J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

SVM-기반 제약 조건과 강화학습의 Q-learning을 이용한 변별력이 확실한 특징 패턴 선택☆

Variable Selection of Feature Pattern using SVM-based Criterion with Q-Learning in Reinforcement Learning

김 차 영1*
Chayoung Kim

요 약

RNA 시퀀싱 데이터 (RNA-seq)에서 수집된 많은 양의 데이터에 변별력이 확실한 특징 패턴 선택이 유용하며, 차별성 있는 특징을 정의하는 것이 쉽지 않다. 이러한 이유는 빅데이터 자체의 특징으로써, 많은 양의 데이터에 중복이 포함되어 있기 때문이다. 해당 이슈 때문에, 컴퓨터를 사용하여 처리하는 분야에서 특징 선택은 랜덤 포레스트, K-Nearest, 및 서포트-벡터-머신 (SVM)과 같은 다양한 머신러닝 기법을 도입하여 해결하려고 노력한다. 해당 분야에서도 SVM-기반 제약을 사용하는 서포트-벡터-머신-재귀-특징-제거 (SVM-RFE) 알고리즘은 많은 연구자들에 의해 꾸준히 연구 되어 왔다. 본 논문의 제안 방법은 RNA 시퀀싱 데이터에서 빅-데이터처리를 위해 SVM-RFE에 강화학습의 Q-learning을 접목하여, 중요도가 추가되는 벡터를 세밀하게 추출함으로써, 변별력이 확실한 특징선택 방법을 제안한다. NCBI-GEO와 같은 빅-데이터에서 공개된 일부의 리보솜 단백질 클러스터 데이터에 본 논문에서 제안된 알고리즘을 적용하고, 해당 알고리즘에 의해 나온 결과와 이전 공개된 SVM의 Welch' T를 적용한 알고리즘의 결과를 비교 평가하였다. 해당결과의 비교가 본 논문에서 제안하는 알고리즘이 좀 더 나은 성능을 보여줌을 알 수 있다.

☞ 주제어 : 서포트-벡터-머신, RNA시퀀싱, 빅-데이터, SVM-RFE알고리즘, Q-learning, 강화학습

ABSTRACT

Selection of feature pattern gathered from the observation of the RNA sequencing data (RNA-seq) are not all equally informative for identification of differential expressions: some of them may be noisy, correlated or irrelevant because of redundancy in Big-Data sets. Variable selection of feature pattern aims at differential expressed gene set that is significantly relevant for a special task. This issues are complex and important in many domains, for example. In terms of a computational research field of machine learning, selection of feature pattern has been studied such as Random Forest, K-Nearest and Support Vector Machine (SVM). One of most the well-known machine learning algorithms is SVM, which is classical as well as original. The one of a member of SVM-criterion is Support Vector Machine-Recursive Feature Elimination (SVM-RFE), which have been utilized in our research work. We propose a novel algorithm of the SVM-RFE with Q-learning in reinforcement learning for better variable selection of feature pattern. By comparing our proposed algorithm with the well-known SVM-RFE combining Welch' T in published data, our result can show that the criterion from weight vector of SVM-RFE enhanced by Q-learning has been improved by an off-policy by a more exploratory scheme of Q-learning.

☞ keyword : Support-Vector Machine, RNA sequencing Big-Data, Support Vector Machine-Recursive Feature Elimination, Q-learning, Reinforcement Learning

1. Introduction

The primary source of information about a machine learning in these days is mostly RNA sequencing Big-data (RNA-seq). In terms of one of numerical systems in computational research area, like a Support Vector Machines-Recursive Feature Elimination (SVM-RFE), based on Support Vector Machine (SVM) criterion , data are usually represented as vectors such as featured patterns, especially in RNA-seq. They may correspond to measurements, being performed on the information gathered from the observation of a phenomenon. Usually all featured pattern genes in RNA-seq are not equally

1 Division of General Studies, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon, Gyeonggi, Korea
* Corresponding author (kimcha0@kgu.ac.kr)
[Received 26 December 2018, Reviewed 23 January 2019(R2 22 February 2019), Accepted 7 March 2018]
☆ A preliminary version of this paper was presented at APIC-IST 2018 and was selected as an outstanding paper.

