

# Analysis of Adverse Drug Reactions Identified in Nursing Notes Using Reinforcement Learning

Eunjoon Jeon<sup>1,\*</sup>, Youngsam Kim<sup>2,\*</sup>, Hojun Park<sup>3</sup>, Rae Woong Park<sup>3,4</sup>, Hyopil Shin<sup>5</sup>, Hyeoun-Ae Park<sup>6</sup>

<sup>1</sup>Technology Research, Samsung SDS, Seoul, Korea

<sup>2</sup>Institute for Cognitive Science, College of Humanities, Seoul National University, Seoul, Korea

<sup>3</sup>Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea

<sup>4</sup>Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Korea

<sup>5</sup>Department of Linguistics, Seoul National University, Seoul, Korea

<sup>6</sup>College of Nursing, Seoul National University, Seoul, Korea

**Objectives:** Electronic Health Records (EHRs)-based surveillance systems are being actively developed for detecting adverse drug reactions (ADRs), but this is being hindered by the difficulty of extracting data from unstructured records. This study performed the analysis of ADRs from nursing notes for drug safety surveillance using the temporal difference method in reinforcement learning (TD learning). **Methods:** Nursing notes of 8,316 patients (4,158 ADR and 4,158 non-ADR cases) admitted to Ajou University Hospital were used for the ADR classification task. A TD( $\lambda$ ) model was used to estimate state values for indicating the ADR risk. For the TD learning, each nursing phrase was encoded into one of seven states, and the state values estimated during training were employed for the subsequent testing phase. We applied logistic regression to the state values from the TD( $\lambda$ ) model for the classification task. **Results:** The overall accuracy of TD-based logistic regression of 0.63 was comparable to that of two machine-learning methods (0.64 for a naïve Bayes classifier and 0.63 for a support vector machine), while it outperformed two deep learning-based methods (0.58 for a text convolutional neural network and 0.61 for a long short-term memory neural network). Most importantly, it was found that the TD-based method can estimate state values according to the context of nursing phrases. **Conclusions:** TD learning is a promising approach because it can exploit contextual, time-dependent aspects of the available data and provide an analysis of the severity of ADRs in a fully incremental manner.

**Keywords:** Drug-Related Side Effects and Adverse Reactions, Electronic Health Records, Machine Learning, Deep Learning, Nursing Records

**Submitted:** July 26, 2019

**Revised:** 1st, October 18, 2019; 2nd, December 27, 2019;  
3rd, February 20, 2020

**Accepted:** March 27, 2020

## Corresponding Author

Hyeoun-Ae Park

College of Nursing, Seoul National University, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea. Tel: +82-2-740-8827, E-mail: hapark@snu.ac.kr (<https://orcid.org/0000-0002-3770-4998>)

\*These authors are contributed equally to this work.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2020 The Korean Society of Medical Informatics

## 1. Introduction

The digitization of healthcare data of patients, commonly processed as Electronic Health Records (EHRs), has enabled researchers to analyze the health conditions of patients on a large scale, which was almost impossible a few decades ago [1].

In line with the widespread use of EHRs, pharmacovigilance monitoring using EHR data has been applied in recent years, with many studies detecting adverse drug reactions (ADRs) to improve patient safety in relation to the use of medicines [2]. The active reporting method saves time and

effort while monitoring ADR cases with medicines that are not frequently prescribed.

However, most active surveillance systems have used structured data in EHRs and structured data only account for about 20% of the total amount of data stored in the health sector, with the remaining 80% of data consisting of unstructured natural language text including medical notes and nursing notes [3]. Substantial amounts of ADR signals are expressed in nursing notes, which clinicians can use to identify and interpret potential ADRs [4].

One of the natural language processing (NLP)-based methods suggested in the previous studies utilizes hand-picked rules and selected terms that are mostly derived from external dictionaries in the target domain [5,6]. Another approach utilizing natural language data is primarily based on machine-learning and deep learning methods [7,8]. However, these methods only produce high precision scores when they are performed in a laboratory situation. Also, these previous studies have viewed ADR detection as a static analysis problem, categorizing a particular phrase of longitudinal text as either relevant or irrelevant to ADRs. Few ADRs are determined by a single event; rather, they are normally caused by a series of an indefinite number of ADR-related events.

