Learning to Diagnose: Assimilating Clinical Narratives using Deep Reinforcement Learning

Yuan Ling, Sadid A. Hasan, Vivek Datla, Ashequl Qadir,
Kathy Lee, Joey Liu, and **Oladimeji Farri**
Artificial Intelligence Lab, Philips Research North America, Cambridge, MA, USA
{yuan.ling,sadid.hasan,vivek.datla,ashequl.qadir}@philips.com
{kathy.lee_1,joey.liu,dimeji.farri}@philips.com

Abstract

Clinical diagnosis is a critical and non-trivial aspect of patient care which often requires significant medical research and investigation based on an underlying clinical scenario. This paper proposes a novel approach by formulating clinical diagnosis as a reinforcement learning problem. During training, the reinforcement learning agent mimics the clinician's cognitive process and learns the optimal policy to obtain the most appropriate diagnoses for a clinical narrative. This is achieved through an iterative search for candidate diagnoses from external knowledge sources via a sentence-by-sentence analysis of the inherent clinical context. A deep Q-network architecture is trained to optimize a reward function that measures the accuracy of the candidate diagnoses. Experiments on the *TREC CDS* datasets demonstrate the effectiveness of our system over various non-reinforcement learning-based systems.

1 Introduction

Clinical diagnosis is a critical aspect of patient care requiring expert medical knowledge and intuition. Given a clinical case narrative such as a patient's past medical history and current condition, a clinician performs complex cognitive processes to infer the probable diagnosis based on his/her experience or up-to-date knowledge obtained from relevant external resources (Norman et al., 2007). Table 1 shows an example clinical narrative with relevant external knowledge, which suggests that *Pulmonary Embolism* is the diagnosis for this clinical scenario.[1]

Clinical Narrative:
An 87 yo woman with h/o osteoporosis, DM2, dementia, depression, and anxiety presents s/p fall with evidence of C2 fracture, chest pain, tachycardia, tachypnea, and low blood pressure.
External Knowledge (partially shown)
From Wikipedia page for Pulmonary Embolism - "Signs and symptoms" Section:
Symptoms of pulmonary embolism are typically sudden in onset and may include one or many of the following: dyspnea (shortness of breath), tachypnea (rapid breathing), chest pain of a "pleuritic" nature (worsened by breathing), cough and hemoptysis (coughing up blood).
From MayoClinic page for Pulmonary Embolism - "Symptoms" Section:
Pulmonary embolism symptoms can vary greatly, depending on how much of your lung is involved, the size of the clots, and whether you have underlying lung or heart disease.
Diagnosis:
Pulmonary Embolism

Table 1: An example clinical narrative with relevant external knowledge and diagnosis.

This paper considers the challenge of inferring the diagnoses of a patient condition based on available documentation in the Electronic Health Record (EHR), specifically free text clinical reports. Earlier work that builds Artificial Intelligence (AI) systems to support clinical decision making, mostly uses structured clinical data (e.g. physiological signals, vital signs, lab tests etc.) stored in the EHR (Lipton et al., 2015; Choi et al., 2015, 2016). They commonly formulate diagnosis inferencing as a supervised classification task.

The efficacy of these models largely depends on the size of the annotated datasets used for training, which requires expert-derived annotations that are expensive to obtain. These models also tend to lack the ability to capture the underlying uncertainties related to generating differential diagnoses (Richardson et al., 1999) and linguistic complex-

---

[1]The clinical narrative with corresponding diagnosis is obtained from the Text REtrieval Conference (TREC) Clinical Decision Support (CDS) track 2016 dataset (Roberts et al., 2016a).

ities (Seidel et al., 2015) of a clinical scenario as they consider medical codes and a finite number of diagnoses for prediction labels.

By contrast, we explore the discriminatory capability of the unstructured clinical narratives to infer the possible diagnoses. To overcome the sparsity in annotated data and adequate representation of ambiguities, we formulate the problem as a sequential decision-making process using deep reinforcement learning while leveraging external knowledge to infer the differential diagnoses.

Our proposed approach is novel as, unlike previous approaches, it focuses on the clinician's cognitive process to infer the most probable diagnoses from clinical narratives. Given a clinical text scenario, a physician typically reviews the sentences sequentially, skipping those s/he deems irrelevant and focusing on those that would contribute to his/her understanding of the clinical scenario.

While assimilating the sentences (i.e. understanding partial information), s/he tries to recognize a logical pattern or clinical progression similar to one or more prior patient encounters towards arriving at a candidate diagnosis. Ultimately, the intuition of the clinician is guided by understanding of these sentences and s/he can make an overall assessment of the scenario based on the narrative and/or additional evidence obtained from relevant external knowledge sources.

Our system replicates this cognitive flow by using a deep reinforcement learning technique. During training, the agent learns the optimal policy to obtain the final diagnoses through iterative search for candidate diagnoses from external knowledge sources via a sentence-by-sentence analysis of the inherent clinical context.

A deep Q-network architecture (Mnih et al., 2015) is trained to optimize a reward function that measures the accuracy of the candidate diagnoses. Our model predicts the differential diagnoses by utilizing the optimum policy learned to maximize the overall possible reward for an action during training. Extensive experiments on the TREC CDS track (Roberts et al., 2015, 2016a) datasets demonstrate the effectiveness of our system over several non-reinforcement learning-based systems.

In recent TREC CDS tracks, clinical diagnosis inferencing from free text clinical narratives has been showcased as a significant milestone in clinical question answering and a path to improv-ing the accuracy of relevant biomedical article retrieval (Roberts et al., 2015, 2016b; Goodwin and Harabagiu, 2016).

