

# Task-oriented Dialogue System for Automatic Diagnosis

Qianlong Liu<sup>1</sup>, Zhongyu Wei<sup>1\*</sup>, Baolin Peng<sup>2</sup>, Xiangying Dai<sup>3</sup>,  
Huaixiao Tou<sup>1</sup>, Ting Chen<sup>1</sup>, Xuanjing Huang<sup>4</sup>, Kam-fai Wong<sup>2,5</sup>

<sup>1</sup>School of Data Science, Fudan University, China

<sup>2</sup>The Chinese University of Hong Kong, Hong Kong

<sup>3</sup>Baidu Inc., China

<sup>4</sup>School of Computer Science, Fudan University, China

<sup>5</sup>MoE Key Lab of High Confidence Software Technologies, China

{17210980040, zywei, 17210980013, 17210980029}@fudan.edu.cn

{blpeng, kfwong}@se.cuhk.edu.hk

daixiangying@baidu.com

xjhuang@fudan.edu.cn

## Abstract

In this paper, we make a move to build a dialogue system for automatic diagnosis. We first build a dataset collected from an online medical forum by extracting symptoms from both patients' self-reports and conversational data between patients and doctors. Then we propose a task-oriented dialogue system framework to make the diagnosis for patients automatically, which can converse with patients to collect additional symptoms beyond their self-reports. Experimental results on our dataset show that additional symptoms extracted from conversation can greatly improve the accuracy for disease identification and our dialogue system is able to collect these symptoms automatically and make a better diagnosis.

## 1 Introduction

Automatic phenotype identification using electronic health records (EHRs) has been a rising topic in recent years (Shivade et al., 2013). Researchers explore with various machine learning approaches to identify symptoms and diseases for patients given multiple types of information (both numerical data and pure texts). Experimental results prove the effectiveness of the identification of heart failure (Jonnalagadda et al., 2017; Choi et al., 2016), type 2 diabetes (Li et al., 2015; Zheng et al., 2017), autism spectrum disorders (Doshi-Velez et al., 2014), infection detection (Tou et al., 2018) etc. Currently, most attempts focus on some

specific types of diseases and it is difficult to transfer models from one disease to another.

In general, each EHR contains multiple types of data, including personal information, admission note, diagnose tests, vital signs and medical image. And it is collected accumulatively following a diagnostic procedure in clinic, which involves interactions between patients and doctors and some complicated medical tests. Therefore, it is very expensive to collect EHRs for different diseases. How to collect the information from patient automatically remains the challenge for automatic diagnosis.

Recently, due to its promising potentials and alluring commercial values, research about task-oriented dialogue system (DS) has attracted increasing attention in different domains, including ticket booking (Li et al., 2017; Peng et al., 2017a), online shopping (Yan et al., 2017) and restaurant searching (Wen et al., 2017). We believe that applying DS in the medical domain has great potential to reduce the cost of collecting data from patients.

However, there is a gap to fill for applying DS in disease identification. There are basically two major challenges. First, the lack of annotated medical dialogue dataset. Second, no available DS framework for disease identification. By addressing these two problems, we make the first move to build a dialogue system facilitating automatic information collection and diagnosis making for medical domain. Contributions are two-fold:

- We annotate the first medical dataset for dialogue system that consists of two parts, one is self-reports from patients and the other is conversational data between patients and

\* Corresponding author

doctors.

- We propose a reinforcement learning based framework for medical DS. Experiment results on our dataset show that our dialogue system is able to collect symptoms from patients via conversation and improve the accuracy for automatic diagnosis.

## 2 Dataset for Medical DS

Our dataset is collected from the pediatric department in a Chinese online healthcare community<sup>1</sup>. It is a popular website for users to inquire with doctors online. Usually, a patient would provide a piece of self-report presenting his/her basic conditions. Then a doctor will initialize a conversation to collect more information and make a diagnosis based on both the self-report and the conversational data. An example is shown in Table 1. As we can see, the doctor can obtain additional symptoms during conversation beyond the self-report. For each patient, we can also obtain the final diagnosis from doctors as the label. For clarity, we term symptoms from self-reports as *explicit symptoms* while those from conversational data as *implicit symptoms*.