Temporal difference (TD) learning is the core algorithm of reinforcement learning that has been successfully applied to a range of complicated prediction problems [9,10]. One inherent property of TD learning we want to highlight is “incrementality”, which refers to TD-based methods not requiring a complete set of data to make a prediction. This property is desirable for longitudinal data, such as nursing notes, which should be analyzed electronically. TD learning has several advantages. It provides continuous analysis, giving an incremental estimate for every new nursing phrase stored in an EHR system; the value estimate is based on time-dependent contextual information, rather than on a snapshot of the time series data; and it can be used seamlessly with continuous feedback.

The goal of this study was to develop a flexible method to deal with such noisy, longitudinal, time series data, such as nursing notes. This goal is two fold. One is to predict the occurrence of ADRs based on the narrative nursing phrases for each person. The other is to devise a method for monitoring the risk of ADRs based on each nursing phrase.

In this paper, we used a TD( $\lambda$ ) model to estimate the state values of nursing phrases that indicate ADR risks. We applied logistic regression to the state values from the TD( $\lambda$ ) model for an ADR classification task to predict whether a phrase was relevant to ADRs. We evaluated the performance

of our proposed method by comparing it with those of four other methods: Naive Bayes (NB), support vector machine (SVM), text-convolutional neural network (CNN), and long short-term memory (LSTM).

## II. Methods

### 1. Data Collection and Processing

This study received Institutional Review Board approval from Ajou University Hospital (No. AJIRB-MED-MDB-17-087). The data analyzed in this study were derived from the EHRs of 380,600 patients hospitalized between June 1994 and July 2015 at Ajou University Hospital in Korea. ADR reports were available for 5,503 patients, of whom 4,158 who could match the control group were selected as the experimental group. Control group subjects were selected to match the experimental group subjects on a 1:1 basis for sex, age (within 1 year), inpatient department, and hospitalization period (within 1 day).

Table 1 presents examples of the nursing phrases regarding patients used in this study. The nursing records of the study subjects comprised 4,625,547 nursing phrases, of which 837,293 were lexically unique (but not semantically unique). The maximum number of phrases recorded for a single patient was 10,625, and the mean number of phrases was 421.

Nursing phrases documented before the occurrence of an ADR were selected. However, because there was no such reference time in the non-ADR cases, a random time point was chosen during the hospitalization period, and nursing phrases documented before that point of time were selected. Reinforcement learning, NB, and SVM used all nursing phrases documented before ADR (or before the random point of time), while text-CNN and LSTM neural network used 288 and 200 nursing phrases, respectively.

We conducted several experiments of ADR classifications using narrative nursing notes. For preprocessing, raw Korean texts materials were POS-tagged, and the words with POS-tagging were stored (e.g., ‘수액 주입중’ became ‘수액’/NNG ‘주입’/NNG ‘중’/NNB) and the constructed forms were stored in a dictionary format with unique index numbers. Some of the tokens were removed in the NB or SVM experimental methods for fine tuning. All experiments were conducted by coded programs using Python, NLTK, Gensim, TensorFlow, and scikit-learn (the used codes and appendixes have been released in [https://github.com/Youngsam/adr\\_analysis\\_paper](https://github.com/Youngsam/adr_analysis_paper)).

## 2. TD Learning

Our proposed model is presented graphically as two separable processes (Figure 1). Figure 1A shows the TD learning