informative: some of them may be noisy, meaningless, correlated or irrelevant. Therefore, they are lack of identifying deferentially expressed genes [1]. A variable selection of feature pattern aims at selecting a subset of the featured genes which is distinguishable relevant for specially problems [1]. It is an important open issue: the hugh amount of data to gather or process should be reduced. That means if training itself might be easier, then the better estimates will be obtained when using relevant featured genes of RNA-seq. Therefore, more sophisticated processing algorithms should be used on smaller dimensional spaces than on the original measure space. Moreover, computational performances might get increasing when non-relevant informations do not interfere the processes [2, 3, 4, 5]. A variable selection of feature pattern has been the subject of intensive researches in the application of identifying differential-expressed genes for the maximum gene relevancy and minimum gene redundancy. It has recently began to be investigated in the machine learning algorithms such as random forest, K-nearest, and SVM. Because of the curse of dimensionality, whatever the domain is, a variable feature selection remains an open issue and non-monotonous. Moreover, the size of expressed genes are extremely larger than those of Big-data samples. That means the distinguishable subset of p variables for the discrimination does not always contain the best discriminate subset of q variables ($q < p$) [6, 7]. Most algorithms for variable selection rely on human heuristics in the machine learning which perform a limited exploration on the whole set of variable combinations [8, 9, 10]. In the field of machine learning, feature selection has been studied. One of the most well-known machine learning algorithms is SVM, which is classical as well as original. And the one of a member of SVM-criterion is Support Vector Machine-Recursive Feature Elimination (SVM-RFE), which have been utilized in our research work.

We propose a novel algorithm of exploiting the efficiency of criteria derived from Support Vector Machines-Recursive Feature Elimination (SVM-RFE) [1] combining off-policy Q reinforcement learning for a variable feature selection in application to differential-expressed genes in RNA-seq. We especially employ an off-policy Q-learning in reinforcement learning to be trained for controling the optimization of the criterion for better weight vectors in the SVM-RFE. Due to the reinforcement learning [2-5], the self-teaching algorithms are designed in the area of Big-Data on open issues. By accompanying algorithm in the area of RNA-seq Big-data, the variable feature selection on the huge amount of data might be helpful for reducing loads on Big-Data issues. Moreover, the variable feature selection might be appropriate for resource demands in other different researches. We think that our proposed algorithm based on reinforcement learning is beyond the feature selection in the area of Big-data because reinforced selection algorithm makes the refined meaningful features. An off-policy Q-learning is regarded as superior to a discount method and a randomness of off-policy Q-learning might cut on the relevance of data because exploration are compromised in a policy of Q-learning.

By comparing our proposed algorithm with the well-known SVM-RFE combing Welch' T, our result can show that the criterion from weight vector of SVM-RFE enhanced by Q-learning has been improved by a greediness of off-policy following a more exploratory of Q-learning.

2. METHODS

2.1 MOTIVATION

Our first purpose of the proposed algorithm is that enhancing the weight vectors of SVM by exploration and exploitation. There is a big difference between on-policy and off-policy. SARSA or TD-learning is one of on-policies and a Q-learning is one of off-policies. Before we are regarding of the two policies, let us give a simple example. There is one decision that a robot moves either to a nearby door or to a distant gate. In terms of Q-learning with low γ by reducing a value, it moves to the nearby door (a goal). An off-policy Q-learning might have more advantage in terms of discounted methods and get captured in limit cycles. Therefore, trying to do random, as known as exploration is extremely important for a long-term goal success. We call it that "Exploration-Exploitation Dilemma" [2-5]

When the determined policy has been given as follows:

Deterministic: function that maps states to actions.

$$\pi : S \rightarrow A \; a = \pi(s).$$

Examples: off-policy in Q-learning:

$$a_t = \pi(s_t) = \text{argmax}_a Q(s_t,a).$$

Stochastic: Probability of an action given a state s.
π: S×A→ [0,1] with $\sum_{a\in As} \pi(s,a) = 1$ for all s
P (a|s) = π(a,s)

The on-policy (TD-learning, SARSA) starts with a simple soft rather than off-policy Q-learning collects information from sometimes random moves evaluates states as if a ε-greedy random-policy was employed and reduce randomness very slowly. Therefore, in terms of on-policy, it is rare to use randomness. Because of that when it comes to the high-dimension and low-sample-size data, it is difficult to make advantage of exploration. We can compare the both TD-learning and Q-learning equations as follows: [2, 3, 4, 5]

