ISSN 1598-0170 (Print)
ISSN 2287-1136 (Online)
http://www.jksii.or.kr

RNAseq 빅데이터에서 유전자 선택을 위한 밀집도-의존 정규화 기반의 서포트-벡터 머신 병합법☆

Combining Support Vector Machine Recursive Feature Elimination and Intensity-dependent Normalization for Gene Selection in RNAseq

김 차 영[1*]
Chayoung Kim

요 약

고처리 시퀀싱과 빅데이터 및 크라우드 컴퓨팅에 혁신이 일어나면서, RNA 시퀀싱도 획기적인 변화가 일어, RNAseq 가 기존의 DNA 마이크로어레이를 대체하여, 빅-데이터를 형성하고 있다. 현재, RANseq 이용한 유전자 조절망(GRN) 까지 연구가 활성화 되고 있는데, 그 중 한 분야가 GRN의 기본 요소인 특징 유전자를 빅-데이터에서도 구별하고 기존에 알려진 것 외에 새로운 역할을 찾는 것이다. 그러나, 이러한 연구 방향에 부합하는 빅-데이터를 처리할 수 있는 컴퓨테이션 방법이 아직까지 매우 부족하다. 따라서 본 논문에서는 RNAseq 빅-데이터를 처리할 수 있도록 기존의 SVM-RFE알고리즘을 밀집도-의존 정규화에 병합하여, NCBI-GEO와 같은 빅-데이터에서 공개된 일부의 데이터에 개선된 알고리즘을 적용하고 해당 알고리즘에 의해 나온 결과의 성능을 평가한다.

☞ 주제어 : 서포트-벡터-머신, RNA시퀀스, 빅-데이터, 밀집도-의존 정규화, SVM-RFE알고리즘

ABSTRACT

In past few years, high-throughput sequencing, big-data generation, cloud computing, and computational biology are revolutionary. RNA sequencing is emerging as an attractive alternative to DNA microarrays. And the methods for constructing Gene Regulatory Network (GRN) from RNA-Seq are extremely lacking and urgently required. Because GRN has obtained substantial observation from genomics and bioinformatics, an elementary requirement of the GRN has been to maximize distinguishable genes. Despite of RNA sequencing techniques to generate a big amount of data, there are few computational methods to exploit the huge amount of the big data. Therefore, we have suggested a novel gene selection algorithm combining Support Vector Machines and Intensity-dependent normalization, which uses log differential expression ratio in RNAseq. It is an extended variation of support vector machine recursive feature elimination (SVM-RFE) algorithm. This algorithm accomplishes minimum relevancy with subsets of Big-Data, such as NCBI-GEO. The proposed algorithm was compared to the existing one which uses gene expression profiling DNA microarrays. It finds that the proposed algorithm have provided as convenient and quick method than previous because it uses all functions in R package and have more improvement with regard to the classification accuracy based on gene ontology and time consuming in terms of Big-Data. The comparison was performed based on the number of genes selected in RNAseq Big-Data.

☞ keyword : Support-Vector Machine, RNAseq, Big-Data, Intensity-dependent Normalization, SVM-RFE

1. Introduction

Most systems biology experiments have some important objectives, which are to identify differentially expressed genes (DEG) in the experimental results and build gene regulatory networks (GRN) to provide qualitative and quantitative models for reviewing the intricate patterns of gene interaction[1,2]. But, the complex patterns of gene expression might evoke specific cellular activities at different levels. The elementary requirement of GRN has been to minimize relevant and maximize distinguishable genes to provide a basis for molecular mechanisms through inferring causality relationships. We call it that as gene selection to produce biologically relevant and extractable data from DNA Microarray and cDNA (RNA-seq). RNA-seq is a revolutionary DNA sequencing technology recently developed that provides a high throughput method for cDNA