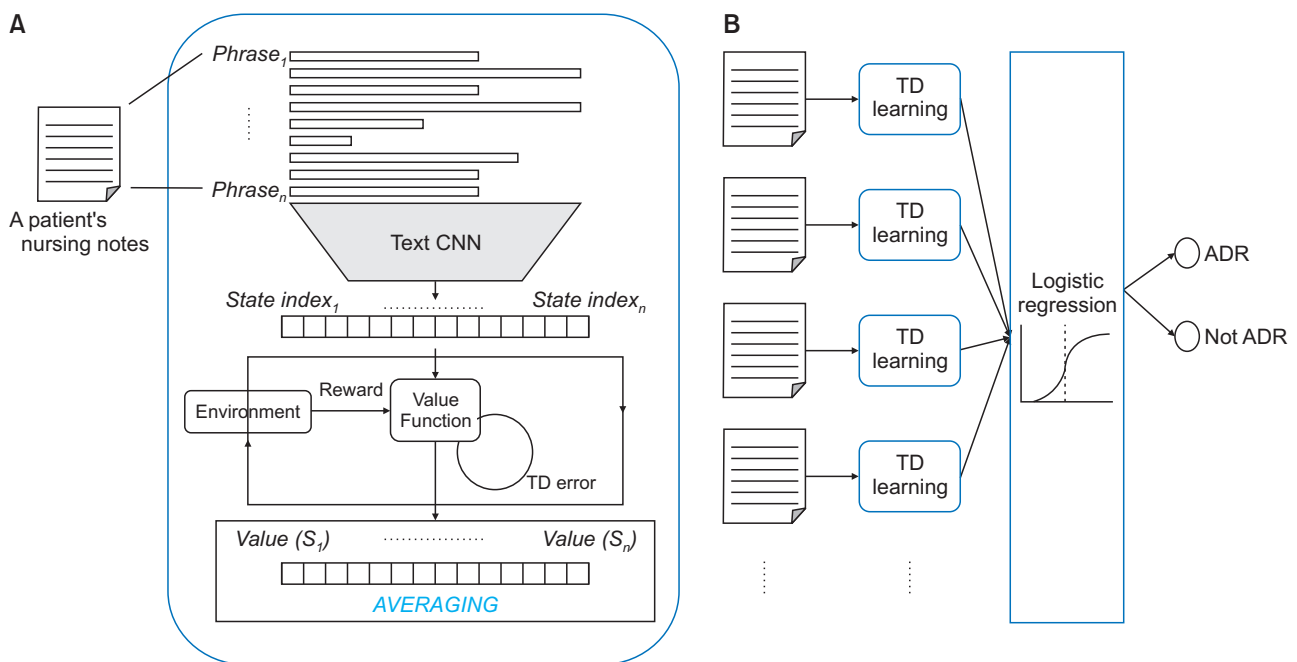
**Table 1. Example nursing phrases of a patient**

Time	Nursing phrase
2012-06-27 05:55:00	Education given to patient about deep breathing technique
2012-06-27 05:55:00	Oral care given
2012-06-27 06:30:00	Decreasing nausea
2012-06-27 08:00:00	Bed rest in place
2012-06-27 08:00:00	Maintenance fluids are given (site, right arm; gage, 28G)
2012-06-27 08:00:00	No pain, no swelling, no redness at IV site
2012-06-27 08:00:00	Education given to patient about dangers of extravasation drugs and symptoms
2012-06-27 09:20:00	No pain, no swelling, no redness at IV site
2012-06-27 09:20:00	Keep fasting
2012-06-27 09:20:00	No thirst
2012-06-27 09:20:00	Observed symptoms of water shortage

IV: intravenous.

process of state values for the seven predefined states. Each nursing phrase is assigned a state index by the trained text-CNN classifier. If a patient has nursing phrases with a size of N, there would be N state indexes (e.g., 0, 1, 0, 1, 6, 5, ...) for each patient. Our value function involves estimating a value for each state while looping the sequence of the states of nursing phrases. In each update, the value function for a state is changed to represent the expected risk of an ADR for that state based on the nursing phrases. After the value function has been trained, the learnt state values can be used for the ADR classification task. Figure 1B summarizes the entire procedure of our classification method. Logistic regression was applied to the validation dataset with the learned state values from the training dataset, and the logistic regression classifier was tested on our test dataset. We collected TRUE or FALSE labels for the nursing phrases for each patient and used the accuracy to calculate the performance of the method.