Q-learning (off-policy): $a_t = \text{argmax}_a Q(s_t,a)$ (plus exploration)
$Q_{t+1}(s_t,a_t) = (1-\eta)Q_t(s_t, a_t) + \eta(r_{t+1} + \gamma \max_a Q_t(s_{t+1},a))$

TD-learning or SARSA (on-policy): $Q_{t+1}(s_t,a_t) = (1-\eta)Q_t(s_t,a_t) + \eta(r_{t+1} + \gamma Q_t(s_{t+1},a_{t+1}))$

Q-learning follows the rule: $V(s_{t+1})=\max_a Q(s_{t+1},a)$, however, a_{t+1} can be anything. That is exploration. SARSA follows the rule: $a_t \sim \pi(s_t,\cdot)$ and updates the rule it learns by the precise value for π(s,a). That means that it does not make an advantage of exploration [2-5]

Initialize $Q(s,a)$ arbitrarily
Repeat (for each episode):
  Initialize s
  Repeat (for each step of episode):
    Choose a from s using policy derived from Q
    Take action a, observe r, s'
    Update
      $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$
    $s \leftarrow s'$;
  Until s is terminal

(그림 1) off-policy Q-learning 알고리즘
(Figure 1) off-policy Q learning [2, 3, 4, 5]

2.2 SVM-RFE algorithm

Guyon et al. proposes a feature pattern selection, SVM-RFE [7]. The purpose is to find a distinguishable subset among variables of feature pattern, which maximizes the performance of the prediction method. It is based on a backward sequential selection. One starts with all the features and removes one feature per one loop. In some research works, because of the large amount of feature genes, some chunks of features will be removed until the distinguishable features are left. When facing highly dimensional and the low size of samples, classification or prediction problem suffers from over-fitting and high-variance gradients [8]. However, some machine learning algorithms likewise SVM-RFE can make good results on the low size of samples with low-variance gradients. SVM-RFE removes irrelevant the gene that is the smallest ranking criteria from the gene set [7]. The criteria of SVM for the score of gene ranking is used as the measurement of the determinant of featured genes. The weight vector w of the SVM defines the gene ranking score,

where w is calculated as.

$$w = \sum_{i=1}^{n} a_i x_i y_i \quad (1)$$

where x_i is the gene expression array of a sample i in the training set, y_i is the class label of i, $y_i \in$[-1, 1] and a_i is the "Lagrangian Multiplier". With a non-zero weight of vectors, a_i are support vectors [7].

```
Algorithm:SVM-RFE
Input: gene set, G={1,2,...n},
Output: gene list for classification based on the ranking criterion, R
1. Initialization Set G={}
2. Do while if G is not empty
      Train SVM in G
      Compute the weight vector by eq(1)
Compute the ranking criterion, CR=w2
      Rank, R the features by sorting based on CR
      Update feature ranked list, FRList based on R
      Eliminate the feature based on R
3. Return the feature ranked list, FRList
```

(그림 2) SVM-RFE 알고리즘 R-언어 구현
(Figure 2) The implementation of SVM-RFE Algorithm in R

The removed variable of SVM-RFE is significantly important. In the method, the removal minimizes the variation of $\|w\|^2$. Hence, the ranking criterion R_c for a given variable i is:

$$R_c = \left| \|w^2\| - \|w^{(i)}\|^2 \right| =$$
$$1/2 \left| \sum_{k,j} a_k a_j y_k y_j (x_k)^T (x_j) - \sum_{k,j} a_k^{(i)} a_j^{(i)} y_k y_j (x_k^{(i)})^T (x_j^{(i)}) \right| \quad (2)$$

where x_k are training examples and y_k are class labels. The algorithm consists in first mapping x into a high dimensional space via a function F [11]. By maximizing the distance between the set of points $F(x_k)$ and the hyper-plane parameterized from (w;b), where w is weight vector and b is bias, while being consistent on the training set. The solution is determined by the Lagrangian theory where, a_k is the solution of the following quadratic optimization problem and $<F(x_k);F(x_l)>$ is the "Gram matrix" of the training examples [11].

