



Feature Selection of SVM-RFE Combined with a TD Reinforcement Learning

Chayoung Kim*

Abstract

Thanks to the reinforcement learning, self-teaching models are designed for Big-Data on the demands in different kinds of applications, such as distributed massively multi-player online games (MMOGs) or gene regulatory networks (GRNs). We try to consider the data representation of taking advantages of the server platform. Before adopting the software on the servers, the feature selection of the huge amount of data might be helpful for reducing server loads on Big-Data issues. Moreover, the feature selection might be appropriate for the prediction of resource demands in other similar researches. Therefore we propose a feature selection by “support vector machine-recursive feature elimination” (“SVM-RFE”) with time-difference (TD) Reinforcement Learning. The comparison results show that the proposed algorithm might have some improvement in terms of burden issues of servers in Big-Data with high relevance features by the self-teaching models.

요약

자기-학습 모델의 발전은 강화학습 덕분에 다양한 방법으로 진행되고 있는데, 해당 분야의 예를 들면, 대규모 온라인 게임 서버 구조와 유전자 조절 네트워크 같은 빅-데이터 분야이다. 본 연구는 해당 분야의 서버 플랫폼을 잘 활용하기 위하여, 데이터 표현 자체에 대한 분석에 더 높은 비중을 두어 진행하였다. 첨가하여, 이러한 연구 포커스가 대용량 처리를 하는 소프트웨어를 적용하기 전에, 빅-데이터 분야의 특징 선택에 대한 서버 로드 이슈를 해결하는 데에 도움이 될 것이며, 비슷한 다른 연구 분야에 대한 결과 예측에도 잘 활용될 것이라고 생각한다. 본 연구는 SVM-RFE에 TD 강화학습을 적용한 방법을 제안한다. 본 연구와 기존 연구 결과를 비교한 것을 보면, 빅-데이터 분야에서 서버의 로드를 향상시켰으며, 자기-학습 모델에 의한 관련성 높은 특징만을 선택함을 볼 수 있다.

Keywords

SVM-RFE, TD reinforcement learning, Big-Data, feature selection, self-teaching

* Div. of General Studies, Assistant Professor
- ORCID: <https://orcid.org/0000-0002-4186-5882>

· Received: Sep. 14, 2018, Revised: Oct. 14, 2018, Accepted: Oct. 17, 2018
· Corresponding Author: Chayoung Kim
Div. of General Studies, 154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Korea, 16227.
Tel.: +82-31-249-9509, Email: kimcha0@kgu.ac.kr

I. Introduction

Many discrimination applications, such as the architecture of massively multi-player online games or gene regulatory networks for differential-expressed genes (DEGs) of RNA-seq infer knowledge from a huge amount of data, Big-Data[1]. There have been main existing issues such as providing guarantees consistency on simultaneous participants of the MMOGs or on the intermediate causality on the GRNs. This knowledge is used to obtain a deeply untestable and relevant meaningful concept that generated the huge data. For example, in terms of Big-Data of DEGs of GRNs, they are consisting of measurements also known as to the featured genes, which is characterizing a basis for the mechanisms through inferring causality between biologically relevant genes[2]-[4].

Also, Interaction information brought into the distributed servers between the huge numbers of participants of MMOGs might be considerable and mainly responsible for encompassing a Big-Data world. They require new criteria for load balancing because the huge amount of history information might be enhancing the most relevant and less redundant form. Based on the measurements associated with a label specifying the category might be estimated dependencies between the examples and their labels by machine learning models[5]-[8].

However, the RNA-seq let researchers gather data sets of ever increasing sizes and the servers of MMOGs learn each other by interactions and bring the huge world, which is not always significant [2]-[4]. This issues might compromise the machine learning models, which is extremely influenced by data qualification, such as redundancy, noise or undermined unreliability. The purpose of feature selection is to try to remove irrelevancies and minimize redundancy patterns to improve the generalizations of the machine learning models[5]-[8]. In terms of RNA-seq, the maximum gene relevancy and minimum gene

redundancy is very significant and essential[2]-[4].

Therefore, we try to consider the data feature representation for taking advantages of the server platform of the MMOGs. Thanks to the reinforcement learning [5][6], the self-teaching models are designed for Big-Data on the above-mentioned issues in different kinds of applications, such as MMOGs or GRNs. Before adopting the software on the servers of MMOGs, the feature selection on the huge amount of data might be helpful for reducing server loads on Big-Data issues. Moreover, the feature selection might be appropriate for the prediction of resource demands in other similar researches. Moreover, we think our proposed methods based on reinforcement learning is beyond the qualification of the server platform of MMOGs because enhanced selection of the data makes the enriched meaningful features.