1 Dept. of Computer Science, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon, Gyeonggi, Korea
* Corresponding author (kimcha0@kgu.ac.kr)
[Received 2 June 2017, Reviewed 12 June 2017(R2 28 July 2017), Accepted 3 August 2017]
☆ A preliminary version of this paper was presented at ICONI 2016 and was selected as an outstanding paper.

sequencing, generating information about mRNA content [3]. Despite of RNA sequencing techniques to generate a big amount of data, there are few computational methods to exploit the huge amount of data. Those of methods for constructing GRN from RNA-seq as a Big-data are extremely lacking and urgently required. The DNA microarray technology has provided us several prospects to identify differentially expressed genes. But it has been recalcitrant and extremely challengeable to select a small subset of highly relevant genes because of the high dimensional biological data, where the number of genes is far larger than that of samples. Guyon et al. [4] proposed support vector machine recursive feature elimination (SVM-RFE) algorithm to recursively remove genes based on their weights and classify the samples with SVM. SVM-RFE approach for gene selection has recently attracted many researchers [5,6]. RNA-seq can be regarded as an attractive approach to replace DNA microarray for analyzing genotypes and identifying transcript factors in a comprehensive manner. However, there are few convenient computational tools for identifying differentially expressed genes from RNA-seq although some recent publications have described their methods for the tasks [1,2,3]. And there are also some computational approaches have been proposed to select genes only using statistical properties of the data without any learning model. In [4], only some statistical methods including generalized linear model likelihood ratio test have validated the results of two significantly activate regulators. It is more popular that the methods evaluate the fitness of subset of selected genes iteratively by a specific learning classifier model, such as SVM.

Therefore, we have suggested a novel gene selection algorithm combining Support Vector Machines and Intensity-dependent normalization, which uses log differential expression ratio (Minus vs Add plot, MA-plot) in RNA-Seq[2]. Here, we exploit DEGseq, a free R package for the differentially expressed genes based on MA-plot methods with random sampling model or technical replicates. The input of DEGseq is uniquely mapped reads from RNA-seq[2]. The output includes the expression values for the samples, a P-value and two kinds of Q-values for each gene to denote its expression difference.

Some literatures [5, 6] develop some alternative algorithms based on SVM-RFE[4], to overcome consuming a huge amount of training time and the problem of over-fitting persist and eliminating only one gene at each iteration to improve the accuracy and narrow down the potential set of cancerous genes. Most of the target methods of that algorithms are combined some statistical test, such as t-statistic [comb], Welch's t-test [5], Bayesian t-test [6], which are incorporated for maximum classification accuracy or formed with two-stage for ensemble, which can be more reliable prediction. But, the proposed algorithm can be exploited right before SVM-RFE to reduce the size of the number of the potentially distinguishable genes. If those algorithms [4-6] might use our proposed algorithm, it can be essential prerequisite for better time consumption and more accurate and reliable because the output of DEGseq can be the input of them. And the proposed our algorithm can be convenient and quick than previous, because it uses all functions in R package. It have more improvement with regard to the classification accuracy based on gene ontology and time consuming in terms of Big-Data. Also, we have some experiments with subsets of Big-Data, such as NCBI-GEO[7,8]. We compared the results of the proposed algorithm with one of the existing algorithms which uses gene expression profiling DNA microarrays. The comparison was quite comparative based on the number of genes selected in Big-Data than the previous because those algorithms [4,5,6] might not consider the case of a huge amount of data. We can find that it can be accomplishing minimum relevancy especially with Big-Data.

2. MATRIALS AND METHODS

2.1 MOTIVATION

For comparisons of our experiment with a well-known result of a Microarray-based technology [9, 10, 11], the publicly available microarray datasets, leukemia [7] were downloaded from their websites. This set evaluated with combining SVM-RFE and MA-plot-based methods by using R-package 'DEGseq' for the comparison. [7] was assayed using Affymetrix Hgu6800 chips. Also, we downloaded the publicly avaialble colon RNA-seq data [8] from Gene Expression Omnibus, which was assayed using Illumina HiSeq 2000. RNA-seq data of 54 samples (normal colon,

primary colorectal cancer (CRC), and liver metastasis) from 18 CRC patients are generated in [8] identifying significant genes associated with aggressiveness of CRC for a prognostic signature with diverse progression and heterogeneity. [8] has validated the results by using only diverse statistical methods without a well-known classifier, such as SVM-RFE.

