EPiC Series in Computing

Volume 60, 2019, Pages 129–138

Proceedings of 11th International Conference on Bioinformatics and Computational Biology

A SVM Model for Candidate Y-chromosome Gene Discovery in Prostate Cancer

Wageesha Rasanjana, Sandun Rajapaksa, Indika Perera, Dulani Meedeniya
Department of Computer Science and Engineering
University of Moratuwa, Sri Lanka
rasanjana.wageesha@gmail.com, {sandunip, indika, dulanim}@cse.mrt.ac.lk

Abstract

Prostate cancer is widely known to be one of the most common cancers among men around the world. Due to its high heterogeneity, many of the studies carried out to identify the molecular level causes for cancer have only been partially successful. Among the techniques used in cancer studies, gene expression profiling is seen to be one of the most popular techniques due to its high usage. Gene expression profiles reveal information about the functionality of genes in different body tissues at different conditions. In order to identify cancer-decisive genes, differential gene expression analysis is carried out using statistical and machine learning methodologies. It helps to extract information about genes that have significant expression differences between healthy tissues and cancerous tissues. In this paper, we discuss a comprehensive supervised classification approach using Support Vector Machine (SVM) models to investigate differentially expressed Y-chromosome genes in prostate cancer. 8 SVM models, which are tuned to have 98.3% average accuracy have been used for the analysis. We were able to capture genes like CD99 (MIC2), ASMTL, DDX3Y and TXLNGY to come out as the best candidates. Some of our results support existing findings while introducing novel findings to be possible prostate cancer candidates.

keywords: Support Vector Machines; prostate cancer; Y-chromosome; differential expression; microarray data; log fold change.

1. Introduction

Advancement in computing technology has enabled the possibility of materializing microscopic information into humanly sensitive data, thus causing a massive growth of bioinformatics and computational biology fields. Collecting biological information of humans and other species, which once was a virtually impossible task has now become trivial. The depth of available cellular information of species has gone from cell level to DNA sequences. Furthermore, gene expression data can be acquired for analysis on computers with the development of Microarray technology [18] [21]. The advancement of retrieving methods and purity of Microarray data has given insights into touching

O. Eulenstein, H. Al-Mubaid and Q. Ding (eds.), BiCOB 2019 (EPiC Series in Computing, vol. 60), pp. 129–138

untouchable aspects in Medicine and Biology [9]; thus, benefitting the treatments for various diseases. Cancer treatments have become highly cellular based as bioinformaticians and scientists have been discovering genes and their interactions with each other over the past decade. These discoveries could cause triggers for treatments of a number of cancers and other sorts of untreatable diseases. In fact, studies have been carried out on the genome level to find differentially expressed genes that are remarkable in cancerous tissues [3] [14] [23]. However, the heterogeneity of these data has caused many hardships in analyzing gene expression profiles, which in turn provoking the development of novel data-analysis techniques.

One of the most common, yet hardly treatable cancers is Prostate cancer, which can be seen often among men around the globe [2]. 84,861 men from the USA [20] and 47,151 men from the UK [5] were diagnosed with prostate cancer in 2015 while 11,631 deaths were reported in the UK in 2016. It has two basic stages as primary prostate tumour and metastatic prostate cancer where primary cancer is of low risk than metastatic cancer. The primary prostate tumour is only located within the prostate gland while the metastatic disease is spread across many other organs of the body [24].

The probability of a man being diagnosed with it rises highly with the age making this cancer common among the geriatric population. The 84,861 men who were diagnosed with prostate cancer in the USA in 2015 includes 12,489 men aged between 60 and 79 and 14,529 men above the age of 80. This has led to research on gene expression data related to prostate cancer, which in turn revealed to be massively heterogeneous [1] [4] [23]. Many of the studies have been carried out in both biological sciences and statistical domains, highlighting information about some of the candidate genes that can be vital in prostate cancer [1] [22]. Since prostate cancer is male-specific, the importance of analyzing the effect of Y-chromosome genes towards the proliferation of the disease has been identified. However, most of the research carried out on prostate cancer is focused on chromosomes other than Y-chromosome.

This paper presents an approach for the identification of candidate Y-chromosome genes that can have an impact on the growth of prostate cancer. Section 2 presents an overview of commonly used analysis techniques in both statistical and supervised learning domains while describing important statistical terms used in our approach. Section 3 outlines the analysis criteria and methodology of the model. Obtained experimental results and comparisons are presented in Section 4. Finally, section 5 concludes the paper with the inference obtained from the results, future extensions, and emphasizes the importance of the research.