We choose four types of diseases for annotation, including upper respiratory infection, children functional dyspepsia, infantile diarrhea and children’s bronchitis. We invite three annotators (one with medical background) to label all the symptom phrases in both self-reports and conversational data. The annotation is performed in two steps, namely symptom extraction and symptom normalization.

**Symptom Extraction** We follow the BIO (begin-in-out) schema for symptom identification (Figure 1). Each Chinese character is assigned a label of "B", "I" or "O". Also, each extracted symptom expression is tagged with *True* or *False* indicating whether the patient suffers from this symptom or not. In order to improve the annotation agreement between annotators, we create two guidelines for the self-report and the conversational data respectively. Each record is annotated by at least two annotators. Any inconsistency would be further judged by the third one. The Cohen’s kappa coefficient between two annotators are 71% and 67% for self-reports and conversations respectively.

<sup>1</sup><http://muzhi.baidu.com>

Self-report	
宝宝嗓子有痰，腹泄并伴有拉水的症状。请问要吃什么药？	The little baby get sputum in throat and have watery diarrhea. what kind of medicine needs to be taken?
Conversation	
.....	
Doctor: 宝宝现在咳嗽拉肚子吗？	Does the baby have a cough or diarrhea now?
Patient: 不咳嗽，拉肚子。	No cough, but diarrhea.
Doctor: 平常呛奶吗？	Does the baby choking milk?
Patient: 偶尔会吐奶。	He vomits milk sometimes.
.....	

Table 1: An example of a user record. Each record consists of two parts: self-report from the patient and the conversation between the doctor and the patient. Underlined phrases are symptom expressions.

Extracted symptom expression	Related concept in SNOMED CT
咳嗽(cough)	咳嗽(cough)
喷嚏(sneez)	打喷嚏(sneezing)
鼻涕(cnot)	鼻涕(cnot)
拉肚子(have loose bowels)	腹泻(diarrhea)
温度37.5-37.7之间(body temperature between 37.5-37.7)	低热(low-grade fever)

Table 2: Examples of extracted symptom expressions and the related concepts in SNOMED CT.

**Symptom Normalization** After symptom expression identification, medical experts manually link each symptom expression to the most relevant concept on SNOMED CT<sup>2</sup> for normalization. Table 2 shows some phrases that describe symptoms in the example and some related concepts in SNOMED CT. The overview of dataset is presented in Table 3.

After symptom extraction and normalization, there are 144 unique symptoms identified. In order to reduce the size of action space of the DS, only 67 symptoms with a frequency greater than or equal to 10 are kept. Samples are then generated, called *user goal*. As we know, each user goal (see Figure 2) is derived from one real world patient record<sup>3</sup>.

## 3 Proposed Framework

A task-oriented DS typically contains three components, namely Natural Language Understanding (NLU), Dialogue Manager (DM) and Natural Language Generation (NLG). NLU detects the user intent and slots with values from utterances; DM

<sup>2</sup><https://www.snomed.org/snomed-ct>

<sup>3</sup>The dataset is available at: [www.sdspeople.fudan.edu.cn/zywei/data/acl2018-mds.zip](http://www.sdspeople.fudan.edu.cn/zywei/data/acl2018-mds.zip)

宝宝嗓子有痰，腹泻并伴有拉水的症状。……  
 O O B I I O B I O O O B I O O O O ……  
 The baby get sputum in throat and have watery diarrhea. ……

Figure 1: An example utterance with annotations of symptoms in BIO format.