2.2 SVM-RFE algorithm

Guyon et al. [4] proposed a gene selection algorithm, Support Vector Machine-Recursive Feature Elimination (SVM-RFE). It starts with the gene set containing the full genes and removes iteratively the gene that is the smallest ranking criterion from the set. It is trained with a linear kernel and gene ranking score is used as the criteria measuring the significance of the gene for classification. Gene ranking score is defined by the weight vector w of the SVMs, and w is calculated as.

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

where i is the number of genes ranging from 1 to n, x_i is the gene expression vector of a sample i in the training set and y_i is the class label of i, $y_i \in [-1,1]$ and a_i is the Lagrangian Multiplier estimated from the training set. The training vectors with non-zero weights a_i are support vectors. Most weights a_i are zero[4].

```
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 1) SVM-RFE 알고리즘 R-언어 구현

(Figure 1) The implementation of SVM-RFE Algorithm in R

2.3 MA-plot-based method

The MA-plot, which is a statistical analysis tool having been widely used to detect and visualize intensity-dependent ratio of microarray data [2]. The normalization of expressed data adjusts the individual hybridization intensities in two-color (Red/Green) microarray assay to balance them appropriately so that meaningful biological comparisons can be made. In addition normalization, log differential expression ratio (M vs A plot) is widely accepted. To visualize intensity-dependent effects, the locally weighted linear regression analysis (LOWESS) is processed [2] to plot the measured M (log-intensity $\log_2 C_1$) and A (log-intensity $\log_2 C_2$). Local variation as a function of intensity can be used to identify differentially expressed genes by calculating an intensity-dependent Z-score. In this plot, array elements are color-coded depending on whether they are less than one standard deviation from the mean (blue), between one and two standard deviations (red), or more than two standard deviations from the mean (green). That means the genes out of the line ($\log_2 C_1 = \log_2 C_2$) in the plot are identified for carrying out further analyses. Let C_1 and C_2 denote the counts of reads mapped to a specific gene obtained from two samples, with $C_i \sim$ binomial (n_i, p_i), i = 1, 2, where n_i denotes the total number of mapped reads and p_i the probability of a read coming from that gene. We define $M = \log_2 C_1 - \log_2$, and $A = (\log_2 C_1 + \log_2) / 2$. It can be proved that under the random sampling assumption the conditional distribution of M given that A, follows an approximate normal distribution. We set x_i is M and y_i is A for LOWESS analysis. For example, the results in a mean $\log_2$(ratio) equal to zero (M=0) is no changes. M=1($\log_2 2$=1) is doubled. A=2 is 4 times than normal. The implementation of DEGseq, a free R package for this purpose.

3. THE PROPOSED ALGORITHM

In the proposed algorithm Fig 2, before SVM-RFE [4], MA-plot-based method [2] is applied firstly for identifying top most genes, which can be the input list of SVM-RFE for further analyzing.

SVM has been trained in each iteration, depending on different sets of G. It is a state-of-the-art technique but has

the flaws, which is consumption of the high amount of training time, elimination of one gene at a time and overfitting problem. We might get some different ranking criteria for SVM-RFE [4] with DEGseq in R, which is the differentially expressed genes set. The extended version of the SVM-RFE algorithm called SVM-T-RFE [5] that is a conjunction of SVM-RFE [4] and Welch's t-test statistic or SVM-BT-RFE [6] merged Bayesian T-test with the weight vector to produce a new ranking score could be used with DEGseq in R. They are aimed at training the algorithm in a much faster manner by eliminating many a genes at a time.

Therefore, the SVM-BT-RFE [6] also restricts the consumption of time if few numbers of top most genes by our proposed algorithm are considered. In the proposed algorithm, the equations (2) is for SVM-T-RFE [5] and (3) for SVM-BT-RFE [6], respectively.

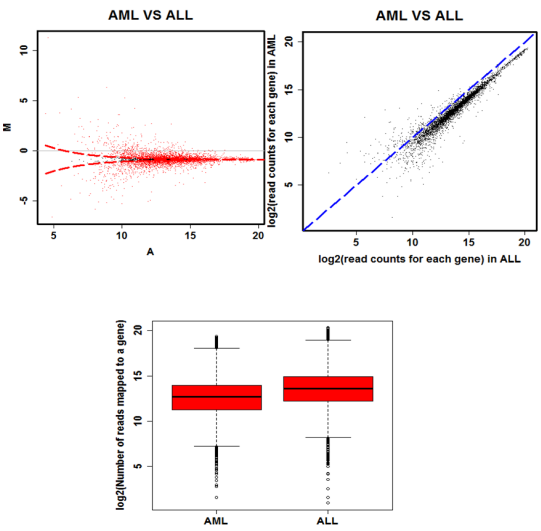
```
Algorithm: DEGseq-SVM-RFE
Input: gene set, G = {1, 2, … n},
Output: gene list for classification based on the
ranking criterion, R
1. Initialization GeneSet1 <-Normal
2. Initialization GeneSet2 <-Cancer
3. Get the output gene list, OL by
Intensity-Dependent Normalization (DEGseq)
4. Cut OL with the threshold, such as
        log2 (fold-change) or p-value or q-value
5. Update G with the output gene list
      New_G=G-G (OL)
2. Do while if New_G is not empty
    Train SVM in New_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) MA-plot기반 방법과 SVM-RFE를 병합한 제안한 알고리즘