In addition to these established use cases, we envisage that our work can also lead to a busy clinician considering relevant differential diagnoses that could otherwise be ignored due to inadvertent diagnostic errors (Nendaz and Perrier, 2012; Graber et al., 2012; Berge and Mamede, 2013). Also, nurse practitioners can use the proposed system as a source of second opinion before contacting a physician towards accurately diagnosing and managing their patients.

2 Related Work

Addressing inference tasks generally requires significant contributions from domain experts and access to a variety of resources (Ferrucci et al., 2013; Lally et al., 2014) such as structured knowledge bases (KBs) (Yao and Van Durme, 2014; Bao et al., 2014; Dong et al., 2015). However, KBs are known to have limitations such as knowledge incompleteness, sparsity, and fixed schema (Socher et al., 2013; West et al., 2014; Bordes et al., 2014), which have motivated researchers to use unstructured textual resources like Wikipedia for various related tasks (Katz et al., 2005; Wu and Weld, 2010; Miller et al., 2016; Chen et al., 2017). In this paper, we also leverage the power of unstructured knowledge sources to address clinical diagnosis inferencing.

Previous clinical diagnosis inferencing works mostly utilized various bio-signals from patients (Lipton et al., 2015; Choi et al., 2015, 2016). EHRs typically store such structured clinical data (e.g. physiological signals, vital signs, lab tests etc.) along with unstructured text documents that contain a relatively more narrative picture of the associated clinical events.

Recently, diagnosis inferencing from unstructured clinical text has gained much attention among AI and Natural Language Processing researchers, with the advent of the *TREC CDS* tracks (Simpson et al., 2014; Roberts et al., 2015, 2016b; Goodwin and Harabagiu, 2016; Zheng and Wan, 2016; Balaneshin-kordan and Kotov, 2016; Prakash et al., 2017; Ling et al., 2017a). Although the main task in the CDS track was to retrieve relevant biomedical articles given a clinical scenario, researchers also explored diagnosis inferencing from clinical narratives as part of the pilot
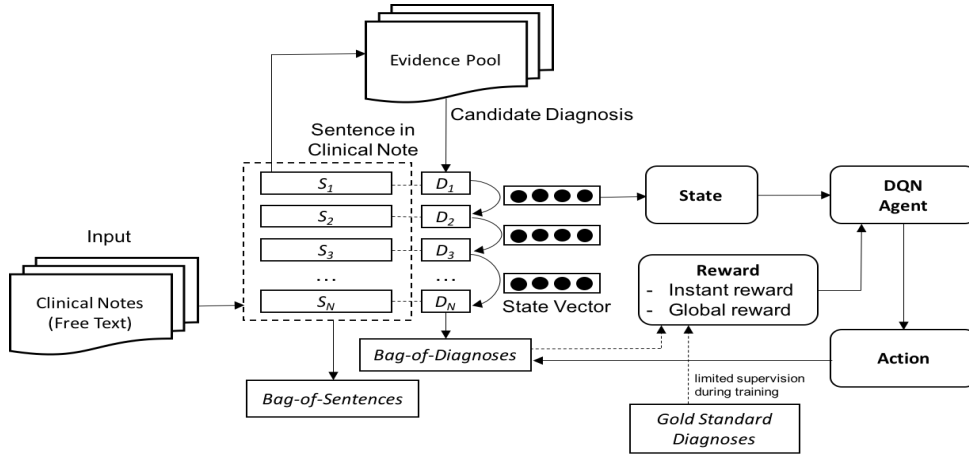
Figure 1: Overall architecture of the DBrain system.

task in 2015 that investigated the impact of diagnostic information on retrieving relevant biomedical articles (Roberts et al., 2015, 2016b).

Existing approaches for diagnosis inferencing mostly propose supervised classification models using various neural network architectures (Lipton et al., 2015; Choi et al., 2015; Prakash et al., 2017). However, such models heavily rely on large labeled data, and lack the ability to capture inherent ambiguities and complexities of a clinical scenario. Moreover, they are limited by the number of diagnosis labels and the use of medical codes to simplify the computational and linguistic difficulties of a clinical case. Other works have explored graph-based reasoning methods to incorporate relevant medical concepts and their associations (Shi et al., 2017; Geng and Zhang, 2014; Goodwin and Harabagiu, 2016; Zheng and Wan, 2016; Ling et al., 2017a).

These approaches do not focus on the intuitive and analytical processes of a clinician to infer the probable diagnoses from a clinical case narrative (Pelaccia et al., 2011; Kushniruk, 2001). By contrast, we propose a novel approach for clinical diagnosis inferencing that formulates the task as a reinforcement learning problem to mimic the clinician's cognitive process for clinical reasoning.

Prior works that use reinforcement learning for clinical decision support tasks focused on other modalities e.g. medical imaging (Netto et al., 2008) or specific domain-dependent use cases, and clinical trials (Poolla, 2003; Shortreed et al., 2011; Zhao et al., 2011), but not for inferencing diagnosis. Recent works have shown the utility of deep reinforcement learning techniques for challenging tasks like playing games and entity extraction via utilizing external evidence (Mnih et al., 2015; Narasimhan et al., 2015, 2016). To the best of our knowledge, we are the first to explore deep-reinforcement learning for clinical diagnosis inference using text data from EHR.

3 Inferencing Diagnoses with Deep Reinforcement Learning

Our proposed approach, DBrain, uses a reinforcement learning formulation that leverages evidence from external resources to mimic the clinician's complex reasoning. The overall architecture of our method is depicted in Figure 1.

DBrain takes free-text clinical narratives as input, and generates differential diagnoses as output. It scans the clinical narrative sentence-by-sentence and each sentence is used as a query to obtain a candidate diagnosis from external knowledge sources. We use a Markov Decision Process (MDP) to model this process. DBrain system creates two pools for each clinical narratives to keep the candidate sentences and the candidate diagnoses, namely: 1) bag-of-sentences, and 2) bag-of-diagnoses. Actions are taken at each step to decide which candidate sentence goes into the bag-of-sentences, and which candidate diagnosis goes into the bag-of-diagnoses.