We attempted to define nursing phrases as state indexes using the categories listed in Table 2. Assigning a state to each nursing phrase is a difficult but necessary process to make our prediction fit into the framework of reinforcement learning. We therefore decided to use a small number of discrete states and created a supervised classifier that returned the state index corresponding to a nursing phrase. For the labeling of each state index, two experts on nursing informatics



**Figure 1. Our proposed model as two separable processes: (A) TD learning process of state values for the seven predefined states and (B) the entire procedure of our classification method. ADR: adverse drug reaction, TD: temporal difference, CNN: convolutional neural network.**

Table 2. Categories of nursing phrases

State index	Category of nursing phrase	Nursing phrase
0	Unknown	Patient came back after receiving CT
1	Drug-related	Injected Epocelin (1 g)
2	Abnormal reaction	Patient is describing skin itching (region, both arms)
3	Doctor related	Notified to doctor
4	Subjective response	Subjective statement: "I feel better"
5	Drug-related and abnormal reaction	Patient vomited twice after taking tramadol
6	Subjective response and drug-related	Subjective statement: "I feel like throwing up after taking the pill"

CT: computed tomography.

Table 3. Examples of annotated ADR-relevant phrases and event types

Nursing phrase	Relevant to ADRs?	State index
Invasive procedure performed	No	0
No signs of infection: no swelling, no redness, and no pain	No	0
Patient reports decreasing headache	No	0
No pain, no swelling, no redness at IV site	No	0
Invasive procedure performed	No	0
No symptoms of infection	No	0
No sign of infection	No	0
No discharge at the tube insertion site	No	0
Measured vital signs: body temperature of 37.2°C	No	0
Subjective statement: "I had muscle pain and stiffness after changing my nutrition"	Yes	6
Check the content of TPN: Oliclinomel + MVH	No	0
Extremities have become stiff and complains about muscular pain	Yes	2
Called the doctor: Dr. xxx	Yes	3
Dr. xxx ordered to stop injecting fluid and keep under observation	Yes	1
Patient reports decreasing pain	No	0
Assessed insertion tube: site, abdomen; condition, sound pressure; type, Barovac	No	0
Patient has been fasting for 2 days	No	0

ADR: adverse drug reaction, IV: intravenous; TPN: total parenteral nutrition.

analyzed the datasets of 298 randomly selected patients with reported ADRs. In the dataset, 347 ADR-relevant phrases were found from among a total of 15,642 phrases, and these were categorized into seven types (Table 3 provides examples of the annotations). To train a classifier for the categorization, we constructed a dataset of 542 phrases by combining 347 ADR-relevant phrases and 195 non ADR-relevant phrases selected randomly from 15,295 phrases. The entire dataset was divided into training, validation, and test sets at a ratio 8:1:1. We used the text-CNN model of Kim [11] to classify each nursing phrase into one of the seven categories. We used the same hyperparameters as Kim [11] with early

stopping and obtained accuracies of 95% for the test set.

Due to the unreliable time delay of the ADR reports, we applied a practice used in reinforcement learning called "reward shaping", whereby additional training rewards are used to guide the learning agent [12]. In our implementation of reward shaping, nursing phrases of all states except 0 or 1 received 1 as the reward. In addition, a reward of 1 was given for a phrase at the time when the official ADR code was reported via a different channel of the EHR data. If a patient had not received any ADR report until discharge, a reward of -1 was given. Figure 2 graphically presents the general process. Regarding reward assignments, we defined that

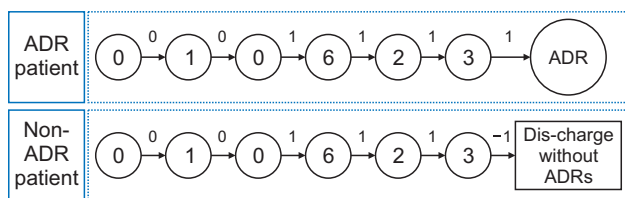


Figure 2. The general process of reward shaping.

every transition to the state of 1 would receive a reward of 0 based on a report that ADRs rarely occur when drugs are prescribed.

Logistic regression was adopted to utilize the predicted state values for the ADR classification task, which involves taking the average prediction value for each patient and returning the corresponding label (ADR or non-ADR). The data were divided into training, validation, and test sets (ratio 8:1:1), which contained 6,652, 832, and 832 subjects, respectively. All of the phrases were substituted with seven categories using the pretrained classifier following the procedure described above. The state values were estimated using the training set, and the logistic regressor was trained using the validation set. The accuracy of the logistic regression was obtained by comparing it with the test set. We experimented with three types of eligibility tracing (replacing, accumulating, and Dutch) and assigned two hyperparameters (0.1 for the discount rate and 0.3 for the trace-decay rate).

