Improving Automated Feature Selection in Cancer Genomics Using Reinforcement Learning

Hannah Z. Wang
Trinity Preparatory School, Winter Park, FL
wangh26@trinityprep.org

Wei Zhang
University of Central Florida, Orlando, FL
wzhang.cs@ucf.edu

***Abstract*—Genomic datasets are often characterized by high dimensionality and a limited number of samples, making it difficult to identify biologically meaningful features. Effective feature selection is crucial in biomedical research, where both genomic and clinical variables are abundant. In this study, we evaluate multiple feature selection techniques on large-scale gene expression datasets from breast cancer, lung adenocarcinoma, and ovarian cancer. We specifically adapt and optimize a novel reinforcement learning (RL) approach to address the challenges of high-dimensional genomic data. The performance of the RL-based method is compared against widely-used techniques, including Least Absolute Shrinkage and Selection Operator (LASSO), Recursive Feature Elimination (RFE), Random Forest, and Principal Component Analysis (PCA). Our experimental results demonstrate that the RL approach outperforms traditional methods in predicting cancer outcomes in most cases, highlighting its potential for identifying biologically meaningful features in genomic analysis.**

***Index Terms*—Feature Selection, Reinforcement Learning, LASSO, RFE, Random Forest, PCA, Gene Expression Data, Cancer Outcome Prediction**

I. Introduction

Powered by the high-throughput sequencing technologies, the new DNA- and RNA-sequencing methods are capable of measuring molecular activities in cells [1]. This allows researchers to address unanswered important biological and biomedical questions: 1) Elucidation of gene expressions generated from mRNA-sequencing data could lead to new molecular mechanisms such as gene regulations, and potentially molecular signals for phenotype prediction [2]; 2) Single nucleotide polymorphisms (SNPs) and copy number variations (CNVs) identified from DNA-sequencing data contribute to human genetic diversity and associated with common human diseases [3]. These high-throughput sequencing datasets show quantitative measures of more than hundreds of thousands of genomic features for a cohort of hundreds of patients. However, due to the unavoidable patient heterogeneity, statistical randomness, and experimental noise in the high-throughput genomic data, extracting valuable information and discovering the underlying patterns is becoming a serious challenge to the computational biology and machine learning communities. Feature selection plays a critical role in extracting meaningful insights from high-throughput genomics data. However, inaccurate information can undermine this process, potentially impeding progress toward accurate and efficient disease outcome prediction.

Traditional feature selection methods such as Least Absolute Shrinkage and Selection Operator (LASSO) [4] and Recursive Feature Elimination (RFE) [5] are widely used in genomic data analysis. While these techniques are effective in reducing dimensionality, they often struggle to capture complex interactions between features, particularly in high-dimensional genomic datasets. Recently, reinforcement learning (RL), a paradigm for sequential decision-making [6], has shown promise in various domains, including feature selection. However, its application to genomic feature selection is still relatively new and underexplored.

In this paper, we adapt and optimize a reinforcement learning based algorithm [7] for feature selection and apply it to three cancer gene expression datasets: breast cancer (BRCA) [8], lung adenocarcinoma (LUAD) [9], and ovarian cancer (OV) [10]. Reinforcement learning, a branch of machine learning, offers a promising alternative for feature selection. In RL, an agent interacts with an environment, receiving rewards based on the outcomes of its actions. The agent refines its strategy through exploration (trying new actions) and exploitation (using known actions that yield high rewards), aiming to maximize cumulative rewards over time. When applied to feature selection, RL can iteratively choose features while evaluating their contribution to the predictive task, making it well-suited for high-dimensional, complex datasets. Unlike traditional methods, which either rank individual features or perform an exhaustive search of feature subsets, RL approaches the problem dynamically, adjusting feature selection based on feedback from the learning process. This framework enables the exploration of feature dependencies and non-linear interactions that are critical for accurate predictive modeling in cancer genomics. Furthermore, RL can adapt to different datasets and tasks, providing a versatile and powerful tool for feature selection. We compare the performance of RL with common feature selection methods including LASSO, RFE, Random Forest, and PCA. Our results demonstrate that the RL algorithm yields more accurate models, offering a promising approach for feature selection in genomics.

II. Related Work

Molecular signatures are markers of a particular cell or a tissue phenotype. Exploring the complex relations of molecular signatures and disease phenotype is the most effective and efficient way to understand the causes of diseases for

patients [11]. Identifying representative molecular signatures (biomarkers) from the tremendous number of genomic features has became a central problem in data-driven clinical decision-making and personalized medicine. The goal is to apply feature selection techniques to identify the most discriminative biological features for predicting a given disease phenotype. Feature selection methods can generally be categorized into the following main approaches [12]:

Filter Methods: Filter methods evaluate each feature independently using statistical criteria, such as correlation with the target variable or mutual information. Common filter methods in cancer genomics include chi-square tests [13], t-tests [14], and correlation-based feature selection [15]. Although these methods are computationally efficient, they often fail to capture feature dependencies or interactions between features. Consequently, they may not perform optimally in highly complex datasets, such as those found in genomics.

