AI and Personal Healthcare: A literature review

Zhiqing Wang

September 2023

1 Introduction

The AI Era has come. Actually, for decades, but with the stunning appearance of ChatGPT and all kinds of novel applications based on Large Language Models(LLMs), deep learning models, and neural networks(etc.), it became a widely accepted fact to everyone, not just professionals. Artificial Intelligence is bringing about a revolution in many scientific research fields. With the fantastic big data processing ability and Natural Language Processing(NLP) features of the AI models, AI has become more than just a co-pilot to human experts. In particular tasks, they show even better performance in terms of prediction accuracy and working speed. In this literature review, we will focus on AI applications, particularly in the fields of personal healthcare, and disease prediction/diagnosis.

2 Diease Prediction based on EHR

Electronic health records(EHR) is a digital version of the patient's traditional paper chart. Which contains a comprehensive range of data including medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results. The massive amount of EHR makes it a perfect input data source for disease prediction powered by Artificial intelligence. There are three existing models that were designed for processing the EHR: BEHRT[3], G-BERT[6] and Med-BERT[5]. Each of these models focuses on a different purpose but shares some common underlying features.

2.1 G-BERT[6]

The G-BERT model was designed for medication recommendation tasks. It was the very first pre-training language model that had been utilized in the healthcare domain. G-BERT has applied the pre-training technique BERT [1] [Devlin et al., 2018] to leverage more training data that was discarded from the training set due to the un-preferred features such as patients with single-visit only. To resolve the absence of a hierarchy of medical knowledge, the model used graph neural networks(GNNs) to derive the ontology embedding of the medical codes ICD9 and ATC.

2.2 BEHRT[3]

BEHRT is a transformer model designed to give predictions of the patient's risk of diagnosing a disease at a certain age by using the EHR as the input data source. This model was also pre-trained using the Masked language model(MLM), which helps the model to focus on the disease track and avoid other noise information contained in the natural language in EHR. BEHRT integrates both multi-head self-attention and positional encoding mechanisms. Its bidirectional attributes further enable the identification of non-directional correspondences among various diseases. Resulting in a prediction range of more than 300 diseases.

2.3 MED-BERT[5]

Med-BERT, while fundamentally also based on the BERT architecture, is adapted specifically for Electronic Health Records (EHR) for disease predictions. What makes it different from the previously discussed models are its three unique embeddings: Code Embedding, Serialized Embedding, and Visit Embedding. The integration of these specialized embeddings empowers Med-BERT with the capacity to manage EHR sequences encapsulating multiple visits, negating the need for separator tokens [SEP] and summarization tokens like [CLS]. As a result, this refined structure provides a more efficient representation, minimizing the loss of critical information and curtailing the presence of superfluous tokens.

Table 1. Comparison of Med-BERT with BEHRT and G-BERT from multiple perspectives.

Criteria	BEHRT	G-BERT	Med-BERT
Type of input code	Caliber code for diagnosis developed by a college in London	Selected ICD-9 code for diagnosis + ATC code for medication	ICD-9 + ICD-10 code for diagnosis
Vocabulary size	301	<4 K	82 K
Pretraining data source	CPRD (primary care data)[63]	MIMIC III (ICU data)[64]	Cerner Health Facts (general EHR)
Input structure	Code + visit + age embeddings	Code embeddings from ontology + visit embeddings	Code + visit + code serialization embeddings
Pretraining sample unit	Patient's visit sequence	Single visit	Patient's visit sequence
Total number of pretraining patients	1.6 M	20 K	20 M
Average number of visits for each patient for pretraining	Not reported but >5	<2	8
Pretraining task	Masked LM	Modified Masked LM	Masked LM + prediction of prolonged length of stay in hospital
Evaluation task	Diagnosis code prediction in different time windows	Medication code prediction	Disease predictions according to strict inclusion/exclusion criteria
Total number of patients in evaluation tasks	699 K, 391 K, and 342 K for different time windows	7 K	50 K, 20 K, and 20 K for three task cohorts

Figure 1: A comparison table from [5]

3 Mental Health Support Chatbots

In the fast-paced modern life, especially after COVID-19, the demand for mental health support and intervention has rapidly grown. Traditionally confined to therapist's offices and clinical settings, mental health care has gradually shifted into the digital domain, with applications like Woebot[2] and Saarthi[4] emerging at the forefront. These innovative platforms not only make psychological assistance more accessible but also seamlessly integrate into the daily lives of users, thereby popularizing mental health support. But behind the user-friendly interfaces of these applications lies the power of Artificial Intelligence (AI), Large Language Models (LLMs), and advanced neural networks that form the backbone of these chatbots, enabling them to offer personalized guidance and feedback. Although, due to the nature of AI, the chatbot essentially lacks empathy, and is usually not compatible with emergencies such as saviour panic attacks or suicide in progress. And it is not a substitute for a professional mental health doctor, who cannot diagnose the patient but only provides possibilities and suggestions.

