Enliven Archive
Intensifying Thoughts

Review Article www.enlivenarchive.org Enliven: Journal of Genetic, Molecular and Cellular Biology

ISSN: 2379-5700

Improving Cancer Genomics Using Support Vector Machine (SVM) Learning

Vincent Marcov

Department of Pediatrics and Genetics, Ain Shams University, Nour Mosque, 38 Abbasia Next, El-Mohamady, Al Waili, Egypt

*Corresponding author: Vincent Marcov, Department of Pediatrics and Genetics, Ain Shams University, Nour Mosque، 38 Abbasia Next، El-Mohamady, Al Waili, Egypt, E-mail: vincentmarcov@protonmail.com

Received Date: 15th March 2018
Accepted Date: 25th April 2018
Published Date: 10th May 2018

Citation: Marcov V (2018) Improving Cancer Genomics Using Support Vector Machine (SVM) Learning. Enliven: J Genet Mol Cell Biol 5(2): 004.

Researchers have used a potent classification tool known as Support Vector Machine (SVM) to perform cancer genomic sub-typing [1]. The current technological development has contributed towards the generation of huge volumes of epigenetic and genomic data [2]. The move has created greater awareness on novel medicine or drug targets, driver genes of cancer, as well as new biomarkers [3]. In this case, the researcher assessed the cancer genomic literature to determine the latest developments of Support Vector Machine (SVM). They intended to add knowledge on the strength of SVM learning as well as how it can be integrated into cancer genomic [4]. Machine learning (ML) uses artificial intelligence to forecast future data, for instance, several methods used to perform probability, statistics, and optimization as artificial neural networks, Naïve Bayes, decision trees, K-nearest neighbor, and logistic regression are usually used as ML techniques [5]. Support Vector Machine learning is popularly relied on to explore intricate datasets with the view of identifying indefinable patterns. Furthermore, it can be counted on to detect a facial model, a person's voice, as well as recognize handwriting [6]. As a genomic disease, cancer is characterized by genetic feature patterns that might assist to categorize the cancer subtypes as well as predict drug benefit, find a tumor-specific biological procedure, as well as the conduct prognosis of the outcome [7,8]. Consequently, it is evident that the artificial intelligence of such Support Vector Machines can greatly assist in determining the patterns of a feature in diverse applications [9].

The support vector machine is usually used in developing a classifier (which is utilized to conduct cancer classification) and it is supported by the microarray gene expression data [10]. In this study, the researchers relied on thirty-eight participants who were recruited as attaining set [11]. Besides, they employed a simplified machine learning algorithm to establish the difference between two known types of blood cancer. The learned SVM model was relied on in testing an autonomous dataset of thirty-four patients [12]. The results obtained revealed that the support vector machine performed excellently in classifying gene features [13]. With regards to drug development to facilitate pharmacotherapies for cancer disease in positively diagnosed patients, the researchers argue that there are limited drugs which were being prescribed to treat the different varieties of terminal cancer [14]. The main challenge that inhibits the development of cancer drugs include high toxic and resistance levels as well as the side effects [15]. The conventional process adopted to facilitate drug development usually take a long time to achieve results, and they are expensive; besides, there are many experimental procedures used in determining the active compounds that can neutralize a biotic target [16].

Presently, the highest margin hyper-planes of the SVM can be utilized to support the screening procedure. The inactive and active compounds are usually separated by the hyper-plane [17]. The researcher indicates that support vector model relied on the cancer cells' genetic characteristics to develop medications used to treat cancer [18]. Additionally, they illustrate the Model was employed to explore massive groups of compounds from open source databanks to establish likely salutary compounds to be used in anti-cancer drugs cancer drugs [19]. Moreover, the researchers describe how support vector achiness can be used to virtually screen toxics emitted by radio-protectors [20]. Radiation is among the numerous therapies which are broadly adopted to kill cancerous cells [21]. In scenarios where by the target protein is recognized in addition when the researchers sought to look for a compound that bonded perfectly with the target protein, the SVM performed excellently as compared to other machine learning approaches [22].