2.3 Off-policy Q learning

Off-policy Q-learning collects information from sometimes random moves evaluates states as if a εgreedy random-policy was employed and reduce randomness very slowly. Q-learning is off-policy TD Control. That means Q-learning is trained how to exploit action-value function, Q, and directly approximates the optimal action-value function, while being independent of the policy being followed [2-5].

$$Q(s_t,a_t) \leftarrow (s_t,a_t)+ \alpha(r_{t+1}+ \gamma max_a \cdot Q(s_{t+1},a') - Q(s_t,a_t)) \quad (3)$$

Off-policy Q-learning evaluates one policy while obeying another, for example, to evaluate the greedy policy, as kown as εgreedy. That means it makes advantage of more exploratory scheme. The off-policy is utilized for behavior should be soft and may be slower, but remains more flexible if alternative ways appear [2-5].

3. THE PROPOSED ALGORITHM

In the proposed algorithm Fig 3, we SVM-RFE with off-policy Q-learning in reinforcement learning for variable selection of feature pattern in some applications. There are some recent research works of variable selection of feature pattern based on SVM. For the criterion of Rakotomamonjy et al. [11], they utilize a gradient descent by the derivatives of $\|w\|^2$ with regard to a scaling array-vector associated to variables. And in [12], the SVM-RFE and gradient of $\|w\|^2$ are fundamentally identified as they have the same ranking criteria. However in our proposed algorithm based on the ε -greedy, the ranking criterion has been slightly effected in an iterative way.

Algorithm: SVM-RFE with off-policy Q-learning
Input: GeneSet, G = {1, 2, à n},
Output: Ordered List based on the criterion, FRList

1. Initialize GeneSetNormal and GeneSetCancer
2. Initialize Q(g, a), ∀g ∈ G, a ∈ A(g), hyper-parameters in A(s) and Q(terminal-state)
3. Do while if NewGene, g^i is not empty
   Choose a^i from g^i using policy derived from Q
   \\ for random, (ϵ-greedy)
   Take action a^i, observe r^{i+1}, g^{i+1}
   $Q(g^i,a^i) = Q(g^i,a^i) +$
   $\alpha[r^{i+1} + \gamma max_a Q(g^{i+1}, a') \quad Q(g^i,a^i)]$
   G=G'
   Train SVM in g^i
   Compute the Ranking Criterion with one of A,
   CR= $|\ \|w^2\| - \|w^{(i)}\|^2\ |$
   Sort CR ; Update FRList ; Eliminate genes on CR
4. Return the feature ranked list, FRList

(그림 3) 제안하는 Q-learning으로 향상된 SVM-RFE 알고리즘

(Figure 3) The proposed algorithm combining SVM-RFE [1] with off-policy Q learning [2-5]

In our proposed SVM-RFE with Q-learning, SVM has been trained in each iteration, depending on different sets of genes, G, because of randomness of εpolicy. In that G, the action policy has been improving for selecting more differential-expressed genes. Moreover, the action policy can be improved by back propagation of the gradient descent. There has been many state-of-the-art techniques in which hyper-parameters such as Lagrangian multiplier or a learning rate has been selected by the system designers. However, in our proposed algorithm, we try to eliminate that flaws of the state-of-the-art technique and give a good change to over-fitting problem by only learning the action policy of off-policy Q-learning in reinforcement learning.

Normally, in terms of "High-Dimension and Low-Samples" [13] like gene expressions array data, the on-policy (TD-learning or SARSA) are regarded to more better solutions which can evaluate or improve the policy used to make decisions often using soft action choice, i.e. $\pi(s,a) > 0 \ \forall a$, commit to always exploring and try to find the best policy that still explores. However, it could become trapped in local minima [2-5]. Because of those being trapped issues, we

choose the off-policy Q-learning for evaluating one policy while following another, for example, trying to evaluate the greedy policy as known as ε-greedy while following a more exploratory scheme. The rules for behaviour like that randomness should be soft policies and not be sufficient and be slower [2-5]. However, the rule remains more adaptable if alternative ways appear, then it might lead to a better result following the greediness [2-5]. Therefore, we decide the use of off-policy Q-learning for more enhancement of the weight vectors of SVM-RFE.