(Figure 2) The proposed algorithm combining SVM-RFE [4] with MA-plot-based method [2]

4. Performance Evaluation

We applied DEGseq to compare the samples from leukemia [7], acute myeloid leukemia (AML) and acute lymphocytic leukemia (ALL). Figure 3 shows the summary generated by DEGseq in R. In Figure 3, there are boxplot

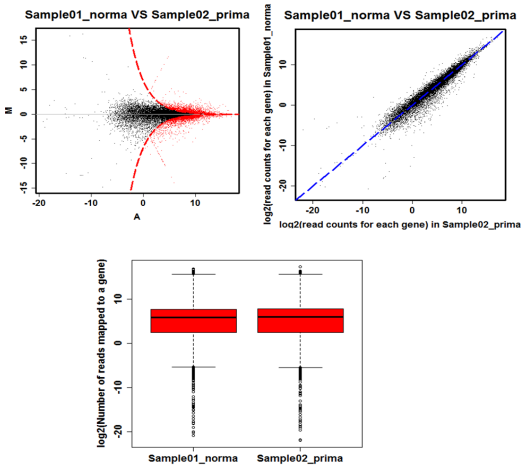


(그림 3) AML과 ALL의 DEGseq 요약

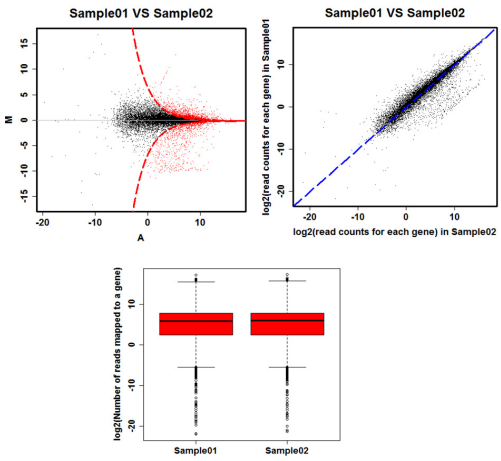
(Figure 3) The summary report page generated by DEGseq in R of AML and ALL [7]