SVMs play a great role in detecting cancerous targets in addition they are widely adopted to forecast the drugability scores. Additionally, 3 to 6 universal descriptors of sites used for a binding protein which account for size, chemical composition, and firmness are employed to design the Support Vector Machine Model [23]. Biometric is greatly enhanced through the process of ascertaining the drug target proteins. The researchers also described a novel structure used in retrieving drug target protein from a protein dataset [24]. Previous techniques disregard useful biochemical structures in favor of protein-protein collaborative networks [25]. Furthermore, the research included extrapolation of cancer drug sensitivity where by computational models were relied on to estimate the rate by which the cancer cells can respond to compounds contained in the medication to reinforce the discovery of cancer pharmacotherapies [26]. Therefore, it is evident that applying support vector machine learning can assist in examining the huge volume of intricate and varied cancer genomic data that can help in the development of drugs.

Table 1: Summary Application of SVM in Cancer Genomic [1]

Applications	SVM Model*	Data	Cancer type	Ref
Classification/Subtyping	Linear SVM	mRNA	Soft tissue sarcomas	14
	Linear SVM	Methylation	Leukemia	17
	SVM-RFE	Methylation	Multiple	20
	Meta-SVM	Multi-Omics	Breast Cancer	28
	Linear SVM	Protein	Multiple	22
	Linear SVM	Proteomics	Breast Cancer	23
	Linear SVM	CNV	Bladder Cancer	24
	Linear SVM	SNP	Breast Cancer	25
Biomarker/Signature	SVM-FRE	mRNA	Multiple	38
	net SVM	Expression& Interaction	Breast Cancer	41
Drug Discovery				
Screen Radiation Protection	RBF-SVM	Normal Cell Culture	All	46
Identify Novel Drug Targets	Linear SVM	Drug gability data set	All	48
Assess target-ligand Interactions	Reg-SVM	Structure-activity data sets	All	49
Identify Drug target Proteins	Biased SVM	A Collected Protein Data Set	All	51
Anti/ non-anticancer Molecule Classification	Linear SVM	Anti, Non-anticancer molecules	NCI-60 Cells	53
Anticancer Drug Sensitivity Prediction	Ensemble SVM	Cell Multi Omics	Cell Lines	55
Predicting Substrates of the cancer Resistance	Linear SVM	BCRP Substrates	Breast Cancer	56
Driver Gene Discovery				
Kinase Mutation Activation	Linear SVM	Kinase data set	All	59
Drivers Versus Passengers	Linear SVM	COSMIC	All	62
Gene Interaction				
	RBF-SVM	Interacting Proteins(DIP)	All	68

*SVM-RFE: SVM recursive feature elimination; RBF- SVM: Radial basis function SVM; net SVM: Network Constrained SVM; Reg-SVM: regression SVM; CNV: Copy Number Variaton; SNP: Single nucleotide Polymorphism; COSMIC: The catalogue of somatic mutations in cancer; BCRP: Human breast cancer resistance Protein.