Wrapper Methods: Wrapper methods evaluate subsets of features based on their impact on model performance. A typical wrapper method involves training a model on various subsets of features and selecting the subset that produces the best results. Recursive Feature Elimination (RFE) [16] and Genetic Algorithms [17] are popular wrapper methods in cancer genomics. While wrapper methods often produce better results than filter methods, they can be computationally expensive, especially when applied to large genomic datasets containing thousands of biological features.

Embedded Methods: Embedded methods integrate feature selection directly into the model training process. Examples include LASSO [4], which applies an L1 penalty to shrink less important feature weights to zero, and decision trees [18], which inherently perform feature selection based on information gain or other splitting criteria. Embedded methods are popular in genomics due to their ability to balance feature selection with model performance. Although PCA [19] is not a primary feature selection method, it reduces dimensionality by transforming features into uncorrelated principal components.

III. Methodology

A. Reinforcement Learning for Feature Selection

In this paper, we adapt the reinforcement learning for feature selection [7] and apply it for the aforementioned cancer genomic datasets. [7] presents a multi-agent reinforcement learning (MARL) approach, which automates the feature subspace exploration, improving both the efficiency of feature selection and the accuracy of predictive models. [7] treats each feature as an independent agent in the multi-agent framework. These agents make decisions (actions) about whether to select or deselect their corresponding features. Compared with single-agent RL, where each agent requires evaluating the entire feature space, multi-agent RL reduces the action space by allowing each agent to handle one feature at a time. The agents work collaboratively and competitively to select the most optimal subset of features. When designing RL approach, there are two major things to consider: state representation and reward function.

State Representation. A state indicates the current feature subspace (*i.e.*, the subset of selected features). The agents must learn from this state to make decisions. In [7], three methods are proposed for constructing state representations: (a) *Meta Descriptive Statistics (MDS)*: This method computes first-level descriptive statistics (*e.g.*, mean, variance) for each selected feature and then further derives second-level statistics from the first-level features. This ensures that the state representation has a fixed length, regardless of how many features are selected at any given time. (b) *Autoencoder-Based Deep Representation (AE)*: A two-level autoencoder approach is used to learn a latent representation of the selected feature subspace. The first autoencoder encodes the selected features into a latent space, and the second autoencoder reduces the dimensionality of this latent representation to ensure a fixed-length state vector. (c) *Graph Convolutional Network (GCN)*: This method captures relationships between features by constructing a dynamic graph where nodes represent features, and edges represent their correlations. A GCN is then applied to learn a latent representation of the selected feature subset. This method is particularly effective in modeling dependencies between features.

B. Extended Reinforced Learning for Cancer Genomics

[7] was designed as a general RL method for feature selection. We extended it to cancer genomics. Figure 1 shows the adapted MARL framework from [7], which operates in two main stages.

(1) *Control Stage: Feature Selection, State Generation, & Transitions Generations*: Each agent selects or deselects its corresponding feature based on its policy network, which takes the current state as input. The selected feature subset is considered the environment, and the actions of all agents lead to a new state. The agents then receive a reward based on the performance of the feature subset. Here, the reward is mainly measured by predictive accuracy of the selected features, compared with the entire features, as well as feature redundancy and relevance, etc. For state generation, according to [7], Meta Descriptive Statistics (MDS) plus Autoencoder-Based Deep Representation (AE) obtains the best performance as it takes both explicit and implicit information into consideration from the selected features. We also adopt this method. In Figure 1, $Q_1, Q_2, ..., Q_n$ indicates the statistical measures (such as standard deviation, minimum, maximum, quartiles) with respect to data samples; $Q_1, Q_2, ..., Q_m$ indicates the statistical measures with respect to features. Similarly, $Latent_1, ..., Latent_t$ is the encoded latent values with respect to data samples, where $Latent_a, ..., Latent_k$ are with respect to features.

(2) *Training Stage*: Each agent independently updates its policy network by performing experience replay on mini-batches of samples from its memory. The Deep Q-Network (DQN) algorithm is used to update the agents' policies, where the goal is to maximize the long-term reward. The control stage and training state cooperate to finally find out the best feature subset with highest reward.