Disease	# of user goal	Ave # of explicit symptoms	Ave # of implicit symptoms
infantile diarrhea	200	2.15	2.71
children functional dyspepsia	150	1.70	3.20
upper respiratory infection	160	2.56	3.55
children's bronchitis	200	2.87	3.64

Table 3: Overview of the dataset. # of user goal is the number of dialogue sessions of each disease, Ave # of explicit symptoms and Ave # of implicit symptoms are the average number of explicit and implicit symptoms among user goals respectively.

tracks the dialogue states and takes system actions; NLG generates natural language given the system actions. In this work, we focus on the DM for automatic diagnosis consisting of two sub-modules, namely, dialogue state tracker (DST) and policy learning. Both NLU and NLG are implemented with template-based models. Typically, a user simulator is designed to interact with the dialogue system (Liu et al., 2017; Peng et al., 2017b; Su et al., 2016; Schatzmann et al., 2006). We follow the same setting as Li et al. (2017) to design our medical DS. At the beginning of a dialogue session, the user simulator samples a user goal (see Figure 2), while the agent attempts to make a diagnosis for the user. The system will learn to select the best response action at each time step by maximizing a long term reward.

### 3.1 User Simulator

At the beginning of each dialogue session, a user simulator samples a user goal from the experiment dataset. At each turn  $t$ , the user takes an action  $a_{u,t}$  according to the current user state  $s_{u,t}$  and the previous agent action  $a_{t-1}$ , and transits into the next user state  $s_{u,t+1}$ . In practice, the user state  $s_u$  is factored into an agenda  $A$  (Schatzmann et al., 2007) and a goal  $G$ , noted as  $s_u = (A, G)$ . During the course of the dialogue, the goal  $G$  ensures that the user behaves in a consistent, goal-oriented manner. And the agenda contains a list of symptoms and their status (whether or not they are requested) to track the progress of the conversation.

Every dialogue session is initiated by the user

```
{
  "disease_tag": "小儿支气管炎(children's bronchitis)",
  "request_slots": {
    "disease": "unknow"
  },
  "explicit_symptoms": {
    "咳嗽(cough)": true,
    "鼻流涕(snot)": true
  },
  "implicit_symptoms": {
    "咽痛(sore throat)": true,
    "发烧( fever)": true,
    "粗糙呼吸音(harsh breath sounds)": false,
    "呕吐( emesis)": false
  }
}
```

Figure 2: An example of user goal. Each user goal consists of four parts, *disease\_tag* is the disease that the user suffers; *explicit\_symptoms* are symptoms extracted from the user self-report; *implicit\_symptoms* are symptoms extracted from the conversational data between the patient and the doctor; *request\_slots* is the *disease* slot that the user would request.

via the user action  $a_{u,1}$  which consists of the requested *disease* slot and all explicit symptoms. In terms of the symptom requested by the agent during the course of the dialogue, the user will take one of the three actions including *True* (if the symptom is positive), *False* (if the symptom is negative), and *not\_sure* (if the symptom is not mentioned in the user goal). If the agent informs correct disease, the dialogue session will be terminated as successful by the user. Otherwise, the dialogue session will be recognized as failed if the agent makes incorrect diagnosis or the dialogue turn reaches the maximum dialogue turn  $T$ .

### 3.2 Dialogue Policy Learning

**Markov Decision Process Formulation for Automatic Diagnosis** We cast DS as Markov Decision Process (MDP) (Young et al., 2013) and train the dialogue policy via reinforcement learning (Cuayahuitl et al., 2015). An MDP is composed of states, actions, rewards, policy, and transitions.

**State  $S$ .** A dialogue state  $s$  includes symptoms requested by the agent and informed by the user till the current time  $t$ , the previous action of the user, the previous action of the agent and the turn information. In terms of the representation vector of symptoms, it's dimension is equal to the number of all symptoms, whose elements for positive symptoms are 1, negative symptoms are -1, not-sure symptoms are -2 and not-mentioned symp-

toms are 0. Each state  $s \in S$  is the concatenation of these four vectors.

**Actions  $A$ .** An action  $a \in A$  is composed of a dialogue act (e.g., *inform*, *request*, *deny* and *confirm*) and a slot (i.e., normalized symptoms or a special slot *disease*). In addition, *thanks* and *close\_dialogue* are also two actions.

**Transition  $T$ .** The transition from  $s_t$  to  $s_{t+1}$  is the updating of state  $s_t$  based on the agent action  $a_t$ , the previous user action  $a_{u,t-1}$  and the step time  $t$ .

**Reward  $R$ .** The reward  $r_{t+1} = R(s_t, a_t)$  is the immediate reward at step time  $t$  after taking the action  $a_t$ , also known as reinforcement.