3.1 MDP Framework

We model the integration of external knowledge sources for clinical diagnosis inferencing as a Markov Decision Process (MDP) (Bellman, 1957; Sutton and Barto, 1998). At each MDP step, the agent takes a sentence from the clinical narrative

and uses it as a query to obtain an external article from the evidence pool so that the sentence can be mapped to a candidate diagnosis. The evidence pool contains external knowledge sources, such as Wikipedia articles (details in Section 4.1).

For each sentence and the corresponding candidate diagnosis, a state vector s is created to encode their information. The state vector comprises information on the importance of the current sentence and the current candidate diagnosis with respect to inferring the most probable diagnoses for a clinical narrative. In a state s, the agent takes an action a to get to the next state, $s' = s + a$. A reward function $r(s, a)$ is used to estimate the reward at each state s after taking an action a.

We estimate a state-action value function $Q(s, a)$, which determines the optimal action a to take in a state s using the Q-learning technique (Watkins and Dayan, 1992). The Q-function is approximated using a deep Q-network (DQN) architecture (Mnih et al., 2015). The trained DQN agent takes state s and reward r as input, and outputs an action a.

Once the training is complete, the sentences in the bag-of-sentences represent the most important sentences, and the diagnoses in the bag-of-diagnoses denote the final predicated diagnoses for the clinical narrative. The overall MDP framework for clinical diagnosis inferencing is presented in Algorithm 1.

Algorithm 1: MDP framework

Input : clinical narrative $C = s_1, s_2, ..., s_n$
Output: bag-of-diagnoses D, bag-of-sentences S
1 $D = \emptyset$ and $S = \emptyset$;
2 **for** *each sentence s_i in C* **do**
3   Use s_i as query, search in knowledge sources, get candidate diagnosis d;
4   Generate state vector v for sentence-diagnosis pair (s_i, d);
5   Calculate reward value r;
6   Send (v, r) to DQN agent, and get action value a_1 and a_2 from agent (where a_1 and a_2 denote actions for diagnoses and sentences, respectively);
7   **if** action == "stop" **then break**;
8   Update D according to a_1;
9   Update S according to a_2;
10 **end**
11 return D, S

For each clinical narrative, the output is a bag-of-diagnoses D and a bag-of-sentences S. For the training phase, the steps in Algorithm 1 for each clinical narrative are run for multiple epochs. During the testing stage, each clinical narrative is processed only once in a single epoch. The next subsections provide details on the state, actions, and the reward function of the MDP framework.

3.1.1 State

The state s in our MDP comprises DBrain system's confidence on the current sentence and the corresponding candidate diagnosis. We represent state s as a continuous real-valued vector containing the following information: 1) $S1$: similarity between the current sentence and the bag-of-sentences, 2) $S2$: similarity between the current sentence and the context of the clinical narrative, 3) $S3$: similarity between the current sentence and the source article context of a candidate diagnosis, 4) $S4$: similarity between the bag-of-sentences and the source article context of a candidate diagnosis, 5) $S5$: similarity between a candidate diagnosis and the bag-of-diagnoses, and 6) number of words in the current sentence.

We compute the aforementioned similarities in two ways: 1) string similarity, which includes n-gram (unigram/bigram/trigram), and Levenshtein distance, 2) similarity/distance measures using one-hot vector representations including Jaccard similarity, cosine similarity, Manhattan distance, Euclidean distance, and fractional distance.

In addition to the above similarities, words in the current sentence are encoded into the state vector using a Long Short Term Memory (LSTM) network and mean pooling. In particular, we take the sequence of words in the current sentence as input, pass their one-hot vector embeddings to the LSTM cells, and output a corresponding vector representation, which combined with the similarities (described above) produces a state vector to serve as the input for the DQN module.

3.1.2 Actions

At each step, there are two kinds of actions for the agent: a_1 for updating the bag-of-diagnoses and a_2 for updating the bag-of-sentences, where a_1 includes: 1) accept the candidate diagnosis, 2) reject the candidate diagnosis, 3) reject all candidate diagnoses, and 4) stop; and, a_2 includes: 1) accept the current sentence, and 2) reject the current sentence.

3.1.3 Reward Function

The agent receives limited supervision from the ground truth diagnoses via a reward function during training. The reward function is chosen in
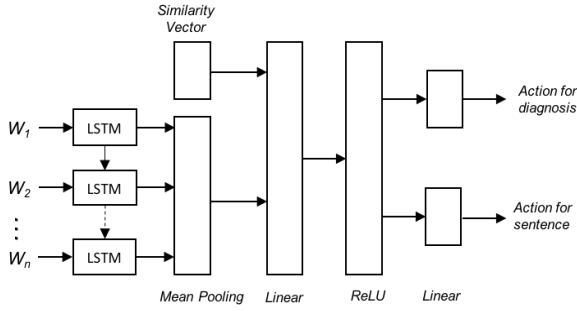
Figure 2: DQN architecture.

a way such that the accuracy of the final diagnoses prediction can be maximized. We consider two types of rewards: instant reward $r_{instant}$ and global reward r_{global}. The overall reward r is computed as:

$$r = r_{instant} + r_{global} \qquad (1)$$

where $r_{instant}$ is calculated based on the match of a candidate diagnosis with gold standard diagnoses as:

$$r_{instant} = \begin{cases} 1, & \text{if candidate diagnosis matches} \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

On the other hand, r_{global} is equal to the number of correct diagnoses minus the number of incorrect diagnoses in the bag-of-diagnoses.