4. Performance Evaluation

In recent studies, they claim that "for feature selection on the gene expression", it is extremely important to select as less as significant distinguishable subsets for better understanding and validation [7-10]. The smaller selected, the better claimed. However, the smaller features of pattern might not justify the high correlated variable feature target solutions remarkable compared to other methods because of "Curse of Dimension"[6]. Moreover, smaller variable features of pattern might not be discriminated within a significant computational performance. Therefore, we try to discriminate the distinguishable variable selection of feature pattern that describe the complicated gene expressions with regard to the computational strengthen in published gene data, such as colon cancer in Alon et al[14].

Fig. 4 shows the original result of Alon et al. on the ribosomal protein cluster. Fig. 5 shows that comparition of the proposed algorithm, SVM-RFE with enhanced weight vectors by Q-learning and the previous SVM-RFE with enhanced with weight vectors by Welch' T. The results comparing are based on the original result by Alon et al [14]. We describe how many distinguishable variable selection of feature pattern are ranked in the output-list. The all-features of pattern selected from Alon et al[14] might not be in our result and also the previous' result. However, our result of SVM-RFE with Q-learning is a little bit of better than those of the previous SVM-RFE combining Welch' T. We get the "gene U14971" (Human Robosomal Protein S9m RNA), "gene X57691" (40S Robosomal Protein S6) and "gene T58861" (60S Robosomal Protein L30E). However, only two "gene R20593" (60S acidic Robosomal Protein P1) and "gene

Gene number	Sequence	Name
T63591	3′ UTR	60S acidic ribosomal protein P0 (human)
R50158	3′ UTR	*Mus musculus* L36 ribosomal protein*
T52642	3′ UTR	Guanylate kinase homolog (vaccinia virus)
R85464	3′ UTR	ATP synthase lipid-binding protein P2 precursor (human)
X55715	Gene	Human Hums3 mRNA for 40S ribosomal protein s3
T52185	3′ UTR	P17074 40S ribosomal protein
T56934	3′ UTR	*Homo sapiens* alpha NAC mRNA (transcriptional coactivator)
T47144	3′ UTR	JN0549 ribosomal protein YL30
T72879	3′ UTR	60S ribosomal protein L7A (human)
T57633	3′ UTR	40S ribosomal protein S8 (human)
T58861	3′ UTR	60S ribosomal protein L30E (*Kluyveromyces lactis*)
T52015	3′ UTR	Elongation factor 1-gamma (human)
T57619	3′ UTR	40S ribosomal protein S6 (*Nicotiana tabacum*)
T72938	3′ UTR	Ribosomal protein L10*
R02593	3′ UTR	60S acidic ribosomal protein P1 (*Polyorchis penicillatus*)
T48804	3′ UTR	40S ribosomal protein S24 (human)
R01182	3′ UTR	60S ribosomal protein L38 (human)
T61609	3′ UTR	*H. sapiens* gene for ribosomal protein Sa, partial cds*
H77302	3′ UTR	60S ribosomal protein (human)
U14971	Gene	Human ribosomal protein S9 mRNA, complete cds
H54676	3′ UTR	60S ribosomal protein L18A (human)
R86975	3′ UTR	40S ribosomal protein S28 (human)
T51560	3′ UTR	40S ribosomal protein S16 (human)
H09263	3′ UTR	Elongation factor 1-alpha 1 (*H. sapiens*)
T49423	3′ UTR	Breast basic conserved protein 1 (human)
T63484	3′ UTR	Human ornithine decarboxylase antizyme (Oaz) mRNA, complete cds
R02593	3′ UTR	60S acidic ribosomal protein P1 (*P. penicillatus*)
R22197	3′ UTR	60S ribosomal protein L32 (human)
T51496	3′ UTR	60S ribosomal protein L37A (human)

(그림 4) Cell Biology: Alon et al〔14〕의 결과
(Figure 4) The result of Cell Biology: Alon et al〔14〕