**Policy  $\pi$ .** The policy describes the behaviors of an agent, which takes the state  $s_t$  as input and outputs the probability distribution over all possible actions  $\pi(a_t|s_t)$ .

**Learning with DQN** In this paper, the policy is parameterized with a deep Q-network (DQN) (Mnih et al., 2015), which takes the state  $s_t$  as input and outputs  $Q(s_t, a; \theta)$  for all actions  $a$ . A Q-network can be trained by updating the parameters  $\theta_i$  at iteration  $i$  to reduce the mean squared error between the Q-value computed from the current network  $Q(s, a|\theta_i)$  and the Q-value obtained from the Bellman equation  $y_i = r + \gamma \max_{a'} Q(s', a'|\theta_i^-)$ , where  $Q(s', a'|\theta_i^-)$  is the target network with parameters  $\theta_i^-$  from some previous iteration. In practice, the behavior distribution is often selected by an  $\epsilon$ -greedy policy that takes an action  $a = \arg \max_{a'} Q(s_t, a'; \theta)$  with probability  $1 - \epsilon$  and selects a random action with probability  $\epsilon$ , which can improve the efficiency of exploration. When training the policy, we use a technique known as experience replay. We store the agent’s experiences at each time-step,  $e_t = (s_t, a_t, r_t, s_{t+1})$  in a fixed size, queue-like buffer  $D$ .

In a simulation epoch, the current DQN network is updated multiple times (depending on the batch size and the current size of replay buffer) with different batches drawn randomly from the buffer, while the target DQN network is fixed during the updating of current DQN network. At the end of each epoch, the target network is replaced by the current network and the current network is evaluated on training set. The buffer will be flushed if the current network performs better than all previous versions.

## 4 Experiments and Results

### 4.1 Experimental Setup

The max dialogue turn  $T$  is 22. A positive reward of +44 is given to the agent at the end of a success dialogue, and a  $-22$  reward is given to a failure one. We apply a step penalty of  $-1$  for each turn to encourage shorter dialogues. The dataset is divided into two parts: 80% for training with 568 user goals and 20% for testing with 142 user goals. The  $\epsilon$  of  $\epsilon$ -greedy strategy is set to 0.1 for effective action space exploration and the  $\gamma$  in Bellman equation is 0.9. The size of buffer  $D$  is 10000 and the batch size is 30. And the neural network of DQN is a single layer network. The learning rate is 0.001. Each simulation epoch consists of 100 dialogue sessions and the current network is evaluated on 500 dialogue sessions at the end of each epoch. Before training, the buffer is pre-filled with the experiences of the rule-based agent (see below) to warm start our dialogue system.

To evaluate the performance of the proposed framework, we compare our model with baselines in terms of three evaluation metrics following Li et al. (2017) and Peng et al. (2017a; 2017b), namely, success rate, average reward and the average number of turns per dialogue session. As for classification models, we use accuracy as the metric.

The baselines include: **(1) SVM:** This model treats the automatic diagnosis as a multi-class classification problem. It takes one-hot representation of symptoms in the user goal as input, and predicts the disease. There are two configurations: one takes both explicit and implicit symptoms as input (denoted as SVM-ex&im), and the other takes only explicit symptoms to predict the disease (denoted as SVM-ex). **(2) Random Agent:** At each turn, the random agent takes an action randomly from the action space as the response to the user’s action. **(3) Rule-based Agent:** The rule-based agent takes an action based on handcrafted rules. Conditioned on the current dialogue state  $s_t$ , the agent will inform disease if all the known symptoms related are detected. If no disease can be identified, the agent will select one of the left symptoms randomly to inform. The relations between diseases and symptoms are extracted from the annotated corpus in advance. In this work, only the first  $T/2.5$ <sup>4</sup> symptoms with high frequency are kept for each disease so that the rule-

<sup>4</sup>2.5 is a hyper-parameter.

based agent could inform a disease within the max dialogue turn  $T$ .