### 3. Other Methods

As previously mentioned, NB and SVM used all nursing phrases documented before ADR, while text-CNN and LSTM neural network used 288 and 200 nursing phrases, respectively. The dataset was split in the same as that in the previous procedure.

#### 1) Naive Bayes

About 9,000 unique phrases that appeared more than 20 times were selected as features for a NB classifier. This type of classifier is a simple probabilistic classifier based on Bayes' theorem with strong (naive) independence assumptions between the features. To use the full potential of the model, information gains were utilized to determine the n-best features using a grid-search method. This search method yielded the best set of features comprising 3,700 phrases, which were used to classify the test dataset.

#### 2) Support vector machine

This study used two SVM models: a linear SVM model and a radial basis function (RBF) SVM model. The nursing

phrases of each patient were preprocessed with TF-IDF (term frequency-inverse document frequency) vectorization, and the processed features were employed for the SVM classifier. To maximize the performance, grid searching was used to find the best hyperparameters for the method. The optimal minimum document frequency was 5, and the best proportion for the maximum document frequency was 0.5. For the RBF SVM setting, a value of 0.1 was used for the parameter of gamma and an integer, and 5 was used for the decision boundary parameter.

#### 3) Text-CNN

We implemented a text-CNN model for our task by replacing the input vectors of words with vectors of nursing notes. In the model of Kim [11], each word in a sentence is treated as a k-dimensional vector, and the sentence is represented as the concatenation of the word vectors. Convolution operations with various filters are then applied to the concatenation, which yields a set of feature maps that are passed with maximum pooling to a fully connected Softmax layer whose output is the probability distribution over labels.

We used a paragraph vector model for sentence embedding and compared the results with those in the lookup embedding condition [13]. We set the dimension size of embedding as 200 to make it compatible with the text-CNN input size. Text-CNN's parameters included the filter sizes of 3, 4, and 5; number of filters of 128; dropout probability of 0.5; batch size of 64; and number of epochs of 20. Applying the text-CNN model to our task required the size of the longitudinal nursing phrases to be fixed to the manageable length of 288.

#### 4) LSTM

The LSTM method [14]—which is a variant of recurrent neural network models—utilizes LSTM as the classifier for a sequence of nursing phrases due to its robustness in dealing with long-sequence data. Similarly, the text-CNN method was used with two types of embedding methods (naive sentence encoding vs. pretrained paragraph vectors), and a sequence of 200 phrases before the time point when an ADR was reported was used for the input, while a series of 200 nursing phrases was randomly selected in non-ADR cases. The parameters used in the LSTM classifier were the LSTM cell size of 256, number of LSTM layers of 2, dropout probability of 0.5, batch size of 64, learning rate of 0.01, and number of epochs of 7.

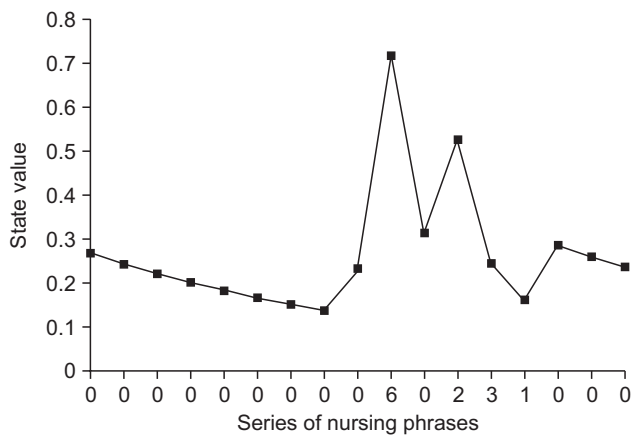


Figure 3. The concrete results obtained from the temporal difference-based predictions in Table 3.

### III. Results

#### 1. Estimation of State Values by TD Learning

Figure 3 displays the concrete results obtained from the TD-based predictions in Table 3. Figure 3 shows that the state value for a phrase with an index of 0 (unknown phrase) increased slightly immediately before one with an index of 6 (subjective response and drug-related phrase), which represents a clear sign of bootstrapping. Thus, the occurrence of ADR-relevant events within a narrow window of time will produce a sharp temporal increase in the state values over that period, thereby providing information about changes in the ADR risk.