3.2 DQN Architecture

In order to learn the Q-value, the iterative updates are derived from the Bellman equation(Sutton and Barto, 1998):

$$Q_{i+1}(s, a) = E[r + \gamma max_{a'} Q_i(s', a')|s, a], \qquad (3)$$

where γ is a discount factor for the future rewards and the expectation is over the whole training process.

It is impractical to maintain the Q-values for all possible state-action pairs. Mnih et al. (2015) proposed a deep Q-network (DQN) architecture, which approximates the Q-value function and predicts $Q(s, a)$ for all possible actions. We extended the DQN architecture in Narasimhan et al. (2015) to fit our problem formulation (Figure 2).

4 Experimental Setup

4.1 External Knowledge Sources

Our work relies on external knowledge sources to provide candidate diagnoses for the sentences

from a clinical narrative. We use two external knowledge sources: Wikipedia pages and MayoClinic pages. We index Wikipedia and MayoClinic using Elasticsearch[2]. As an example, Wikipedia and MayoClinic pages for the diagnosis "pulmonary embolism" are partially displayed in Table 1.

4.1.1 Wikipedia

We select 37,245 Wikipedia pages under the "clinical medicine" category[3]. Each page title is used as the diagnosis name and the texts from the *Signs and symptoms* subsection are used as an evidence for mapping candidate diagnosis. As shown in Table 1, "*Sign and symptom*" section describes symptoms of "pulmonary embolism". These symptoms have a higher chance of appearing in a clinical narrative if the documented diagnosis is "pulmonary embolism".

4.1.2 MayoClinic

The MayoClinic[4] disease corpus contains 1,117 pages, which include sections of *Symptoms*, *Causes*, *Risk Factors*, *Treatments and Drugs*, *Prevention*, etc. Each MayoClinic page title is regarded as one diagnosis. We select sentences from the "*Symptoms*" section as the external source of evidence for mapping candidate diagnoses.

4.2 Candidate Diagnosis Mapping

Each sentence from a clinical narrative is used as a query to search in both Wikipedia and MayoClinic corpora. Each search returns top 10 results per corpus. If there is any common diagnoses, we return the top ranked diagnosis as the candidate diagnosis. Otherwise, we consider the top ranked diagnosis from Wikipedia as the candidate diagnosis since Wikipedia has a higher coverage for ground truth diagnoses in both training and testing dataset. Table 2 presents the diagnoses coverage for Wikipedia and MayoClinic in our training and test set, where the test set numbers essentially denote the maximum possible recall of our systems.

	Wikipedia	MayoClinic
Training Set	93.33%	80.00%
Test Set	96.67%	86.67%

Table 2: Diagnoses coverage.

---

[2]https://www.elastic.co/
[3]https://en.wikipedia.org/wiki/Category:Clinical_medicine
[4]http://www.mayoclinic.org/diseases-conditions

4.3 Datasets of Clinical Narratives

We use the 2015 and 2016 *TREC CDS* track datasets (Roberts et al., 2015, 2016a) for our experiments. Each dataset contains 30 topics, where each topic is a medical case narrative that describes a patient scenario. A topic example is partially shown as the clinical narrative in Table 1 (see accompanied dataset for all topics).

Each topic contains "*description*", "*summary*", and "*diagnosis*" fields. "*description*" includes a comprehensive description of the patient's situation, whereas "*summary*" contains an abridged version of the most important information. In addition, the *2016* dataset includes a "*note*" field for each topic, which resembles an actual clinical note in terms of linguistic complexity. We use "*description*", "*summary*" and "*note*" fields separately to generate more samples with same/similar patient situations.

We use all fields from the *2016* dataset for training our systems, while "*description*" and "*summary*" fields from the *2015* dataset are used separately for testing (see dataset statistics in Table 3).

	Train	Test- *description*	Test- *summary*
# of Topics	30	30	30
# of Samples	90	30	30
Total # of Sent.	703	152	45
Avg. # of Sent.	7.8	5.1	1.5

Table 3: Dataset statistics.

4.4 Evaluation Metrics

We use precision, recall and F-score as the evaluation metrics. Precision is the fraction of correctly predicted diagnoses among all predicted diagnoses. Recall is the fraction of correctly predicted diagnoses among all gold standard diagnoses. F-score is calculated based on precision and recall as follows:

$$F = \frac{2 \times precision \times recall}{precision + recall} \quad (4)$$

Instead of using an exact match for comparing predicted diagnosis and gold diagnosis, we use paraphrases and disease synonyms based on the human disease network (Schriml et al., 2012) to compare two diagnosis terms.

4.5 Systems for Comparison

We explore a supervised method using Support Vector Machines (SVM), an information retrieval-based method (*IR-based*), and two heuristic methods (*KG-based* and *Concept-based*) to systematically evaluate the performance of our DBrain system. In addition, we also compare the performance among different representational variations of the DQN architecture.

4.5.1 Supervised Method

We build a supervised method using SVM (Cortes and Vapnik, 1995). Each sentence s_i in a clinical narrative is used as a query to search in knowledge sources. We use the top retrieved Wiki page, p as the candidate diagnosis. For each sentence, we get a sentence-page pair (s_i, p). If the page title indicates the correct diagnosis for a clinical narrative, we label the sentence-page pair (s_i, p) as a positive example, otherwise, the pair is labeled as a negative example.

The feature space for SVM contains 13 features[5] denoting the similarity between a sentence from the clinical narrative and an external knowledge source page: cosine similarity, Damerau-Levenshtein distance, Jaccard similarity, JaroWinkler distance (Winkler, 1995), Levenshtein distance, weighted Levenshtein distance, longest common subsequence, metric longest common subsequence (Bakkelund, 2009), N-gram similarity (Kondrak, 2005), optimal string alignment, Q-gram distance (Ukkonen, 1992), Sorensen-Dice coefficient, and the relevance score returned from Elasticsearch. The similarity scores are concatenated to generate a vector. Finally, the similarity vector and positive/negative labels are used as input to train the SVM model. During testing, each clinical narrative generates multiple sentence-page pairs and the positive diagnoses predicted by the SVM model are considered as the final diagnoses.