Disease	SVM-ex&im	SVM-ex
Infantile diarrhea	0.91	0.89
Children functional dyspepsia	0.34	0.28
Upper respiratory infection	0.52	0.44
Children’s bronchitis	0.93	0.71
Overall	0.71	0.59

Table 4: Accuracy of classification models

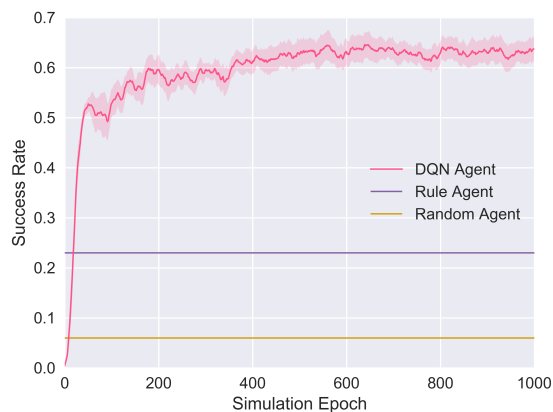


Figure 3: Learning curve of policy learning

Model	Success	Reward	Turn
Random Agent	0.06	-24.36	17.51
Rule Agent	0.23	-13.78	17.00
DQN Agent	0.65	20.51	5.11

Table 5: Performance of three dialogue systems on 5K simulated dialogues

## 4.2 Results

Table 4 shows the accuracy of two SVM-based models. The result shows that the implicit symptoms can greatly improve the accuracy of disease identification for all the four diseases, which demonstrates the contribution of implicit symptoms when making diagnosis for patients. Figure 3 shows the learning curve of all the three dialogue systems and Table 5 shows the performance of these agents on testing set. Due to the large action space, the random agent performs badly. The rule-based agent outperforms the random agent in a large margin. This indicates that the rule-based agent is well designed. We can also see that the RL-based DQN agent outperforms rule-based agent significantly. Moreover, DQN agent outperforms SVM-ex by collecting additional implicit

symptoms via conversing with patients. However, there is still a gap between the performance of *DQN agent* and *SVM-ex&im* in terms of accuracy, which indicates that there is still rooms for the improvement of the dialogue system.

## 5 Related Works

In 2003, an ontology-based dialogue system that supports electronic referrals for breast cancer is proposed (Milward and Beveridge, 2003), which can deal with the informative response of users based on the medical domain ontologies. In addition, there are two works where deep reinforcement learning is applied for automatic diagnosis (Tang et al., 2016; Kao et al., 2018). However, their models need extra human resources to categorize the diseases into different groups and the data used is simulated that can not reflect the situation of the real patients.

## 6 Conclusions and Future Works

In this paper, we propose a reinforcement learning based framework of dialogue system for automatic diagnosis and build a dataset for training DS which is derived from the dialogue text between real patients and doctors. Experiment results on a self-constructed dataset show that our dialogue system is able to collect additional symptoms via conversation with patients and improve the accuracy for automatic diagnosis.

The relationship between diseases and symptoms is an external knowledge which is thought to be useful for the automatic diagnosis. One of our future directions is to explore models that can incorporate external knowledge for better policy learning.

## Acknowledgments

We thank the anonymous reviewers for their detailed and insightful comments on this paper. The work is partially supported by National Natural Science Foundation of China (Grant No. 61702106), Shanghai Science and Technology Commission (Grant No. 17JC1420200, Grant No.17YF1427600 and Grant No. 16JC1420401). Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the funding agencies.