4 A section written by ChatGPT regarding the three BERT-based models

As of my last update in September 2021, there has been growing interest in adapting the BERT (Bidirectional Encoder Representations from Transformers) model to various domains, including healthcare. Given the sensitive and specialized nature of medical data, domain-specific models can be crucial for improving performance in healthcare-related tasks. Let's dive into each of the models you mentioned:

1. **MED-BERT**: - **Overview**: MED-BERT is an adaptation of the original BERT model, tailored for medical data. It is pretrained on medical corpora to better capture medical terminologies, relationships, and context. - **Applications**: MED-BERT can be utilized for several tasks in the medical domain, including medical text classification, disease prediction based on clinical notes, and medical entity recognition, among others. - **Advantages**: Since MED-BERT is trained on medical data, it is expected to better understand the context and semantics of clinical text, leading to more accurate results for healthcare-specific tasks compared to a generic BERT model.

2. **G-BERT**: - **Overview**: As of my last update, G-BERT might not be a widely-recognized variant specifically tailored for healthcare. However, there might have been developments or papers post-2021 that introduced G-BERT in the medical domain. - **Applications and Advantages**: Specific applications and advantages would be based on the precise nature and training of G-BERT, which would require access to recent publications or resources for accurate information.

3. **BEHRT**: - **Overview**: BEHRT (Bidirectional Encoder Representations from Transformers for Healthcare) is another adaptation of BERT

tailored for healthcare. It's designed to work with electronic health records (EHRs). Unlike free-text clinical notes, EHRs often contain structured data, and BEHRT is optimized to work with this kind of information. - **Applications**: BEHRT can be used for a range of EHR-based tasks like predicting disease onset, treatment outcomes, patient stratification, and more. - **Advantages**: EHRs contain vast amounts of patient information but are structured differently than traditional texts. BEHRT's training allows it to capture the nuances and relationships within EHRs effectively.

General Consideration for Healthcare BERT Models: All these models, when designed for healthcare, need to be used with caution. Healthcare data is sensitive, and any predictions or inferences drawn from it can have direct consequences on patient care. Additionally, issues related to data privacy, consent, and bias in training data need to be rigorously addressed.

In conclusion, domain-specific models like MED-BERT and BEHRT exemplify the potential of transfer learning in healthcare. By starting with a generic model and fine-tuning it on specialized data, researchers can harness the power of large-scale models while catering to the unique requirements of the medical field.(ChatGPT, 2023)

5 Discussion

ChatGPT has definitely generated a neatly structured section to show its understanding of the mentioned models. It managed to generate a generally correct answer to every model. The key point should be it was only using the most general keywords that were mentioned in each paper.

References

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[2] Fitzpatrick. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial. *JMIR Ment Health*, 4(2):e19, Jun 2017.

[3] Yikuan Li, Shishir Rao, José Roberto Ayala Solares, Abdelaali Hassaine, Rema Ramakrishnan, Dexter Canoy, Yajie Zhu, Kazem Rahimi, and Gholamreza Salimi-Khorshidi. BEHRT: Transformer for electronic health records. *Scientific Reports*, 10(1), apr 2020.

[4] Komal Rani, Harshit Vishnoi, and Manas Mishra. A mental health chatbot delivering cognitive behavior therapy and remote health monitoring using nlp and ai. In *2023 International Conference on Disruptive Technologies (ICDT)*, pages 313–317, 2023.

[5] Laila Rasmy, Yang Xiang, Ziqian Xie, Cui Tao, and Degui Zhi. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *npj Digital Medicine*, 4(1), may 2021.

[6] Junyuan Shang, Tengfei Ma, Cao Xiao, and Jimeng Sun. Pre-training of graph augmented transformers for medication recommendation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 5953–5959. International Joint Conferences on Artificial Intelligence Organization, 7 2019.