and barplot in terms of figures and fold change log of normalization and Z-score comparing two groups AML and ALL [7]. To visualize intensity-dependent ratio, M (log-intensity $\log_2 C_1$) and A (log-intensity $\log_2 C_2$) can be y-axis and x-axis, respectively. The first of Figure3 is differentially expressed genes(DEG) on the MA-plot[2], the second is scatter plot comparing the number of reads for each gene for AML and ALL and the third is box plot of read counts for each gene for AML and ALL[7].

The output of DEGseq is the ordered differentially expressed genes list based on Z-score. We compared two samples of leukemia [7], where the data set is consist of 7,130 genes with 48 samples for ALL and 26 samples for AML. We need few more data sets taken into consideration for meaningful biological comparisons. For our experiment, grounded on the p-values of the genes whose signature p-values are less than equal to 0.001 are considered as statistically significant genes. On the purpose of bio-scientist, the gene list for SVM-RFE could be made. Therefore, only 4118 genes was taken for the input of the SVM-RFE after DEGseq. When we run the previous SVM-RFE with the whole gene list of leukemia [7], it takes quite a bit of long time consumed. When we run the SVM-RFE with the list



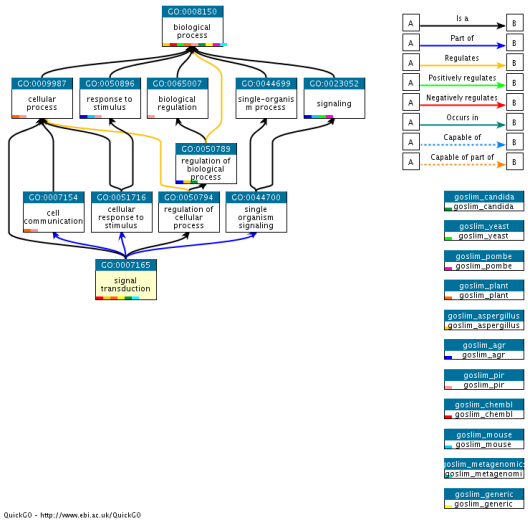
(그림 4)건강한 대장과 초기 대장암의 DEGseq 요약
(Figure 4) The summary report page generated by DEGseq of Normal and Primary Cancer [8]



(그림 5)초기 대장암과 전이 대장암의 DEGseq 요약
(Figure 5) The summary report page generated by DEGseq of Primary and Metastasis Cancer [8]

out of the result of DEGseq, it take a few minutes in terms of R packages. The proposed algorithm have provided as convenient and quick method because it uses all packages in R. And the comparison was outperformed based on the number of genes selected than the previous.

We also applied DEGseq to compare another samples



(그림 6) QUICKGO의 GO:0007165의 유전자 가계도
(Figure 6) The Signaling Pathway Ancestor's Chart of GO:0007165

colon RNA-seq data [8], where the data set is consist of 23,505 genes with 19 samples for normal and 26 samples for cancer. Figure 5 and Figure 6 show the result after DEGseq on Normal and Primary Cancer and on Primary and Metastasis Cancer. As shown that, it is not clear there is meaningful differences between Normal and Primary and Primary and Metastasis. We select the input lists based on p-values for SVM-RFE after DEGseq. In [8], TREM1 and CTGF were identified as two activated regulators associated with CRC aggressiveness.