Therefore, we propose a novel algorithm of exploiting the improved effectiveness of criteria derived from “support vector machines-recursive feature elimination” (“SVM-RFE”) [2] combined with temporal-difference (TD) reinforcement learning[5][6] for feature selection in application to differential-expressed genes (DEGs) of RNA-seq. In the application of MMOGs, our proposed model can learn the patterns (features) of past accumulated server loads. By learning the overhead features, the server can take actions promptly to achieve better supports in advance. We employ a TD reinforcement learning to learn how to control the weight vectors of the SVM-RFE. Our experiment results show that the criteria of the weight vector improved by the back-propagation of TD errors accomplishes better results and performs well over the Big-Data sets.

II. Materials and Methods

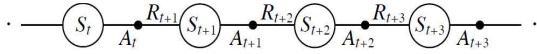
2.1 Data

The well-known statistical analysis, the “MA-plot” has been widely used to visualize and perceive the

intensity-dependent ratio of micro-array data[2]-[4]. For comparisons our results with the well-known leukemia data-sets [3] from their web-sites. It was analyzed by the previous developed “SVM-RFE” [2] and the “MA-plot” [2]-[4].

2.2 Methods

The TD, (as known as SARSA) [5][6] starts with a simple policy, samples state spaces with and improves the policy. SARSA means: “S(tate) → A(ction) → R(eward) → S(tate) → A(ction)”



TD prediction is for a control i.e action-selection, generalization of the iteration of the policy and balances between exploration and exploitation. The Q function is that “ $Q(s^t, a^t) \leftarrow (s^t, a^t) + \alpha(r^{t+1} + \gamma Q(s^{t+1}, a^{t+1}) - Q(s^t, a^t))$ ”. After per transition from a non-terminal state s^t , this Q function is updated. If s^{t+1} is a terminal, then the zero is defined in the “ $Q(s^{t+1}, a^{t+1})$ ”. This update rule is defined in every element of “ $(s^t, a^t, r^{t+1}, s^{t+1}, a^{t+1})$ ” [5][6].

The TD reinforcement learning algorithm attempt to estimate or enhance the policy, which is defined to make better action-selection, known as the on-policy. The on-policy often uses soft action choice, i.e. $\pi(s; a) > 0 \forall a$, performs for the exploration and tries to find out the best policy that develops. It might be trapped into local minima. The on-policy TD reinforcement learning means that while learning the best optimal policy per iteration, it uses the current estimate of the best optimal policy to engender the behavior for the exploration and converges to an optimal policy, while all “S(tate) → A(ction)” pairs have been visited a limited times and the best policy converges into the limit with the greedy policy (“ $\epsilon = 1/t$ ”)[5][6].

```

Initialize  $Q(s, a), \forall s \in \mathcal{S}, a \in \mathcal{A}(s)$ , arbitrarily, and  $Q(\text{terminal-state}, \cdot) = 0$ 
Repeat (for each episode):
  Initialize  $S$ 
  Choose  $A$  from  $S$  using policy derived from  $Q$  (e.g.,  $\epsilon$ -greedy)
  Repeat (for each step of episode):
    Take action  $A$ , observe  $R, S'$ 
    Choose  $A'$  from  $S'$  using policy derived from  $Q$  (e.g.,  $\epsilon$ -greedy)
     $Q(S, A) \leftarrow Q(S, A) + \alpha[R + \gamma Q(S', A') - Q(S, A)]$ 
     $S \leftarrow S'; A \leftarrow A'$ ;
  until  $S$  is terminal
    
```

Fig. 1. On-policy TD control algorithm

2.3 SVM-RFE

The “SVM-RFE” algorithm has been proposed by Guyon et al. for selecting featured patterns that are highly relevant for a cancer classification. The goal is to find a smallest subset of size, r among feature patterns which maximizes for the performance of the predication[2].