Gene Rank	SVM-RFE with off-policy Q-learning	SVM-RFE with Welch's T	Alon et al. [14]
1	J02854	M26383	T63591
2	K03474	T47377	R50518
3	X57351	R62549	T52642
4	T94579	U37012	R85464
5	T47377	T52185	X55715
6	H20709	T62972	T52185
7	U37012	R44884	T56934
8	M59807	J00231	T47144
9	D14812	**T58861**	T72879
10	T88723	R36977	T57633
11	R39465	X63629	T58861
12	T48904	M59807	T52015
13	M65028	D14812	T57619
14	L08044	M20543	T72938
15	T57882	M22382	R02593
16	**U14971**	H08393	T48804
17	X07979	T84051	R01182
18	U14973	T51539	T61609
19	**T57619**	M81651	H77302
20	T98835	R42127	U14971
21	K03001	L06132	H54676
22	M33680	X75208	R86975
23	X02492	**R02593**	T51560
24	H88360	T98555	H09263
25	T52342	D00860	T49423
26	T96832	X17025	T63484
27	T51023	H65355	R02593
28	T72175	X01060	R22197
29	X67325	J04102	T51496
30	**T58861**	T57468	
The number of genes in the rank	3	2	

(그림 5) 제안하는 Q-learninng 으로 강화된 Weight Vector의 SVM-RFE과 Welth' T〔7〕의 Weight Vector에 의한 SVM-RFE의 결과 비교
(Figure 5) Comparition of the proposed algorithm, SVM-RFE with Q-learning and the previous SVM-RFE with Welth' T〔7〕

T58861" (60S Robosomal Protein L30E) are in the result of the previous SVM-RFE combining Welch' T.

Moreover, we found out that the "SVM-RFE" itself is well-known for the most acceptable methods in many research areas, because the gene "gene T58861" (60S Robosomal Protein L30E) are in the both result, ours and the previous one. Therefore, we can assure that there are some issues that should be improved for better results in terms of using SVM-RFE itself.

Our recent research, C. Kim[15] is regard to the SVM-RFE enhanced with minimum-redundance maximum-relevance (MRMR). The results of the research [11] are regard to "how many distinguishable features in the same places in the rank list". The research [15] makes advantages of only machine learning without enhancing weigh vectors by the reinforcement learning. Therefore, we can find out that the proposed algorithm can improve the computational performance by enhancing the previous SVM-RFE with MRMR[7] using enhanced weight vector by Q-learning for a better qualified learning algorithms. Morever, based on the recent works of reinforcement learning [16, 17], we will improve our results on ribosomal protein cluster [14].

5. Conclusion

We have suggested a novel algorithm of exploiting the efficiency of criteria derived from Support Vector Machines-Recursive Feature Elimination (SVM-RFE) [1] with off-policy Q-learning in reinforcement learning [2-5] for variable feature selection in application to differential-expressed genes of RNA-seq Big-data. We employ an off-policy Q-learning in reinforcement learning to learn how to control the optimization of the criteria based on the weight vectors of the SVM-RFE. We exploit a gradient descent by the derivatives of $\left| \|w^2\| - \|w(i)\|^2 \right|$ and $\max_a Q(s_t,a)$+ exploration scheme. The ranking criterion based on the εgreedy has been slightly effected in an iterative way of off-policy Q-learning. The SVM of our proposed algorithm has been trained according to different sets of G because of randomness of εgreey. In that G, the action policy has been improving for selecting more differential-expressed genes by back propagation of gradient descent. Our proposed algorithm try to eliminate the over-fitting problem by learning the action policy of off-policy Q-learning in reinforcement learning. By comparing our proposed algorithm with the previous SVM-RFE combining Welch' T [7], we can show that the criterion based on weight vector of SVM-RFE can be improved by the greedy policy following a more exploratory scheme of off-policy Q-learning.

참고문헌(Reference)

[1] I. Guyon, J. Weston, S. Barnhill, V. Vapnik: Gene selection for cancer classification using support vector machine. Mach. Learn Vol 46, pp. 389-422, 2002. https://doi.org/10.1023/A:1012487302797

[2] D. Silver, A. Huang, C. J. Maddison, A.Guez, L.t Sifre, G. V. D. Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. Nature, Vol 529, No. 7587, pp. 484-489, 2016. http://dx.doi.org/10.1038/nature16961

[3] R. S. Sutton. Learning to predict by the methods of temporal differences. Machine learning, Vol 3, No. 1, pp. 9-44, 1988. https://doi.org/10.1007/BF00115009

[4] R. S. Sutton, A. G. Barto. Reinforcement learning: An introduction, volume 1. MIT press Cambridge, 1998. https://doi.org/10.1016/S1364-6613(99)01331-5