#### 2. Accuracy of TD-Based Logistic Regression

The performance of TD-based logistic regression was measured based on the accuracy of ADR and non-ADR cases. Prior to the main evaluation, we investigated the effects of hyperparameters, such as the discount rate ( $\gamma$ ) and trace decay rate ( $\lambda$ ) based on the results obtained in the experiments with parameter settings of 0.1 for the learning rate, 0.1 for the discount rate, and 0.3 for the trace-decay rate. The results for various types of eligibility traces in the TD learning method were 0.51 for replacing, 0.61 for accumulating, and 0.63 for Dutch traces. The score was highest for Dutch traces, which were recently introduced in the literature on reinforcement learning [10].

The accuracies of all the methods used in this study are presented in Table 4. The accuracy was highest for NB (0.64) and was comparable for the TD-based method and SVM (0.63). The text-CNN method, which used non-pretrained and pretrained embedding, exhibited an accuracy of 0.58, and the LSTM model with non-pretrained embedding ex-

Table 4. Accuracies of various methods for ADR prediction

Method	Accuracy
TD-based logistic regression	0.63
Naive Bayes	0.64
SVM (linear)	0.63
SVM (RBF)	0.63
Text-CNN	0.58
Text-CNN with pretrained embedding	0.58
LSTM	0.61
LSTM with pretrained embedding	0.57

ADR: adverse drug reaction, TD: temporal difference, SVM: support vector machine, RBF: radial basis function, CNN: convolutional neural network, LSTM: long short-term memory.

hibited a higher accuracy (0.61), while the performance of the LSTM with pretrained embedding showed the lowest accuracy (0.57) of all of the methods. It was found that using pretrained embedding did not improve the accuracies of the text-CNN and LSTM models.

### IV. Discussion

In this study, we experimentally assessed various well-known methods for the analysis of potential adverse drug events. Although the TD learning methods did not show the best performance, the TD-based predictions seemed to be at least comparable to those obtained using NB and SVM methods. As noted in previous researches [2,3], we observed that ADR-relevant phrases include both major ADRs (e.g., hematologic toxicity, hematuria, and seizure-like symptoms) and minor ADRs (e.g., cough, nausea, and itching). This heterogeneity seemed to make predictions of adverse drug events very difficult for all the experimental methods in general.

TD-based methods are based on incremental learning; thus, they are easy to apply to the incremental monitoring of events such as ADRs. Other supervised methods can, of course, be utilized to predict ADRs using real-time data, but the principles of the approaches do not fully exploit the properties of adverse drug events, which are highly contextualized and temporally dependent. Also, most deep learning methods must split input data when the size of the data is too big to handle. In particular, recurrent neural network (RNN)-based models still have difficulty capturing long-term dependencies in sequences [14]. TD learning methods take advantage of these features since they are designed to use the context in making predictions [15].

Transforming raw nursing phrases into seven classes sig-

nificantly reduced the computational complexity of the task in this study. Although this involves a very laborious annotation process, it makes the simple application of TD learning possible. Because we predefined classes of nursing phrases, we could utilize the technique of reward shaping [12], which is a convenient heuristic for reinforcement learning. The implementation of reward shaping was unavoidable because the EHR data could not provide sufficient information for the method we used. Our manual observations of the raw data revealed that ADRs are often underreported, which implies that many ADR-relevant phrases were not documented in the datasets of non-ADR cases. This observation agrees with research that has shown that only 6% to 10% of all ADRs are reported globally [16].

In this study, the event classes were not obtained automatically; thus, manual annotation of the nursing phrases and standardization of the nursing notes into diverse semantic groups were required. Although knowledge-free systems are often preferred because of their convenience, we think that our knowledge-based systems have some benefits. This solution seems to be a better approach because it guarantees highly accurate event representations and can be extended with external databases. Mapping all recorded nursing phrases to standard terminologies, such as ICNP and SNOMED-CT, will allow various machine-learning methods to be used to analyze nursing phrases, thereby making the data more useful. The annotations will be more effective if they are combined with logical inference-based prediction. In the future, we believe that such a knowledge-based approach in reinforcement learning will be preferred to model-free reinforcement learning methods due to its expandability to existing knowledge resources.

