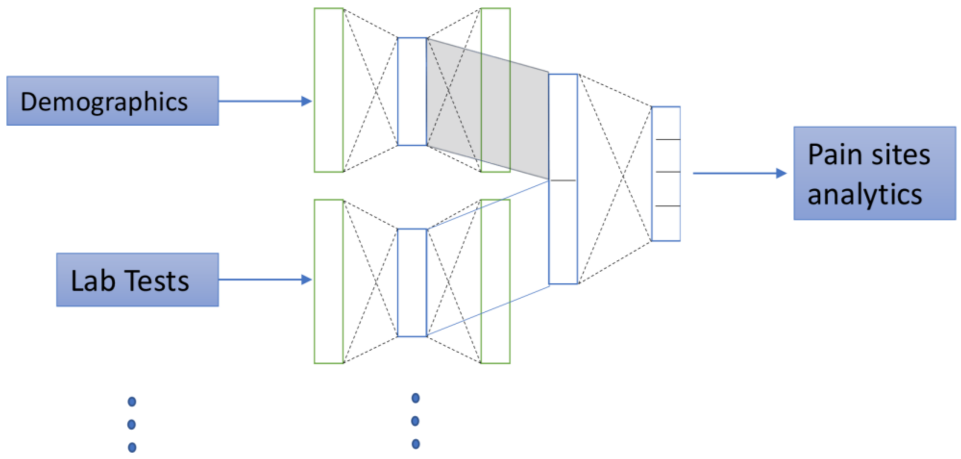

Figure 2. Federated version for heterogeneously fused pain patient representation.

IV. Conclusion

In this paper, we report on our experience of empowering pain patient representation by exploiting expert knowledge through a discriminator-based autoencoder to boost downstream tasks in pain management settings. The proposed weakly supervised neural representation model is potential to be applied in retrieval task to help primary care doctors leverage knowledge embedded in previous cases to provide quality care to pain patients. It is also extendable to a more challenging setting where multi-source fused patient data are aligned to enable fine-grained decision-makings.

In subsequent work, we will explore the application of federated learning [20] to enable the DANN function in a secure and distributed manner without the need for data aggregation so as to protect user privacy and ease doctors' concerns.

Acknowledgements

This research is supported by Nanyang Technological University, Nanyang Assistant Professorship (NAP) and the Singapore Ministry of Education Academic Research Fund Tier 1 (Grant No. 2017-T1-001-270).

References

[1] Allen, R. J. (2006). Physical agents used in the management of chronic pain by physical therapists. Physical medicine and rehabilitation clinics of North America, 17, 315-345.

[2] Culliton, P., Levinson, M., Ehresman, A., Wherry, J., Steingrub, J. S., & Gallant, S. I. (2017). Predicting severe sepsis using text from the electronic health record. arXiv preprint arXiv:1711.11536.

[3] Dahlhamer, J., Lucas, J., Zelaya, C., Nahin, R., Mackey, S., DeBar, L., . . . Helmick, C. (2018). Prevalence of chronic pain and high-impact chronic pain among adults—United States, 2016. Morbidity and Mortality Weekly Report, 67, 1001.

[4] Goldberg, D. S., & McGee, S. J. (2011). Pain as a global public health priority. BMC public health, 11, 770.

[5] Guo, X., Yu, H., Miao, C., & Chen, Y. (2019). Agent-based Decision Support for Pain Management in Primary Care Settings. Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19), (pp. 6587-6589).

[6] Jansen, M. J., Viechtbauer, W., Lenssen, A. F., Hendriks, E. J., & Bie, R. A. (2011). Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: a systematic review. Journal of physiotherapy, 57, 11-20.

[7] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[8] Loria, S. (2018). textblob Documentation. Tech. rep., Technical report.

[9] Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Scientific reports, 6, 26094.

[10] Oh, J., Makar, M., Fusco, C., McCaffrey, R., Rao, K., Ryan, E. E., . . . others. (2018). A generalizable, data-driven approach to predict daily risk of Clostridium difficile infection at two large academic health centers. infection control & hospital epidemiology, 39, 425-433.

[11] Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., . . . others. (2018). Scalable and accurate deep learning with electronic health records. NPJ Digital Medicine, 1, 18.

[12] Reuben, D. B., Alvanzo, A. A., Ashikaga, T., Bogat, G. A., Callahan, C. M., Ruffing, V., & Steffens, D. C. (2015). National Institutes of Health Pathways to Prevention Workshop: the role of opioids in the treatment of chronic pain. Annals of internal medicine, 162, 295-300.

[13] Scascighini, L., Toma, V., Dober-Spielmann, S., & Sprott, H. (2008). Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes. Rheumatology, 47, 670-678.

[14] Shah, N. H., Bhatia, N., Jonquet, C., Rubin, D., Chiang, A. P., & Musen, M. A. (2009). Comparison of concept recognizers for building the Open Biomedical Annotator. BMC bioinformatics, 10, p. S14.

[15] Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2017). Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE journal of biomedical and health informatics, 22, 1589-1604.

[16] Si, Y., & Roberts, K. (2019). Deep Patient Representation of Clinical Notes via Multi-Task Learning for Mortality Prediction. AMIA Summits on Translational Science Proceedings, 2019, 779.

[17] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature medicine, 25, 44.

[18] Voigt, P., & Bussche, A. v. (2017). The EU General Data Protection Regulation (GDPR): A Practical Guide (1st ed.). Springer Publishing Company, Incorporated.

[19] Wells, N., Pasero, C., & McCaffery, M. (2008). Improving the quality of care through pain assessment and management. In Patient safety and quality: An evidence-based handbook for nurses. Agency for Healthcare Research and Quality (US).

[20] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated Machine Learning: Concept and Applications. ACM Transactions on Intelligent Systems and Technology, 10, 12:1--12:19.

[21] Yang, Z., Huang, Y., Jiang, Y., Sun, Y., Zhang, Y.-J., & Luo, P. (2018). Clinical assistant diagnosis for electronic medical record based on convolutional neural network. Scientific reports, 8, 6329.

[22] Yu, K.-H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. Nature biomedical engineering, 2, 719.