4.5.2 *IR-based* Method

The IR-based method has the similar setting as the supervised method. Each sentence s_i is used as a query to obtain top 5 pages as candidate diagnoses. Each page is associated with a relevance score. We combine the results from each sentence in the narrative, and use the cumulative relevance scores to get top 5 ranked diagnoses pages per clinical narrative.

4.5.3 *KG-based* Method

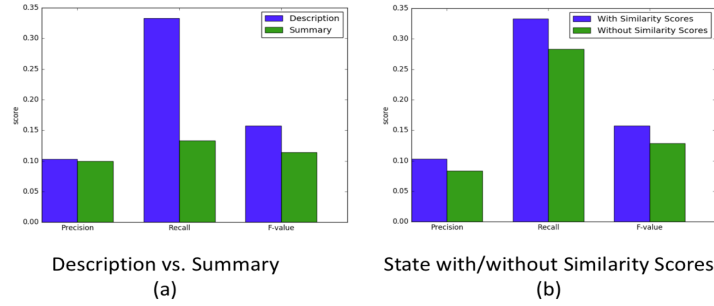We create a knowledge graph (KG)-based method, which uses Wikipedia pages under the "clinical

---

[5]https://github.com/tdebatty/java-string-similarity

Figure 3: **(a)** Description vs. Summary. **(b)** State vectors with or without similarity scores.

medicine" category to build a knowledge graph. The hierarchy of each Wikipedia page is preserved to encode its distinguishing characteristics with respect to other pages. Each page consists of several sections and is related to other medical conditions. We build a directed graph (digraph) by using these relations, where each node is a medical condition, diagnosis, test, procedure, medication or any other clinical concept, and each edge is a relation between two nodes.

The constructed knowledge graph contains $\sim 100K$ nodes and $\sim 1M$ edges, where leaf nodes represent medical symptoms and are connected to relevant diseases and medical conditions. Based on this graph, we infer the clinical diagnoses given a list of signs and symptoms extracted from a clinical narrative using a clinical information extraction engine. This method produces a ranked list of diagnoses. We take the top 5 ranked results as the diagnoses.

4.5.4 *Concept-based* **Method**

We compare our system with the concept graph-based method proposed by Ling et al. (2017a). This method builds a concept graph by integrating knowledge from structured and unstructured sources to infer top 5 ranked diagnoses from a clinical narrative.

4.5.5 **Representational Variations of DQN**

As discussed in Section 3.1.1, we use LSTM and mean pooling to encode words in a sentence. We compare the DQN-LSTM model with two variations (Figure 6) (Narasimhan et al., 2015): 1) DQN-BOW, which uses a bag-of-words approach to represent words in a sentence, and 2) DQN-Rand, where instead of using the DQN agent to choose actions, we randomly choose an action in each step.
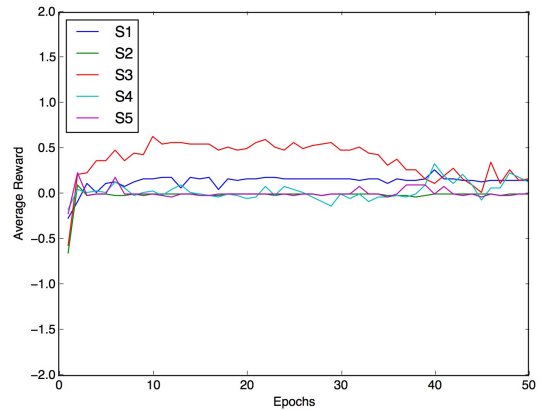

Figure 4: Evolution of reward with different similarities.

4.6 **DQN Settings**

For the DQN learning, we use a replay memory of size 50K, and a discount of 0.99. The embedding dimension is 300. All other settings are kept similar to Narasimhan et al. (2015).

5 **Results and Discussion**

5.1 **Description vs. Summary**

We use *Description* and *Summary* separately as clinical narratives for our experiments to evaluate their impact on the performance of our system. Figure 3 (a) shows Precision, Recall, and F-scores for *Description* and *Summary*. We can see that the results for *Description* is better than *Summary*. One reason is that *Description* has more average number of sentences than *Summary*. It is important for the reinforcement learning agent to infer candidate diagnoses from a sufficient number of sentences. Only one or two sentences may not be adequate for this purpose. Therefore, in the following experiments, we only use *Description* for system comparisons.
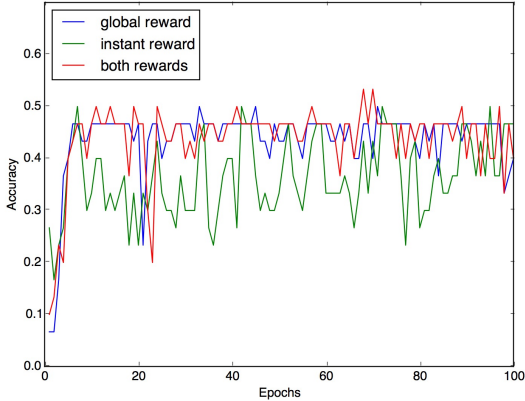
Figure 5: Evolution of accuracy.

5.2 State Vector Variations

In Figure 3 (b), we compare results for similarity scores in state vector to omitting similarity scores on state vector. We see that the inclusion of similarity vectors with the mean pooling of the context of a current sentence inside the DQN architecture provides better results for our model.