In Table 1, there is the ordered list on AML and ALL with DEGseq in 4118 genes and without DEGseq in whole 7129 genes. And those on Normal and Primary and Primary and Metastasis, respectively. In the gene list, we can find that the ordered list is different from the ordered list of the Golub et al. [7], which we can see that in red of the 7129 genes. They developed a method called "neighborhood analysis", which is, one defines an "idealized expression pattern" corresponding to a gene that is uniformly high in one class and uniformly low in the other and used a technique called self-organizing maps (SOMs) and cross-validation tools for regression models. Golub et al. [7] did not use SVM-RFE and our proposed algorithm only SVM-RFE. So, we compare the order of top most genes of our proposed algorithm with

(표 1) AML/ALL과 CRC의 중요 순위 유전자 결과
(Table 1) The result comparison on AML/ALL and CRC in yellow and in orange, respectively

	With DEGseq(4118)	Without DEGseq(7129)	Primary Vs. Metastatis	Normal Vs.Primary
1	U90552_s_at	M27891_at	SNAR-B2	CD55
2	U54804_at	M19507_at	MT1E	FITM2
3	Z26634_at	Y00787_s_at	TMC5	MORC4
4	M63379_at	M63138_at	TMCO1	ADH5
5	U90915_at	U05255_s_at	VEGFA	SCG2
6	D49818_at	M33680_at	CXCL14	RPL41
7	J03040_at	M11722_at	HSPB1	POLR1B
8	D87437_at	J04164_at	GBA3	EVL
9	L08835_rna2_s_at	M77232_rna1_at	ABL1	UBE2D3
10	U79294_at	M11147_at	LRIG3	CD247
11	U15009_at	M92287_at	LILRB4	TKT
12	U10473_s_at	M17733_at	PLSCR1	NUDCD1
13	U32581_at	X51466_at	CPB2	KRT19
14	M62958_at	X17042_at	SULT2A1	CHTF18
15	HG4243-HT4513_at	HG3549-HT3751_at	SPTSSA	RAD23A
16	U56418_at	M28130_rna1_s_at	ANO10	FAM82A1
17	U54778_at	U51240_at	SNORA8	RAB15
18	Z12830_at	M14328_s_at	PLGRKT	PPP1R35
19	L36983_at	X95735_at	PLOD2	WDR12
20	HG358-HT358_at	M26708_s_at	APCDD1	SOX9

the those of i Golub et al. [7]. The top most list is different, but all of meaningful genes are in the ordered list by investigating the gene ontology based on QUICKGO, which gives a biological process, synonyms, the ancestor chart and child terms of a go term [12]. In Figure 6, we can see the one of the examples that the ordered list can have the meaningful results. In Table 1, we can see that the ordered list on on Normal and Primary Cancer and Primary and Metastasis Cancer in the last two columns. We can find the ordered list are totally different. But, when we compare the result of gene ontology terms based on QUICKGO[12], there are some meaningful gene ontologies such GO:0002376 immune system process, GO:0006958 complement activation, classical pathway, GO:0035743 CD4-positive, alpha-beta T cell cytokine production, GO:0045087 innate immune response and GO:0045926 negative regulation of growth. Likewise the result of AML and ALL, there are some different orders on Normal and Primary and Primary and Metastasis. But, whole meaningful genes are in the ordered list in terms of gene ontology. We also apply the extended version of SVM-RFE and consider interaction with other genes for further accurate results.

5. Conclusion

We have suggested a novel gene selection algorithm combining Support Vector Machines and Intensity-dependent normalization, which uses log differential expression ratio (Minus vs Add plot, MA-plot) [2] in RNA-Seq. we exploit DEGseq, a free R package for the differentially expressed genes based on MA-plot methods. The input of DEGseq is uniquely mapped reads from RNA-seq. The output includes the expression values, a P-value, Z score, fold change log normalization and two kinds of Q-values. The proposed algorithm can exploit DEGseq right before SVM-RFE [4] to reduce the size of the number of the potentially distinguishable genes. The proposed algorithm can be convenient and quick than previous, because it uses all functions in R package. It have more improvement with regard to the time consuming in terms of Big-Data. We compared the results of the proposed algorithm with one of the existing algorithms [13, 14] with downloaded samples from Big-Data, such as NCBI-GEO [7, 8]. The comparison was quite comparative based on the gene ontology of the selected order list than the previous. We can find that it can be accomplishing minimum relevancy in terms of Big-Data.