III. Proposed Algorithm

We propose “SVM-RFE” with the TD reinforcement learning for feature selection in some applications. There is the proposed algorithm in Fig 2. Because the choice of criteria is depending on the parameters of “SVM-RFE”, it could be necessary to train the TD reinforcement learning with different sets of feature patterns. The TD reinforcement learning tries to learn how to control the criteria based on the weight vectors of the “SVM-RFE”. For the criterion of Leray et al. [9], they use a back-propagation (gradient descent) by the derivatives of $\|w\|^2$ with respect to scaling vector associated to variables[7].

The SVM - RFE and gradient of $\|w\|^2$ are as same as they have the identical ranking criterion [9]. For previous developed feature selection algorithms, choice criteria are subject to the human-interaction(heuristic individual) feature evaluation functions such as hyper-parameters. Several methods use individual evaluation of the features for ranking them and do not take into considerations for their dependencies or their correlations[2][4].

<p>Algorithm: SVM-RFE with the TD reinforcement learning</p> <p>Input: GeneSet, S ,</p> <p>Output: Ranked Gene List based on the criterion, FRList</p> <hr/> <ol style="list-style-type: none"> 1. Initialize GeneSetNormal and GeneSetCancer 2. Initialize $Q(s, a)$, $\forall s \in S, a \in A(s)$, hyper-parameters in $A(s)$ and $Q(\text{terminal-state})$ 3. Choose A from S using policy derived from Q (ϵ-greedy) 4. Do while if NewGene, S' is not empty <ul style="list-style-type: none"> Take action A, observe R, S' Choose A' from S' using policy derived from Q (ϵ-greedy) $Q(S,A) = Q(S,A) + \alpha[R + \gamma Q(S', A') - Q(S,A)]$ Train SVM in S' Compute the Ranking Criterion with a of A', $CR = \frac{1}{2} (\ w^2\ - \ w^{(1)}\ ^2)$ S=S' ; A=A' Sort CR Update FRList Eliminate genes on CR 5. Return the ranked list, FRList

Fig. 2. Proposed algorithm

This may be problematic for selecting discriminate relevant sets of feature patterns. In our proposed “SVM-RFE” with TD reinforcement learning, SVM has been trained in each iteration, depending on different sets of action reinforcement. The reinforcement has been improving for selecting more differential-expressed genes by back propagation of gradient descent. In our proposed algorithm, we try to eliminate that flaws of the heuristic technique and give a good change to over-fitting problem by learning TD.

IV. Evaluation of the proposed algorithm

For feature selection on the gene expression micro-array data, it is informative to select a minimal relevant gene subset. The topmost-ranked feature genes might not be the most relevant featured genes.

Although the “SVM-RFE” is known for the correlation-based, some of the previous results are

unstable. Currently, that is known that there is smaller is the best featured genes with gene samples. However, because of small size of samples and high correlation among genes, finding the smallest relevant gene subset might not justify the superiority of the gene selection algorithm comparing with others. That is because of “Curse of Dimension”[4].

The smaller relevant subsets might not be extracted in a controllable computation limits. In the controllable limits, we might select the featured subsets that describe the complicated cancerous reaction and regulation in “AML/ALL” leukemia data set[3]. Some results might cover the cancer mechanisms or some might not. The proposed “SVM-RFE” with the TD reinforcement learning is comparing with the previous developed “SVM-RFE” based on the results on Golub. et. al.[3].

The “AML/ALL” leukemia consists of 72 samples (47 for ALL and 25 for AML). In the “AML/ALL”, there are 7,129 feature genes. We try to describe how many featured genes are almost the same.

Fig. 3 shows that the featured genes of “SVM-RFE” with the TD reinforcement learning, “SVM-RFE” [2] and the ranked list of Golub. et. al.[3]. The all genes selected in “AML/ALL” leukemia [3] are not the same with our result. However gene, M27891 (Cystatin), Y00787(IL-8 precursor), M63138(cathepsin D), M92287(cyclin D3), X17042(Proteoglycan), M28130 (IL-8), X95735(Zyxin), M69043(MAD-3), and M96326 (Azurocidin) are in the relevant topmost 25. This result indicates that the proposed algorithm can improve the performance by using TD reinforcement learning prediction in terms of generalization of the iteration method and balances between exploration and exploitation. It also was proved that the smaller is the best featured genes with gene samples in controllable computational limits with the accuracy in smaller size of samples and high correlation among genes. Fig. 3 shows that the feature genes of “SVM-RFE” with TD reinforcement learning is considerably comparative with Golub. et. al.[3].