## References

- Edward Choi, Andy Schuetz, Walter F Stewart, and Jimeng Sun. 2016. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association* 24(2):361–370.
- Heriberto Cuayahuitl, Simon Keizer, Oliver Lemon, et al. 2015. Strategic dialogue management via deep reinforcement learning. *CoRR*.
- Finale Doshi-Velez, Yaorong Ge, and Isaac Kohane. 2014. Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. *Pediatrics* 133(1):e54–e63.
- Siddhartha R Jonnalagadda, Abhishek K Adupa, Ravi P Garg, Jessica Corona-Cox, and Sanjiv J Shah. 2017. Text mining of the electronic health record: An information extraction approach for automated identification and subphenotyping of hfpef patients for clinical trials. *Journal of cardiovascular translational research* 10(3):313–321.
- Hao-Cheng Kao, Kai-Fu Tang, and Edward Y Chang. 2018. Context-aware symptom checking for disease diagnosis using hierarchical reinforcement learning .
- Li Li, Wei-Yi Cheng, Benjamin S Glicksberg, Omri Gottesman, Ronald Tamler, Rong Chen, Erwin P Bottinger, and Joel T Dudley. 2015. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Science translational medicine* 7(311):311ra174–311ra174.
- Xiujun Li, Yun-Nung Chen, Lihong Li, Jianfeng Gao, and Asli Celikyilmaz. 2017. End-to-end task-completion neural dialogue systems. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. volume 1, pages 733–743.
- Bing Liu, Gokhan Tur, Dilek Hakkani-Tur, Pararth Shah, and Larry Heck. 2017. End-to-end optimization of task-oriented dialogue model with deep reinforcement learning <https://arxiv.org/abs/1711.10712>.
- David Milward and Martin Beveridge. 2003. Ontology-based dialogue systems. In *Proc. 3rd Workshop on Knowledge and reasoning in practical dialogue systems (IJCAI03)*. pages 9–18.
- Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533.
- Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjing Liu, Yun-Nung Chen, and Kam-Fai Wong. 2017a. Adversarial advantage actor-critic model for task-completion dialogue policy learning <https://arxiv.org/abs/1710.11277>.
- Baolin Peng, Xiujun Li, Lihong Li, Jianfeng Gao, Asli Celikyilmaz, Sungjin Lee, and Kam-Fai Wong. 2017b. Composite task-completion dialogue policy learning via hierarchical deep reinforcement learning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. pages 2231–2240.
- Jost Schatzmann, Blaise Thomson, Karl Weilhammer, Hui Ye, and Steve Young. 2007. Agenda-based user simulation for bootstrapping a pomdp dialogue system. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*. Association for Computational Linguistics, pages 149–152.
- Jost Schatzmann, Karl Weilhammer, Matt Stuttle, and Steve Young. 2006. A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies. *The knowledge engineering review* 21(2):97–126.
- Chaitanya Shivade, Preethi Raghavan, Eric Fosler-Lussier, Peter J Embi, Noemie Elhadad, Stephen B Johnson, and Albert M Lai. 2013. A review of approaches to identifying patient phenotype cohorts using electronic health records. *Journal of the American Medical Informatics Association* 21(2):221–230.
- Pei-Hao Su, Milica Gasic, Nikola Mrksic, Lina Rojas-Barahona, Stefan Ultes, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2016. Continuously learning neural dialogue management <https://arxiv.org/abs/1606.02689>.
- Kai-Fu Tang, Hao-Cheng Kao, Chun-Nan Chou, and Edward Y Chang. 2016. Inquire and diagnose: Neural symptom checking ensemble using deep reinforcement learning. In *Proceedings of NIPS Workshop on Deep Reinforcement Learning*.
- Huaixiao Tou, Lu Yao, Zhongyu Wei, Xiahai Zhuang, and Bo Zhang. 2018. Automatic infection detection based on electronic medical records. *BMC bioinformatics* 19(5):117.
- TH Wen, D Vandyke, N Mrkšić, M Gašić, LM Rojas-Barahona, PH Su, S Ultes, and S Young. 2017. A network-based end-to-end trainable task-oriented dialogue system. In *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017-Proceedings of Conference*. volume 1, pages 438–449.
- Zhao Yan, Nan Duan, Peng Chen, Ming Zhou, Jian-she Zhou, and Zhoujun Li. 2017. Building task-oriented dialogue systems for online shopping. In *AAAI*. pages 4618–4626.
- Steve Young, Milica Gašić, Blaise Thomson, and Jason D Williams. 2013. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE* 101(5):1160–1179.

Tao Zheng, Wei Xie, Liling Xu, Xiaoying He, Ya Zhang, Mingrong You, Gong Yang, and You Chen. 2017. A machine learning-based framework to identify type 2 diabetes through electronic health records. *International journal of medical informatics* 97:120–127.