References

1. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, et al.(2018) Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics-Proteomics 15: 41-51.
2. Yichuan Tang (2013) Deep Learning Using Linear Support Vector Machines. Bioinformatics 1306: 1-34.
3. Cruz JA, Wishart DS (2006) Applications of Machine Learning in Cancer Prediction and Prognosis. Cancer Inform 2: 59-77.
4. Cicchetti DV (1992) Neural Networks and Diagnosis in the Clinical Laboratory: State of the art. Clin Chem 38:9-10.
5. Simes RJ (1985) Treatment Selection for Cancer Patients: Application of Statistical Decision Theory to the Treatment of Advanced Ovarian Cancer. J Chronic Dis 38: 171-186.
6. Aruna S, S P Rajagopalan (2011) A Novel SVM Based CSSFFS Feature Selection Algorithm for Detecting Breast Cancer. International Journal of Computer Applications 31: 5-6.
7. Swathi M (2017) Clustering Enhancement Using Similarity Indexing to Reduce Entropy. Enliven: Bioinform 4: 001.
8. Noble William Stafford (2004) Support Vector Machine Applications in Computational Biology. Kernel Methods in Computational Biology 3: 92.
9. Vural S, Wang X, Guda C (2016) Classification of Breast Cancer Patients Using Somatic Mutation Profiles and Machine Learning Approaches. BMC Syst Biol. 10: 62.
10. Li T, Zhang C, Ogihara M (2004) A Comparative Study of Feature Selection and Multiclass Classification Methods for Tissue Classification Based on Gene Expression. Bioinformatics 20: 2429-2437.
11. Model F, Adorján P, Olek A, Piepenbrock C (2001) Feature Selection for DNA Methylation-Based Cancer Classification. Bioinformatics 17: S157-S164.
12. Noble, William Stafford (2006) What is a Support Vector Machine?. Nature Biotechnology 24: 1565.

13. Vapnik V (1963) Pattern Recognition Using Generalized Portrait Method. Automation and Remote Control 12: 774-780.
14. Kim S (2016) Weighted K-Means Support Vector Machine for Cancer Prediction. Springerplus 5: 1162.
15. Aizerman, Mark A (1964) Theoretical Foundations of the Potential Function Method in Pattern Recognition Learning. Automation and Remote Control 2: 821-837.
16. Alhasan Alkuhlani, Mohammad Nassef, Ibrahim Farag (2017) Multistage Feature Selection Approach for High-Dimensional Cancer Data. Soft Computing 21: 6895-6906.
17. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. (1999) Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 286: 531-537.
18. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, et al.(2018) Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics-Proteomics 15: 41-51.
19. Chapelle O (2007) Training a Support Vector Machine in the Primal. Neural Comput 19: 1155-1178.
20. Swathi M (2017) Drug Prediction of Cancer Genes Using SVM. Enliven: Pharmacovigilance and Drug Safety 4: 001.
21. Guo S, Yan F, Xu J, Bao Y, Zhu J et al. (2015) Identification and Validation of the Methylation Biomarkers of Non-Small Cell Lung Cancer (NSCLC).Clin Epigenetics 7: 3.
22. Kristin P. Bennett, Colin Campbell (2000) Support Vector Machines: Hype or Hallelujah? ACM SIGKDD Explorations Newsletter 2: 1-13.
23. Sun Z, Fu X, Zhang L, Yang X, Liu F, et al. (2004) A Protein Chip System for Parallel Analysis of Multi-Tumor Markers and its Application in Cancer Detection. Anticancer Res 24: 1159-1166.
24. Moler EJ, Chow ML, Mian IS (2000) Analysis of Molecular Profile Data Using Generative and Discriminative Methods. Physiol Genomics 4: 109-126.
25. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M et al. (2000) Support Vector Machine Classification and Validation of Cancer Tissue Samples Using Microarray Expression Data. Bioinformatics 16: 906-914.
26. TyanovaS, Albrechtsen R, Kronqvist P, Cox J, Mann M, et al. (2016) Proteomic Maps of Breast Cancer Subtypes. Nature Communications 7: 10259.
27. Segal NH, Pavlidis P, Noble WS, Antonescu CR, Viale A, et al. (2003) Classification of Clear-Cell Sarcoma as a Subtype of Melanoma by Genomic Profiling. J Clin Oncol. 21: 1775-1781.
28. Segal NH, Pavlidis P, Antonescu CR, Maki RG, Noble WS et al. (2003) Classification and Subtype Prediction of Adult Soft Tissue Sarcoma by Functional Genomics. Am J Pathol 163: 691-700.
29. Swathi M (2018) Enhancement of K-Mean Clustering for Genomics of Drugs. Enliven: J Genet Mol Cell Biol 5: 001.