2. Background

Common practice in big data analysis is to build statistical models that can interpret the probabilistic behaviour of data. In this paper, interest lies within gene expression profiles and their differential expressions. Biological data are highly variable and very sensitive due to their microscopic scale, especially when it comes to gene expressions. The variance of data highly depends on each of the gene expression values, which vary largely with respect to the tissue from which they have been collected. In order to remove the variability and noise in data, data pre-processing should be taken place as a common practice. Table 1 and Figure 1 depict information about average and variance of expression is 5 randomly selected genes from GSE6919 dataset.

Table 1. Average expression across 171 patient-samples

AR	IGHV3_23	EIF2AK2	RPS19	PLAGL1
1760.5	3.2	1342.7	10398.5	81.8

In statistics, the log2 transform is widely used to get rid of the high variance of data due to its simplicity [8] [13] [19]. It is possible to get rid of the dependency between the mean and variance of data by using this transform. This dependency is generally known as heteroskedasticity. This is important when dealing with noise (error) in microarray data because errors largely dependent on the population mean, especially when finding the quantitative change of a statistical variable.

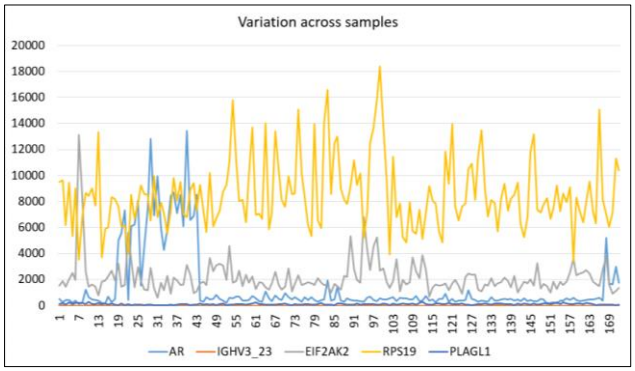

Figure 1. Variation of expression across 171 samples

Fold change (FC) is a statistical quantity used to measure the quantitative change of a variable from one state to another. FC can be calculated between tumour (or cancer) samples and normal samples in microarray gene expression data. It is an indication of the amount of expression change that has occurred when the gene transformed from a normal state to a cancerous state. When the data are log-transformed before calculating the fold change, the resulting value is interpreted as log fold change (LogFC). In addition to that, there are many other statistical quantities, which can be used to evaluate the differential expression of genes such as t-statistics or Bayesian log odds (B-statistics) [19].

Apart from these statistical estimation methods, machine-learning methods have been widely used to analyze gene expression data and sequence data [11]. Since gene expression data are highly variable, statistical methods may provide incorrect estimates when the expression differences are insignificant between tissue samples. Pirooznia et al. [16] have carried out a study to perform comparisons between several machine learning algorithms regarding the applicability in microarray expression analysis. This study concludes that SVM results in better performance compared to other supervised learning algorithms for microarray expression analysis.

A study done by Khosravi et al. [10] found a set of Y-chromosome genes in cancerous tissues which exhibit highly differentiated expression levels compared to other normal tissues. This phenomenon is used in many of the other studies to classify cancer candidate genes by analyzing their differential gene expression profiles. For some genes, expression patterns either can be up or down regulated and those genes are classified as cancer candidates. The occurrence of up or down regulation during the metastatic transformation process is highlighted in a comprehensive study done by Chandran et al. [4]. This information is used extensively to extract candidate Y-chromosome genes having differential expressions between normal tissue cells and cancerous cells in our research. Moreover, SVM is used in our approach to predict the genes since it has been concluded as the best learning algorithm for microarray data analysis [16]. However, further laboratory testing and comprehensive analysis are required to enhance these results and confirm them as truly vital candidates in prostate cancer.