In Figure 4, we display the evolution of rewards by comparing with different similarities (used separately) as listed in Section 3.1.1. We see that $S3$, the similarity between the current sentence and the source article context of a candidate diagnosis, has better performance compared to other similarities.

5.3 Reward Functions

Figure 5 shows the learning curve of our DBrain system by measuring accuracy over epochs for different rewards functions. By using instant reward only, the accuracy trend over epochs on training set is not stable. Global reward function becomes stable after $\sim$ 10 epochs. By combining instant reward with global reward, the accuracy is slightly better than just using global reward. Therefore, we use the combined reward function in other experiments.

5.4 System Comparison Results

Table 4 presents the evaluation results of our system in comparison to other considered systems.

From these results we can see that the DBrain system achieves the best precision and F-value scores over other methods demonstrating the effectiveness of our reinforcement learning formulation. The concept-based approach shows an impressive recall score although with a loss in preci-

	Precision	Recall	F-Value
Supervised Method			
SVM	4.44	33.33	7.84
IR Method			
IR-based	7.33	36.67	12.22
Heuristic Methods			
KG-based	7.33	36.67	12.22
Concept-based	8.96	**44.78**	14.93
Our System			
DQN-BOW	**11.94**	20.00	14.11
DQN-LSTM	10.28	33.33	**15.71**

Table 4: Evaluation results (%).

sion. On the other hand, DQN-LSTM achieves the best F-Value, which is better than DQN-BOW, illustrating the importance of having a better representation of words as input. All the improvements of our system (DQN-LSTM) are statistically significant ($p < 0.05$) over SVM using the paired samples t-test (David and Gunnink, 1997) except for the methods that compute scores for the top 5 diagnoses as output (IR and heuristic-based).

Overall, the low F-measures demonstrate the difficulty of the task, as they are consistently low for all methods. We use exact sentences from a clinical narrative as queries to search for the diagnoses in the knowledge sources. Thus, sometimes our system is not able to identify the correct diagnosis due to noise in the query (see Table 6). This can be rectified with forming the query by extracting relevant clinical concepts from a sentence as shown in Ling et al. (2017b). Another reason for low F-scores is that some ground-truth diagnoses (from the training and test set) are missing in both MayoClinic and Wikipedia (Table 2). A knowledge source with a better coverage for diagnoses may offer additional room for improvements.

Figure 6 shows the evolution of average rewards for DQN-LSTM, DQN-BOW, and DQN-Rand. DQN-Rand performs poorly, which again demonstrates the importance of using a DQN agent to learn the best strategies for actions.

5.5 Example Outputs from DBrain System

We present two detailed examples to show how our DBrain system predicts the diagnoses for two test set topics. Table 5 shows that our system can correctly predict the diagnosis "Hypothyroidism" while Table 6 shows an example where the DBrain system failed to predict the correct diagnosis as the candidate diagnoses list mapped from the sentences of the clinical narrative did not contain the correct diagnosis.
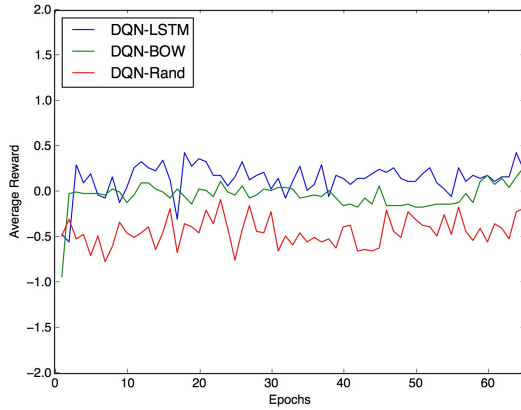
Figure 6: Evolution of reward with representational variations.

6 Conclusion

We present a novel approach for clinical diagnosis inferencing that mimics the cognitive process of clinicians using deep reinforcement learning via leveraging evidence from external resources. Our experiments on the *TREC CDS* datasets demonstrate that the DBrain system learns to diagnose by digesting clinical narratives sentence by sentence and achieves better results than supervised, IR-based, and heuristic-based methods. Furthermore, our experiments using different variations such as *Description vs. Summary* for clinical narratives, *Instant vs. Global vs. Combined* for reward functions, *State Vector with/without Similarity Scores* as input to the DQN module along with various representational variations for the DQN architecture reveal that *Description*, *Combined* reward function, *State Vector with Similarity Score*, and *DQN-LSTM* provide the best results to infer the probable diagnoses, respectively.

Example 1
Input:
Description: A 56-year old Caucasian female complains of being markedly more sensitive to the cold than most people. She also gets tired easily, has decreased appetite, and has recently tried home remedies for her constipation. Physical examination reveals hyporeflexia with delayed relaxation of knee and ankle reflexes, and very dry skin. She moves and talks slowly.
Ground-Truth Diagnosis: Hypothyroidism
Stage 1:
Sentence 1: A 56-year old Caucasian female complains of being markedly more sensitive to the cold than most people.
Candidate Diagnosis: Triple X syndrome
Action for the agent: reject the candidate diagnosis
Stage 2:
Sentence 2: She also gets tired easily, has decreased appetite, and has recently tried home remedies for her constipation.
Candidate Diagnosis: Colorectal cancer
Action for the agent: reject the candidate diagnosis
Stage 3:
Sentence 3: Physical examination reveals hyporeflexia with delayed relaxation of knee and ankle reflexes, and very dry skin.
Candidate Diagnosis: Hypothyroidism
Action for the agent: accept the candidate diagnosis
Stage 4:
Sentence 4: She moves and talks slowly.
Candidate Diagnosis: Conjugate gaze palsy
Action for the agent: reject the candidate diagnosis
Output:
Bag-of-Diagnoses: {Hypothyroidism}

Table 5: DBrain predicts the correct diagnosis.