[5] V. Mnih , K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing atari with deep reinforcement learning. NIPS 2013. http://arxiv.org/abs/1312.5602

[6] T. Tieleman and G. Hinton. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, Vol 4, No. 2, 2012. http://www.citeulike.org/user/ppolon/article/13997430

[7] S. Hansen Using Deep Q-Learning to Control Optimization Hyperparameters, Published 2016 in ArXiv. http://arXiv:1602.04062

[8] Z. Zhou, J. Wang, Y. Wang, Z. Zhu, J. Du, X. Liu and J. Quan, "Visual Tracking Using Improved Multiple Instance Learning with Co-training Framework for Moving Robot," KSII Transactions on Internet and Information Systems, vol. 12, no. 11, pp. 5496-5521, 2018.

https://doi.org/10.3837/tiis.2018.11.018

[9] D. Zhao, B. Guo and Y. Yan, "A Sparse Target Matrix Generation Based Unsupervised Feature Learning Algorithm for Image Classification," KSII Transactions on Internet and Information Systems, vol. 12, no. 6, pp. 2806-2825, 2018. http://dx.doi.org/10.3837/tiis.2018.06.020

[10] M. u Qiao, Haitao Zhao, Shengchun Huang, Li Zhou and Shan Wang, "An Intelligent MAC Protocol Selection Method based on Machine Learning in Wireless Sensor Networks," KSII Transactions on Internet and Information Systems, vol. 12, no. 11, pp. 5425-5448, 2018. https://doi.org/10.3837/tiis.2018.11.014

[11] A. Rakotomamonjy, "Variable selection using SVM-based criteria", J. Machine Learn. Res., Vol 3, pp. 1357-1370, 2003. https://doi.org/10.1162/153244303322753706

[12] P. Leray, P. Gallinari, "Feature selection with neural networks", Behaviormetrika, Vol 26, Jan. 1999. https://doi.org/10.2333/bhmk.26.145

[13] B. Liu, Y. Wei, Y. Zhang, and Q. Yang, "Deep Neural Networks for High Dimension, Low Sample Size Data", IJCAI-17, pp. 2287-2293, Aug., 2017. https://doi.org/10.24963/ijcai.2017/318

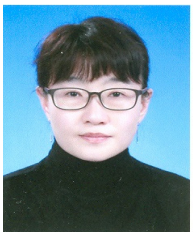
[14] U. Alon, N. Barkai, D. A. Notterman, K. Gish, S. Ybarra, D. Mack, and A. J. Levine, "Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays", Proc. Natl. Acad. Sci. USA Vol. 96, No. 12, pp. 6745–6750, June 1999 Cell Biology, https://doi.org/10.1073/pnas.96.12.6745

[15] C. Kim, "A MA-plot-based Feature Selection by MRMR in SVM-RFE in RNA-Sequencing Data", Journal of KIIT. Vol. 16, No. 12, pp. 25-30, Dec., 2018. https://doi.org/10.14801/jkiit.2018.16.12.25

[16] A. Amiranashvili A. Dosovitskiy V. Koltun and T. Brox, "TD OR NOT TD: ANALYZING THE ROLE OF TEMPORAL DIFFERENCING IN DEEP REINFORCEMENT LEARNING", in ICLR 2018. https://dblp.org/rec/bib/journals/corr/abs-1806-01175

[17] B. Amos, I. Rodriguez, J. Sacks, B. Boots J. and Kolter, "Differentiable MPC for End-to-end Planning and Control", in NeurIPS 2018, https://dblp.org/rec/bib/journals/corr/abs-1810-13400

◐ 저 자 소 개 ◑



김 차 영(Chayoung Kim)
1996년 숙명여자대학교 전산학과(이학사)
1998년 숙명여자대학교 전산학과(이학석사)
2006년 고려대학교 컴퓨터학과(이학박사)
2005년~2008년 한국과학기술정보연구원 선임초청연구원
2008년~2017년 경기대학교, 컴퓨터과학과, 대우교수
2018년~현재 경기대학교, 융합교양대학, 조교수
관심분야 : 빅데이터, 머신러닝, 딥러닝 강화학습, IoT
E-mail : kimcha0@kgu.ac.kr