3. Analysis Criteria and Methodology

The proposed approach employs a dataset extracted from GEO Dataset under the accession number GSE6919 with platform ID GPL8300 [15]. The dataset has 171 patient-samples those, which are acquired from four distinct conditions; normal prostate tissue samples without any pathological alterations, samples adjacent to the primary prostate tumour, primary prostate tumour samples, and metastatic prostate cancer samples. Each of these distinct sample types contains 18, 63, 65 and 25 samples respectively. Healthy prostate tissue samples without any pathology and samples adjacent to a prostate tumour are altogether labelled to normal category while primary tumour samples and metastatic cancer samples are put into the cancerous category. Throughout this paper, we refer to normal samples and cancerous samples according to the above categorization. These patient-samples, each having 12625 gene expression values, are used for categorization purpose while training, testing and prediction performed on gene-probe samples where one gene-probe sample has 171 values. Overall, the microarray dataset is a matrix having 12625 rows and 171 columns. The proposed approach consists of a number of steps combining both statistical and supervised learning methods. The dataset is restructured into 8 categories each of which contains more than one sample category as illustrated in Table 2. The purpose of the categorization is the unique identification of differentially expressed genes throughout the cancer expansion process from the normal stage to the metastatic stage. Our study compares every normal sample with every cancerous sample creating the need for 8 categories. Therefore, 8 SVM models were built having one model for each category. SVM models are trained using a subset of ranked gene probes from the whole gene set of 12625 genes while the Y-chromosome gene set with 45 genes, is extracted for the classification. *Limma* package in R is used for the probe ranking process since it has been widely used among the computational biology researchers for the statistical expression analysis of genes.

Table 2. Categorization of samples

Category	No. of Samples
NOR-MET (Normal & Metastatic samples)	43
ADJ-MET (Adjacent to tumour & Metastatic samples)	88
NOR-ADJ-MET (Normal, Adjacent to a tumour & Metastatic samples)	106
NOR-TUM (Normal & Primary tumour samples)	83
ADJ-TUM (Adjacent to tumour & Primary tumour samples)	128
NOR-ADJ-TUM (Normal, Adjacent to tumour & Primary tumour samples)	146
TUM-MET (Primary tumour & Metastatic samples)	90
ALL (All samples)	171

The gene probes are ranked according to log-fold-change (LogFC) value. Large LogFC values of the top genes provide evidence for the significance of differential expression pattern in them. Under-expressed genes exhibit a negative LogFC value while over-expressed genes exhibit the opposite. In order to signify the patterns in the training dataset, it is divided into two subsets following their positive and negative LogFC values. Highly over-expressed genes tend to have significantly increased expression in the metastatic region while highly under-expressed metastatic gene expressions are significantly shrunk. Thus, highlighting patterns in both cases. This distinct pattern will be slowly diminished with the reduction of LogFC value. Seemingly, top-ranked genes show a greater pattern in differential expression while others barely display a pattern. As per the investigation carried out, none of the top genes was from Y-chromosome. Concisely, Y-chromosome genes that are differentially expressed in the prostate cancer cannot be accurately identified just by the statistical ranking method, but another more sophisticated approach is required. Therefore, a combined approach containing a statistical ranking method and a supervised learning model was used in this research.

3.1 Model Building and Classification

For each SVM model, ~200 genes were classified as vital for cancer and ~200 genes were considered to have no effect for cancer. These classifications were done based on LogFC values. Top-ranked genes from each category having the highest LogFC values (LogFC > 1.5) [17] are classified as vital (boolean 1) and the genes from the bottom of the ranking table having lowest LogFC are considered as neutral (boolean 0); thus, creating a training dataset of ~400 genes out of 12625 total genes. Alternatively, all *NA* values were replaced by zero. We have used a 70:30 dataset split ratio based on the size of the available dataset [7]. Therefore, our training set has ~280 samples, leaving ~120 samples for the test data. In statistics, training datasets less than ~100 samples in size tend to have a greater variance resulting in low accuracy models. Thus, around 85% of data is recommended to use as the training set. Since our training set contains more samples, we refrained from using a higher split ratio.

Using the dataset both linear and non-linear SVM models were evaluated with 10-fold cross-validation. This evaluation resulted in better accuracy for the linear model as illustrated by the bar plot in Figure 2. Therefore, a linear SVM model was adapted for our approach. Moreover, a cost-grid ranging from 0.05 to 35 was used to find the best regularization parameter for each model as illustrated in Figure 3. The *train* function from the *Limma* package selects the optimum cost value from the range and sets to the training model. The accuracy tent to be constant after cost is 25; hence, we limited the range up to 35. After classification, the prediction results from each of the 8 models were observed to derive conclusions. Figure 4 shows the architectural setup of the procedure followed in our study.