Example 2
Input:
Description: A 31-year-old woman with no previous medical problems comes to the emergency room with a history of 2 weeks of joint pain and fatigue. Initially she had right ankle swelling and difficulty standing up and walking, all of which resolved after a few days. For the past several days she has had pain, swelling and stiffness in her knees, hips and right elbow. She also reports intermittent fevers ranging from 38.2 to 39.4 degrees Celsius and chest pain.
Ground-Truth Diagnosis: Rheumatic fever
Stage 1:
Sentence 1: A 31-year-old woman with no previous medical problems comes to the emergency room with a history of 2 weeks of joint pain and fatigue.
Candidate Diagnosis: Premenstrual syndrome
Action for the agent: reject the candidate diagnosis
Stage 2:
Sentence 2: Initially she had right ankle swelling and difficulty standing up and walking, all of which resolved after a few days.
Candidate Diagnosis: Caput succedaneum
Action for the agent: reject the candidate diagnosis
Stage 3:
Sentence 3: For the past several days she has had pain, swelling and stiffness in her knees, hips and right elbow.
Candidate Diagnosis: Synovial osteochondromatosis
Action for the agent: reject the candidate diagnosis
Stage 4:
Sentence 4: She also reports intermittent fevers ranging from 38.2 to 39.4 degrees Celsius and chest pain.
Candidate Diagnosis: Dientamoebiasis
Action for the agent: reject the candidate diagnosis
Output:
Bag-of-Diagnoses: {}

Table 6: DBrain fails to predict the correct diagnosis.

References

Daniel Bakkelund. 2009. An lcs-based string metric. *Olso, Norway: University of Oslo*.

Saeid Balaneshin-kordan and Alexander Kotov. 2016. Optimization method for weighting explicit and latent concepts in clinical decision support queries. In *ICTIR*, pages 241–250. ACM.

Junwei Bao, Nan Duan, Ming Zhou, and Tiejun Zhao. 2014. Knowledge-based question answering as machine translation. *Cell*, 2(6).

Richard Bellman. 1957. A Markovian Decision Process. *Journal of Mathematics and Mechanics*, 6:679–684.

K. Berge and S. Mamede. 2013. Cognitive diagnostic error in internal medicine. *European Journal of Internal Medicine*, 24.

Antoine Bordes, Sumit Chopra, and Jason Weston. 2014. Question answering with subgraph embeddings. *arXiv preprint arXiv:1406.3676*.

Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. 2017. Reading wikipedia to answer open-domain questions. *arXiv preprint arXiv:1704.00051*.

Edward Choi, Mohammad Taha Bahadori, Andy Schuetz, Walter F. Stewart, and Jimeng Sun. 2016. Retain: Interpretable predictive model in healthcare using reverse time attention mechanism. *CoRR*, abs/1608.05745.

Edward Choi, Mohammad Taha Bahadori, and Jimeng Sun. 2015. Doctor ai: Predicting clinical events via recurrent neural networks. *arXiv preprint arXiv:1511.05942*.

C. Cortes and V. N. Vapnik. 1995. Support Vector Networks. *Machine Learning*, 20:273–297.

Herbert A David and Jason L Gunnink. 1997. The paired t test under artificial pairing. *The American Statistician*, 51(1):9–12.

Li Dong, Furu Wei, Ming Zhou, and Ke Xu. 2015. Question answering over freebase with multicolumn convolutional neural networks. In *ACL*, pages 260–269.

David Ferrucci, Anthony Levas, Sugato Bagchi, David Gondek, and Erik T Mueller. 2013. Watson: beyond jeopardy! *Artificial Intelligence*, 199:93–105.

Shichao Geng and Qin Zhang. 2014. Clinical diagnosis expert system based on dynamic uncertain causality graph. In *Information Technology and Artificial Intelligence Conference (ITAIC), 2014 IEEE 7th Joint International*, pages 233–237. IEEE.

Travis R Goodwin and Sanda M Harabagiu. 2016. Medical question answering for clinical decision support. In *CIKM*, pages 297–306. ACM.

M. L. Graber, S. Kissam, V. L. Payne, A. N. Meyer, A. Sorensen, and N. Lenfestey. 2012. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Quality & Safety*, 21.

Boris Katz, Gregory Marton, Gary C Borchardt, Alexis Brownell, Sue Felshin, Daniel Loreto, Jesse Louis-Rosenberg, Ben Lu, Federico Mora, Stephan Stiller, et al. 2005. External knowledge sources for question answering. In *TREC*.

Grzegorz Kondrak. 2005. N-gram similarity and distance. In *String processing and information retrieval*, pages 115–126. Springer.

Andre W. Kushniruk. 2001. Analysis of complex decision-making processes in health care: Cognitive approaches to health informatics. *Journal of Biomedical Informatics*, 34:365–376.

Adam Lally, Sugato Bachi, Michael A Barborak, David W Buchanan, Jennifer Chu-Carroll, David A Ferrucci, Michael R Glass, Aditya Kalyanpur, Erik T Mueller, J William Murdock, et al. 2014. Watsonpaths: scenario-based question answering and inference over unstructured information. *Yorktown Heights: IBM Research*.

Yuan Ling, Yuan An, and Sadid A. Hasan. 2017a. Improving clinical diagnosis inference through integration of structured and unstructured knowledge. In *Proceedings of the 1st Workshop on Sense, Concept and Entity Representations and their Applications*, pages 31–36, Valencia, Spain. Association for Computational Linguistics.

Yuan Ling, Sadid A. Hasan, Vivek Datla, Ashequl Qadir, Kathy Lee, Joey Liu, and Oladimeji Farri. 2017b. Diagnostic inferencing via improving clinical concept extraction with deep reinforcement learning: A preliminary study. In *Proceedings of Machine Learning for Healthcare*.