In conclusion, we introduced a novel architecture of TD learning for the analysis of nursing notes in this paper. Many practical problems of the TD learning method have been explained, and the proposed solutions have been investigated with the aim of determining adequate parameters for practical applications. The TD model can be used for monitoring of nursing notes, and the results obtained for its predictive power in evaluating an ADR classification task showed that the performance of TD-based methods is comparable to those of other machine learning-based methods, while taking advantage of incremental learning.

## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## Acknowledgments

This work was supported by a Seoul National University Research Grant in 2016.

## ORCID

Eunjoo Jeon (<http://orcid.org/0000-0003-3340-6376>)

Youngsam Kim (<http://orcid.org/0000-0003-3342-9340>)

Hojun Park (<http://orcid.org/0000-0003-4965-872X>)

Rae Woong Park (<http://orcid.org/0000-0003-4989-3287>)

Hyopil Shin (<http://orcid.org/0000-0002-7279-014X>)

Hyeoun-Ae Park (<http://orcid.org/0000-0002-3770-4998>)

## References

1. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 2012;13(6):395-405.
2. Oliveira JL, Lopes P, Nunes T, Campos D, Boyer S, Ahlberg E, et al. The EU-ADR Web Platform: delivering advanced pharmacovigilance tools. *Pharmacoepidemiol Drug Saf* 2013;22(5):459-67.
3. Kho AN, Rasmussen LV, Connolly JJ, Peissig PL, Starren J, Hakonarson H, et al. Practical challenges in integrating genomic data into the electronic health record. *Genet Med* 2013;15(10):772-8.
4. Ahn HJ, Park HA. Adverse-drug-event surveillance using narrative nursing records in electronic nursing records. *Comput Inform Nurs* 2013;31(1):45-51.
5. Lee S, Choi J, Kim HS, Kim GJ, Lee KH, Park CH, et al. Standard-based comprehensive detection of adverse drug reaction signals from nursing statements and laboratory results in electronic health records. *J Am Med Inform Assoc* 2017;24(4):697-708.
6. Park MY, Yoon D, Lee K, Kang SY, Park I, Lee SH, et al. A novel algorithm for detection of adverse drug reaction signals using a hospital electronic medical record database. *Pharmacoepidemiol Drug Saf* 2011;20(6):598-607.
7. Jagannatha AN, Yu H. Structured prediction models for RNN based sequence labeling in clinical text. *Proc Conf Empir Methods Nat Lang Process* 2016;2016:856-65.
8. Hughes M, Li I, Kotoulas S, Suzumura T. Medical text classification using convolutional neural networks. *Stud Health Technol Inform* 2017;235:246-50.
9. Sutton RS. Temporal credit assignment in reinforcement learning [dissertation]. Amherst, MA: University of

- Massachusetts Amherst; 1984.
10. Van Seijen H, Mahmood AR, Pilarski PM, Machado MC, Sutton RS. True online temporal-difference learning. *J Mach Learn Res* 2016;17(1):5057-96.
  11. Kim Y. Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; 2014 Oct 25-29; Doha, Qatar. p. 1746-51.
  12. Ng AY, Harada D, Russell S. Policy invariance under reward transformations: theory and application to reward shaping. *Proceedings of the Sixteenth International Conference on Machine Learning (ICML)*; 1999 Jun 27-30; Bled, Slovenia. p. 278-87.
  13. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997;9(8):1735-80.
  14. Trinh TH, Dai AM, Luong MT, Le QV. Learning longer-term dependencies in RNNs with auxiliary losses. *Proceedings of the 35th International Conference on Machine Learning (ICML)*; 2018 Jul 10-15; Stockholm, Sweden. p. 4972-81.
  15. Bahdanau D, Brakel P, Xu K, Goyal A, Lowe R, Pineau J, et al. An actor-critic algorithm for sequence prediction. *Proceedings of the 5th International Conference on Learning Representations (ICLR)*; 2017 Apr 24-26; Toulon, France.
  16. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2009;32(1):19-31.