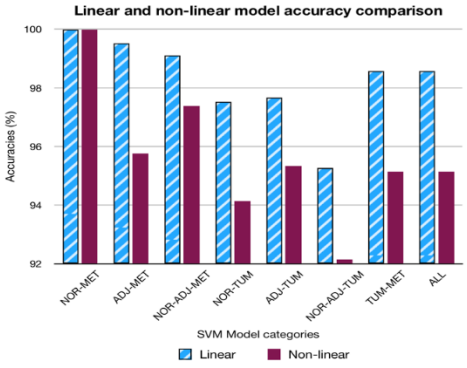

Figure 2. Comparison of linear & non-linear model accuracies

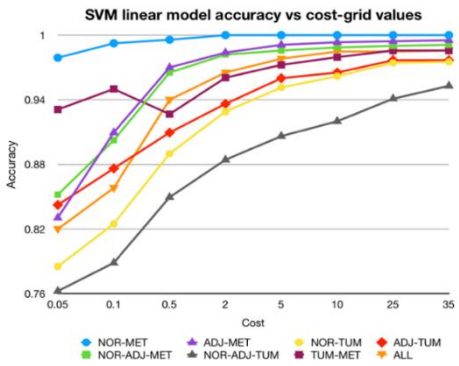

Figure 3. Variation of the accuracy of SVM models against a range of cost values

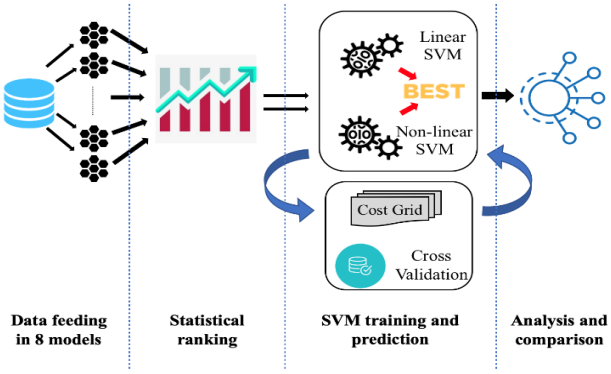

Figure 4. The architecture of the experimental setup

4. Results and comparison

We analyzed 171 samples taken from 171 patients and from different conditions in the human body. Those contain both cancerous and healthy samples from tissues in both prostate gland and other organs (only metastatic samples are taken from other organs). All the samples contain 12625 total gene probes, which we use as samples for training, testing and prediction, summing up to 2,158,875 out of which 67 *NA* values are replaced by zero. We built 8 SVM models to identify and interpret the Y-chromosome genes that are differentially expressed across different tissues. In addition, we simultaneously investigated over-expressed genes and under-expressed genes under those 8 models. Finally, we compared our results with the existing findings from the literature. Table 3 and Table 4 display results and comparison of expression change between candidate Y-chromosome genes across different models.

Table 3. Under-expressed genes

Types	Genes	SVM Model categories							
		NOR-MET	ADJ-MET	NOR-ADJ-MET	NOR-TUM	ADJ-TUM	NOR-ADJ-TUM	TUM-MET	ALL
Under-Expressed	VAMP7	v	v	v				v	v
	USP9Y	v	v	v					
	ASMTL	v	v	v		v	v	v	v
	KDM5D	v	v	v				v	
	DDX3Y	v	v		v	v	v		v
	CD99	v	v	v	v	v	v	v	v
	IL3RA				v		v		v
	RBMY1J							v	
	UTY							v	
	EIF1AY							v	v

Genes with changes of expression pattern in the majority of classes

Initially, we considered the significantly differentiated samples, which are highly cancerous and highly normal. NOR-MET model contained genes from normal prostate tissue samples (samples that are neither diagnosed nor pathologically altered) and metastatic cancer samples (samples taken from different metastatic cancer locations such as lungs, liver or lymph nodes). In the ADJ-MET model, we analyzed genes from tissue samples that are adjacent to the prostate gland and genes from metastatic locations. In both models, we have achieved over 99% model accuracy and the outputs were almost identical in both cases. We identified seven over-expressed Y-chromosome genes and six under-expressed genes namely; *VAMP7*, *USP9Y*, *ASMTL*, *KDM5D*, *DDX3Y*, *SLC25A6* and *CD99*, *LOC101928634*, *AKAP17A*, *TXLNGY*, *SLC25A6*, *RPS4Y1* as illustrated in columns 3, 4 and 5 of Table 3. *USP9Y* and *DAZ4* genes do not show over-expression in the NOR-MET model. NOR-ADJ-MET model was created to justify our findings of NOR-MET and ADJ-MET models. The results under-expression from this category were similar to the first two categories except for the loss of *DDX3Y* and over-expression was different for some genes. When the normal and adjacent samples are combined into one category, the expression patterns become less significant compared to when they are evaluated separately. Thus, causing difficulty for accurate classification.