Zachary C Lipton, David C Kale, Charles Elkan, and Randall Wetzell. 2015. Learning to diagnose with lstm recurrent neural networks. *arXiv preprint arXiv:1511.03677*.

Alexander Miller, Adam Fisch, Jesse Dodge, Amir-Hossein Karimi, Antoine Bordes, and Jason Weston. 2016. Key-value memory networks for directly reading documents. *CoRR*, abs/1606.03126.

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.

Karthik Narasimhan, Tejas D. Kulkarni, and Regina Barzilay. 2015. Language understanding for text-based games using deep reinforcement learning. In *EMNLP*, pages 1–11.

Karthik Narasimhan, Adam Yala, and Regina Barzilay. 2016. Improving information extraction by acquiring external evidence with reinforcement learning. *arXiv preprint arXiv:1603.07954*.

Mathieu Nendaz and Arnaud Perrier. 2012. Diagnostic errors and flaws in clinical reasoning: mechanisms and prevention in practice. *Swiss Medical Weekly*, 142.

Stelmo Magalhães Barros Netto, Anselmo Cardoso de Paiva, Areolino de Almeida Neto, Aristofanes Correa Silva, and Vanessa Rodrigues Coelho Leite. 2008. *Application on Reinforcement Learning for Diagnosis Based on Medical Image*. INTECH Open Access Publisher.

Geoff Norman, Meredith Young, and Lee Brooks. 2007. Non-analytical models of clinical reasoning: the role of experience. *Medical Education*, 41(12):1140–1145.

Thierry Pelaccia, Jacques Tardif, Emmanuel Triby, and Bernard Charlin. 2011. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Medical Education Online*, 16(0).

Radhika Poolla. 2003. A reinforcement learning approach to obtain treatment strategies in sequential medical decision problems. *Graduate Theses and Dissertations, University of South Florida*.

Aaditya Prakash, Siyuan Zhao, Sadid A. Hasan, Vivek Datla, Kathy Lee, Ashequl Qadir, Joey Liu, and Oladimeji Farri. 2017. Condensed Memory Networks for Clinical Diagnostic Inferencing. In *AAAI*, pages 3274–3280.

W. S. Richardson, M. C. Wilson, G. H. Guyatt, D. J. Cook, and J. Nishikawa. 1999. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. *JAMA : The Journal of the American Medical Association*, 281(13):1214–1219.

Kirk Roberts, Dina Demner-Fushman, Ellen Voorhees, and William R Hersh. 2016a. Overview of the TREC 2016 Clinical Decision Support Track. In *TREC*.

Kirk Roberts, Matthew S. Simpson, Dina Demner-Fushman, Ellen M. Voorhees, and William R. Hersh. 2016b. State-of-the-art in biomedical literature retrieval for clinical cases: a survey of the TREC 2014 CDS track. *Information Retrieval Journal*, 19(1-2):113–148.

Kirk Roberts, Matthew S. Simpson, Ellen Voorhees, and William R Hersh. 2015. Overview of the TREC 2015 Clinical Decision Support Track. In *TREC*.

Lynn Marie Schriml, Cesar Arze, Suvarna Nadendla, Yu-Wei Wayne Chang, Mark Mazaitis, Victor Felix, Gang Feng, and Warren Alden Kibbe. 2012. Disease ontology: a backbone for disease semantic integration. *Nucleic acids research*, 40(D1):D940–D946.

Bastian M. Seidel, Steven Campbell, and Erica Bell. 2015. Evidence in clinical reasoning: a computational linguistics analysis of 789,712 medical case summaries 1983–2012. *BMC Medical Informatics and Decision Making*, 15(1).

Longxiang Shi, Shijian Li, Xiaoran Yang, Jiaheng Qi, Gang Pan, and Binbin Zhou. 2017. Semantic health knowledge graph: Semantic integration of heterogeneous medical knowledge and services. *BioMed Research International*, 2017.

Susan M. Shortreed, Eric Laber, Daniel J Lizotte, T Scott Stroup, Joelle Pineau, and Susan A Murphy. 2011. Informing sequential clinical decision-making through reinforcement learning: an empirical study. *Machine learning*, 84(1-2):109–136.

Matthew S Simpson, Ellen M Voorhees, and William Hersh. 2014. Overview of the TREC 2014 Clinical Decision Support Track. In *TREC*.

Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Ng. 2013. Reasoning with neural tensor networks for knowledge base completion. In *NIPS*, pages 926–934.

Richard S. Sutton and Andrew G. Barto. 1998. *Reinforcement Learning: An Introduction* . The MIT Press, Cambridge, Massachusetts, London, England.

Esko Ukkonen. 1992. Approximate string-matching with q-grams and maximal matches. *Theoretical computer science*, 92(1):191–211.

Christopher J. C. H. Watkins and Peter Dayan. 1992. Q-learning. In *Machine Learning*, pages 279–292.

Robert West, Evgeniy Gabrilovich, Kevin Murphy, Shaohua Sun, Rahul Gupta, and Dekang Lin. 2014. Knowledge base completion via search-based question answering. In *WWW*, pages 515–526. ACM.

William E Winkler. 1995. Matching and record linkage. *Business survey methods*, 1:355–384.

Fei Wu and Daniel S. Weld. 2010. Open information extraction using wikipedia. In *ACL*, pages 118–127.

Xuchen Yao and Benjamin Van Durme. 2014. Information extraction over structured data: Question answering with freebase. In *ACL (1)*, pages 956–966. Citeseer.

Yufan Zhao, Donglin Zeng, Mark A Socinski, and Michael R Kosorok. 2011. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, 67(4):1422–1433.

Ziwei Zheng and Xiaojun Wan. 2016. Graph-based multi-modality learning for clinical decision support. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 1945–1948. ACM.