Gene Rank	SVM-RFE	SVM-RFE with TD reinforcement learning	ALL[2]	AML[2]
1	M27891	M27891	D26156	L08246
2	M17733	M19507	D38073	M16038
3	Y00787	Y00787	L47738	M19045
4	M19507	M63138	M13792	M23197
5	U05255	U05255	M29696	M27891
6	Z19554	M33680	M31211	M28130
7	Z23090	M11722	M31303	M55150
8	M69043	J04164	M31523	M57710
9	M11147	M77232	M91432	M62762
10	hum_alu	M11147	M92287	M63138
11	U68105	M92287	S50223	M69043
12	M27783	M17733	U05259	M80254
13	M23613	X51466	U20998	M81695
14	M92287	X17042	U22376	M83652
15	J04088	HG3549-HT3751	U26266	M84526
16	M11722	M28130	U29175	M96326
17	X59417	U51240	U32944	U46751
18	M14328	M14328	U35451	U50136
19	L38941	X95735	X15949	U82759
20	L20941	M26708	X59417	X04085
21	M77232	X97267	X63469	X17042
22	M96326	M91036	X74262	X85116
23	M28130	Z19554	Y08612	X95735
24	U01317	M69043	Z15115	Y00787
25	M95678	M96326	Z69881	Y12670
The number of genes in high rank	7	9		

Fig. 3. Comparison of the proposed algorithm with SVM-RFE based on Golub

V. Conclusion

We propose the algorithm to select some featured genes in a controllable computation and to ensure the training model before evaluation for the relevancy of featured genes. This makes to consider that the training model are explored and enhanced by featured genes. We exploit the efficiency of criteria derived from “support vector machines-recursive feature elimination” (“SVM-RFE”) [2] combined with the TD reinforcement learning [5][6] for feature selection in application to differential-expressed genes of RNA-seq. We employ the TD reinforcement learning to learn how to control the criteria based on the weight vectors of the “SVM-RFE”. We exploit a gradient descent by the derivatives of the weight vectors of “SVM-RFE”.

The ranking criterion based on the ϵ -greedy has been slightly effected in an iterative way of TD reinforcement learning. The SVM has been trained in each iteration, depending on different sets of reinforcement leaning policy. The reinforcement leaning policy has been improving for selecting more differential-expressed genes by back propagation of gradient descent. We try to eliminate that flaws of the heuristic technique and give a good change to over-fitting problem by learning the learning policy of TD reinforcement learning. The result show that the reinforced criteria on the weight vector improved by the back-propagation of TD errors acquires better results.

References

- [1] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. V. D. Drissi, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search", Nature, Vol. 529, No. 7587, pp. 484-489, Jan. 2016.
- [2] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machine", Mach. Learn, Vol. 46, pp. 389-422, Jan. 2002.
- [3] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander, "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring", Science, Vol. 286, No. 5439, pp. 531-537, Oct. 1999.
- [4] Y. Tang, Y. Q. Zhang, and Z. Huang, "Development of Two-Stage SVM-RFE Gene Selection Strategy for Microarray Expression Data

- Analysis", IEEE ACM Transactions on Computational Biology and Bioinformatics, Vol. 4, No. 3, pp. 365-381, Aug. 2007.
- [5] R. S. Sutton, "Learning to predict by the methods of temporal differences", Machine learning, Vol. 3, No. 1, pp. 9-44, Aug. 1988.
- [6] R. S. Sutton, and A. G. Barto, "Reinforcement learning", An introduction, MIT press Cambridge, second edition, ISBN: 9780262039246, 2018.
- [7] 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, pp. 26-31, <https://www.youtube.com/watch?v=LGA-gRkLEsI>, 2012.
- [8] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning", arXiv preprint arXiv:1312.5602, NIPS 2013, 2013.
- [9] P. Leray and P. Gallinari, "Feature selection with neural networks", Behaviormetrika, Vol. 26, pp. 6-16, Jan. 1999.

Author

Chayoung Kim



Feb., 2006 : Ph.D., Computer Science, Korea University
Nov., 2005. ~ Dec., 2008. : Senior Project Researcher, KISTI
Mar., 2009. ~ Feb., 2018. : Adjunct Professor, Dept. of Computer Science, Kyonggi

University

Mar. 2018. ~ present : Assistant Professor, Div. of General Studies, Kyonggi University

Reserach Interest : Big-Data, Machine Learning, Deep Learning, Reinforcement Learning, Development of AI game