Then we analyzed genes that are differentially expressed during the tumour growth process. In contrast, healthy normal prostate tissue samples and primary prostate tumour samples were considered. First, we compared normal prostate gland samples and primary prostate tumour samples, which is the NOR-TUM model. In the ADJ-TUM model, we considered samples that are adjacent to primary prostate tumour and samples taken directly from a primary prostate tumour. In addition, we analyzed the NOR-ADJ-TUM model as well to justify the results. We could analyze them as a special scenario where these genes exhibit differential expressions only within a primary tumour and no differential expression in the metastatic stage. In which, they have resulted as candidates from NOR-TUM, ADJ-TUM and NOR-ADJ-TUM categories while displaying no expression changes in NOR-MET, ADJ-MET and NOR-ADJ-MET models. We could identify only one common under-expressed gene (*CD99*) across these 3 models though there are many common over-expressed genes such as *BPY2*, *XKRY2* and *SRY* etc. These genes that only exhibit early changes expression pattern are shown in gray colour in Table 4.

Table 4. Over-expressed genes

Types	Genes	SVM Model categories							
		NOR-MET	ADJ-MET	NOR-ADJ-MET	NOR-TUM	ADJ-TUM	NOR-ADJ-TUM	TUM-MET	ALL
Over-Expressed	LOC101928634	^	^	^					^
	AKAP17A	^	^	^					^
	TXLNGY	^	^		^	^	^	^	^
	SLC25A6	^	^				^		^
	RPS4Y1	^	^				^		^
	USP9Y		^						
	DAZ4		^	^					
	IL9R				^			^	
	BPY2				^	^	^		^
	XKRY2				^	^	^		^
	SRY				^	^	^		
	ASMT				^	^	^	^	
	TSPY10				^	^	^	^	^
	TTTY15				^	^	^		
	PCDH11Y				^	^	^		^
	LOC100509646					^	^		
	ZFY					^			
	NLGN4Y					^			
	CDY1B					^	^		
	CSF2RA					^			
	SHOX					^	^		
	VCY1B							^	^