참 고 문 헌(Reference)

[1] H. Bolouri, “Modeling genomic regulatory networks with big data”, Trends in Genetics, Vol. 30, No. 5, pp.182, 2014. https://doi: 10.1016/j.tig.2014.02.005.

[2] Y. H. Yang, S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed, “Normalization for cDNA microarray data”, Nucleic Acids Res, Vol.30, No.4, pp.e15, 2002. https://doi.org/10.1093/bioinformatics/btg146

[3] M. Zhu, J. Dahmen, G. Stacey, and J. Cheng, “Predicting gene regulatory networks of soybean nodulation from RNA-Seq transcriptome data”, BMC Bioinformatics, Vol.14, p.278, 2013. https://doi: 10.1186/1471-2105-14-278

[4] 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

[5] X. Li, S. Peng, J. Chen, B. Li, H. Zhang, and M. Lai, “SVM-T-RFE: a novel gene selection algorithm for identifying metastasisrelated genes in colorectal

cancer using gene expression profiles", Biochem. Biophys. Res. Commun. Vol.419, pp.148-153, 2012. https://doi.org/10.1016/j.bbrc.2012.01.087.

[6] S. Mishra., D. Mishra, "SVM-BT-RFE: An improved gene selection framework using Bayesian T-test embedded in support vector machine (recursive feature elimination) algorithm" Karbala International Journal of Modern Science, Vol.1, pp.86-96, 2015. https://doi.org/10.1016/j.kijoms.2015.10.002

[7] 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, 1999. https://doi.org/10.1126/science.286.5439.531

[8] S-K Kim, S-Y Kim, J-H Kim, S-A Roh, D-H Cho, Y-S Kim, J-C Kim, "A nineteen gene-based risk score classifier predicts prognosis of colorectal cancer patients", Molecular Oncology, Vol.8, Iss. 8, pp.1653-1666, 2014. https://doi.org/10.1016/j.molonc.2014.06.016

[9] C. Kim, "Combining Support Vector Machine Recursive Feature Elimination and MA-plot-based methods for Gene Selection in cDNA(RNA-seq) data" ICONI 2016. http://www.iconi.org

[10] M. Ezzeldin, A. Bashir1, H. S. Shon, D. G. Lee, H. Kim and K. H. Ryu, "Real-Time Automated Cardiac Health Monitoring by Combination of Active Learning and Adaptive Feature Selection" KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 7, NO. 1, Jan 2013. https://doi.org/10.3837/tiis.2013.01.007

[11] B. Ahn, E. Abbas, J.-A. Park and H.-J. Choi, "Increasing Splicing Site Prediction by Training Gene Set Based on Species," KSII Transactions on Internet and Information Systems, vol. 6, no. 11, pp. 2784-2799, 2012. https://doi.org/10.3837/tiis.2012.10.002

[12] QuickGO http://www.ebi.ac.uk/QuickGO-Beta/

[13] L. Wanga, Y. Wangb, Q. Chang, "Feature selection methods for big data bioinformatics", Methods, Vol.111, No.1, pp21-31, 2016. https://doi.org/10.1016/j.ymeth.2016.08.014.

[14] S. A. Zadeh, M. Ghadiri, V. S. Mirrokni, M. Zadimoghaddam, "Scalable Feature Selection via Distributed Diversity Maximization" AAAI 2017, pp.2876-2883. http://www.aaai.org/Conferences/AAAI/aaai17.php

◐ 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~현재 경기대학교, 컴퓨터과학과, 대우교수
관심분야 : 빅데이터, 머신러닝, 딥러닝 강화학습, IoT, 클라우드 컴퓨팅,
E-mail : kimcha0@kgu.ac.kr