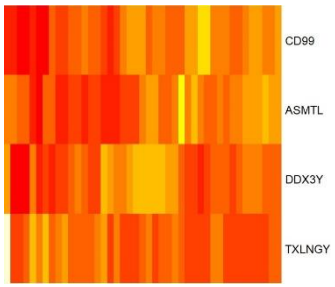

Figure 5.1. NOR-MET

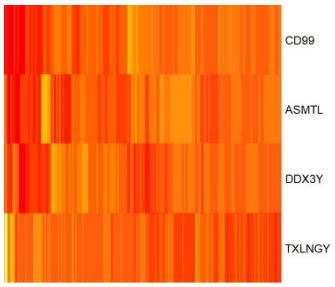

Figure 5.2. ADJ-MET

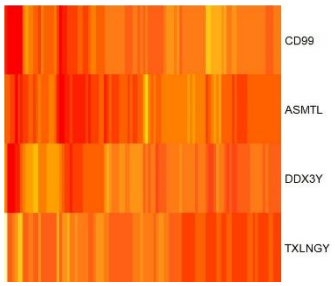

Figure 5.3. NOR-ADJ-MET

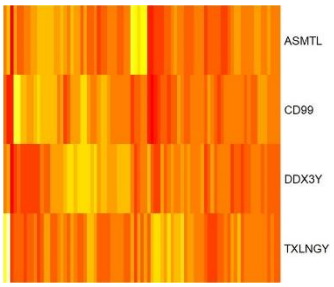

Figure 5.4. NOR-TUM

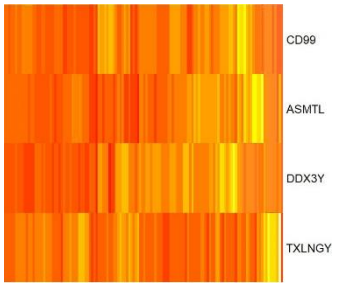

Figure 5.5. ADJ-TUM

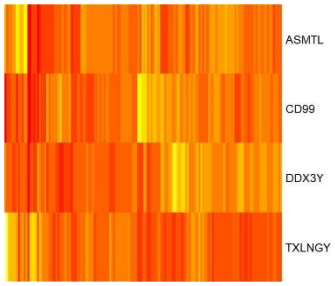

Figure 5.6. NOR-ADJ-TUM

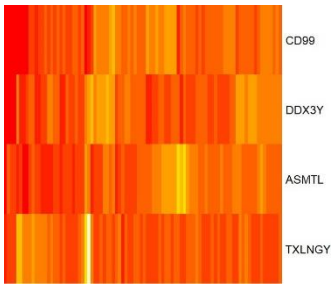

Figure 5.7. TUM-MET

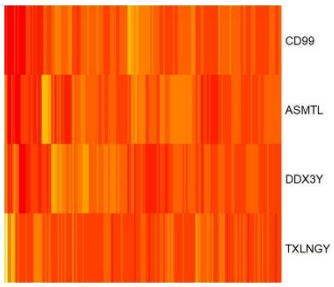

Figure 5.8. ALL

Finally, we analyzed the genes from the TUM-MET (primary tumour samples and metastatic samples) model and ALL model. In the TUM-MET model, we attempt to focus on the cancer spread phenomena starting from a primary tumour to other organs. In ALL model we focused on identifying the genes that generally have a differential expression due to prostate cancer. Out of the resulting candidates, some of the Y-chromosome genes can be recognized as most vital since they significantly vary in expression level across all the categories as illustrated in Table 3 and Table 4. Figure 5.1 to Figure 5.8 depict heatmaps of row-normalized expression for critically identified genes across 8 models in which intensity decrease from red to yellow. High intensity conforms to high normalized-value in each of 4 rows. Next chapter presents important conclusions about our findings and insights into future research.

5. Conclusion and Future Work

The analysis carried out by the categorical SVM model with a minimum accuracy of 95%, results in a set of decisive Y-chromosome genes namely CD99 (also known as MIC2), ASMTL, DDX3Y, and TXLNGY. Those genes are highlighted in yellow colour in Table 3 and Table 4. It is highly probable that the aforementioned Y-chromosome genes to be actively involved in prostate cancer generation and metastasis process when considering the high accuracy obtained for the SVM models. There are many biological studies carried out focusing on the CD99 gene and its involvement in prostate and other types of cancers [3][25]. Apart from that, Lau et al. [12] have found information about the involvement of many Y-chromosome candidates including ASMTL, ILR3, and RPS4Y1 in prostate cancer. In addition, the genes highlighted in gray colour rows might play vital a role in prostate tumour generation. Early medical precautions targeted on them may be able to prevent the cause of developing the tumour. These Y-chromosome genes do not exhibit significant expression patterns when compared to the top-ranked genes in differential expression analysis but the changes in their expressions from normal tissue to cancerous tissue are significant for closer observations. Table 3 and Table 4 contains information about many other genes from our findings, which are correlated with Lau's work. Moreover, Dasari et al. [6] have done similar work to add stability to our findings.

However, it should be highlighted that future work is needed to provide confirmation about these Y-chromosome genes as to how they relate to the progression of prostate cancer. Microarray data may contain noise, which cannot be removed completely by data pre-processing. Therefore, the results obtained through the computational methods can never be accurate enough for direct acceptance. We suggest carrying out thorough narrowed down laboratory experiments on these genes to investigate the actual role they involved in the disease. In fact, it will be highly beneficial for the biological community to find out new cellular level treatments for prostate cancer.

6. References

[1] Burmester, James K., et al. 2004. Analysis of candidate genes for prostate cancer. *Human heredity* 57, no. 4 (2004): 172-178.

[2] Cancer Research UK, Prostate cancer statistics, https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer, October 2018.

[3] Carter, H. Ballentine. 2004. Prostate cancers in men with low PSA levels—must we find them?. *The New England journal of medicine* 350, no. 22 (2004): 2292.

[4] Chandran, Uma R., et. al 2007. Gene expression profiles of prostate cancer reveal involvement of multiple molecular pathways in the metastatic process. *BMC cancer* 7, no. 1 (2007): 64.

[5] Cheng, Liang, Nagabhushan, Moolky, Pretlow, Theresa P., Amini, Saeid B. and Pretlow, Thomas G. 1996. Expression of E-cadherin in primary and metastatic prostate cancer. *The American journal of pathology* 148, no. 5 (1996): 1375.

[6] Dasari, Vijay K. et al. 2001. Expression analysis of Y chromosome genes in human prostate cancer. *J Urol.* (2001).

[7] Dobbin, Kevin K., and Richard M. Simon. 2011. Optimally splitting cases for training and testing high dimensional classifiers. *BMC medical genomics* 4, no. 1 (2011): 31.

[8] Friedman, Nir, Cai, Long and Xie, X. Sunney 2006. Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. In: *Physical review letters*. (2006).

[9] Golub, Todd R., et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, no. 5439 (1999): 531-537.

[10] Khosravi, Pegah, Zahiri, Javad, Gazestani, Vahid H., Mirkhalaf, Samira, Akbarzadeh, Mohammad, Sadeghi, Mehdi, Goliaei, Bahram 2014. Analysis of candidate genes has proposed the role of y chromosome in human prostate cancer. Iranian journal of cancer prevention. (2014); 7(4):204.

[11] Larranaga, Pedro et al. 2005. Machine Learning in Bioinformatics. In: *Briefings in Bioinformatics.* (2005): 86-112.

[12] Lau, Yun-Fai C., and Zhang, Jianqing 2000. Expression analysis of thirty one Y chromosome genes in human prostate cancer. *Mol. Carcinog.* (2000), 308-321.

[13] Lin, Simon M., Du, Pan, Huber, Wolfgang and Kibbe, Warren A. 2008. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research 36*, 2 (2008), e11-e11.

[14] Myers, Jennifer S., von Lersner, Ariana K., Robbins, Charles J., and Sang, Qing-Xiang Amy (2015). Differentially Expressed Genes and Signature Pathways of Human Prostate Cancer. *PLOS ONE*, 10(12), p.e0145322.

[15] NCBI, GEO Accession viewer. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6919, October 2018.

[16] Pirooznia, Mehdi, Yang, Jack Y., Yang, Mary Qu and Deng, Youping 2008. A Comparative Study of Different Machine Learning Methods on Microarray Gene Expression Data. In: *BMC Genomic* (2008).

[17] Raza, Khalid, Mishra, Akhilesh 2012. A Novel Anticlustering Filtering Algorithm for the Prediction of Genes as Drug Target. In: *American Journal of Biomedical Engineering.* . (2012): 206-211.

[18] Salome, J. Jacinth 2012. Efficient Retrieval Technique for Microarray Gene Expression. *International Journal of Information Retrieval Research (IJIRR), 2*(2), (2012), 43-51.

[19] Satagopan, Jaya M. Olson, Sarah H., Elston, Robert C. 2017. Statistical interactions and Bayes estimation of log odds in case-control studies. Statistical methods in medical research. (April 2017); 26(2):1021-38.

[20] Siegel, Rebecca L., Miller, Kimberly D. and Jemal, Ahmedin Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians 67*, 1 (2017), 7-30.

[21] Slonim, Donna K. and Yanai, Itai 2009. Getting Started in Gene Expression Microarray Analysis. *PLoS Computational Biology 5*, 10 (2009), e1000543.

[22] Thibodeau, S. N., et al. 2015. Identification of candidate genes for prostate cancer-risk SNPs utilizing a normal prostate tissue eQTL data set. *Nature communications* 6 (2015): 8653.

[23] Wang, Lipo, Chu, Feng and Xie, Wei 2007. Accurate cancer classification using expressions of very few genes. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 4, no. 1 (2007): 40-53.

[24] Yuan, Ye, Kishan, Amar U. and Nickols, Nicholas G. Treatment of the primary tumor in metastatic prostate cancer. *World Journal of Urology* (2018).

[25] Zaccarini, Daniel, J. et al. 2018. Expression of TLE-1 and CD99 in Carcinoma: Pitfalls in Diagnosis of Synovial Sarcoma. In: *Applied Immunohistochemistry & Molecular Morphology.* (2018): 368-373.