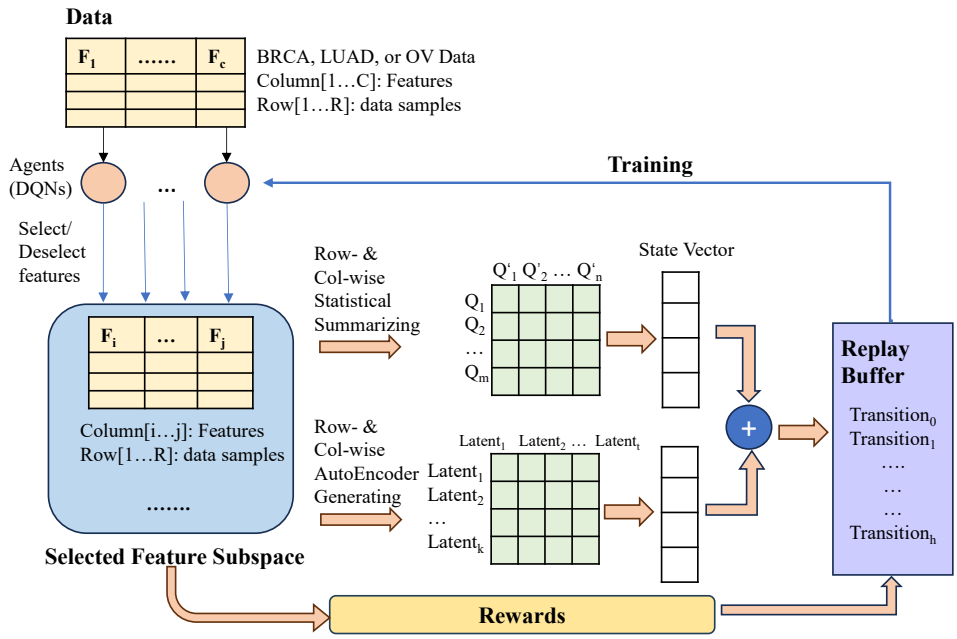

Fig. 1: The Framework of the Reinforcement Learning Approach for Automated Feature Selection in Cancer Genomics.

C. Comparative Algorithms

We compared the performance of our RL-based feature selection algorithm with the following traditional methods:

- *Least Absolute Shrinkage and Selection Operator (LASSO)*: A popular feature selection method utilizing regularization. It is based on L1 regularization, which shrinks coefficients of less important features to zero. In particular, λ is the regularization parameter that controls the strength of the penalty. λ controls the trade-off between achieving a good fit and penalizing the complexity of the model. A higher λ increases regularization, resulting in fewer selected features. LASSO is particularly useful in high-dimensional datasets, where many features may be irrelevant or redundant.
- *Recursive Feature Elimination (RFE)*: A feature selection technique that aims to identify the important features by recursively removing the least important ones. It works by training a model, ranking the features based on their importance, and then eliminating the least significant features in a step-by-step process. RFE is a powerful tool for improving model performance by selecting the most relevant features while eliminating noise and reducing overfitting.
- *Random Forrest (RF)*: An ensemble learning method that consists of multiple decision trees built from random subsets of data and features, and aggregates their predictions to improve accuracy and reduce overfitting. One of the key advantages of Random Forest is its ability to rank features by their importance. Each decision tree in the forest uses different subsets of features, and the importance of a feature is calculated based on how much it contributes to reducing the impurity (*e.g.*, Gini impurity or entropy) in the tree's decision nodes. Random Forest can capture complex interactions between features, making it robust for feature selection in datasets with nonlinear relationships. Unlike some other methods, Random Forest does not require explicit feature selection; it automatically ranks and selects important features based on its internal mechanism.
- *Principal Component Analysis (PCA)*: A dimensionality reduction technique used to transform a large set of features into a smaller set while preserving as much variance as possible. However, PCA is not a feature selection method in the traditional sense because it doesn't select original features but rather creates new features (called principal components) that are combinations of the original features. PCA could be effective when the goal is to reduce dimensionality and improve model efficiency, particularly in cases where there are many correlated features. However, it is not ideal for selecting individual and interpretable features.

IV. Experiments

A. Datasets

The RL framework and baseline methods were tested on TCGA datasets for breast invasive carcinoma (BRCA) [8], lung adenocarcinoma (LUAD) [9], and ovarian serous cystadenocarcinoma (OV) [10]. RNA-seq mRNA expression data for each cancer type was downloaded from the UCSC Xena Hub [20], using transformed RSEM normalized counts for 20,531 genes. Clinical information for the three cancer studies was obtained from cBioPortal [21]. In the breast cancer study, patients were classified based on estrogen receptor status (ER+ vs. ER-) and triple-negative status (TN+ vs. TN-). Triple-negative breast cancer patients lack estrogen receptors,

progesterone receptors, and excess HER2 protein. For lung and ovarian cancer studies, patients were classified based on their survival time.

B. Experiment Setting

In the experimental setup for all methods, the data is split into training and testing sets in a ratio of 80:20. Classifiers are trained on the training set using the selected genomic features from large-scale gene expression datasets and tested on the testing set with the same selected features. All approaches except for RL are based on scikit-learn [22]. Additional settings are shown below for these methods.

- *LASSO*: We use cross validation to find the optimal regularization weight λ, which is set to 0.03 in the experiments.
- *RFE*: The number of selected features is set to 100, and the maximum iterations is set to 500 in our experiments. Logistic regression classifier is applied for prediction.
- *Random Forrest (RF)*: The hyperparameter in *Random Forest* is the number of estimators, which is set to 100 in the experiments.
- *PCA*: The number of principle components are set to 10, 50, and 30 for BRCA, LUAD and OV datasets, respectively, and the logistic regression classifier is applied for the final classification.
- *Multi-agent Reinforcement Learning*: We set the batch size during training to 32 and use the Adam Optimizer with the learning rate of 0.01.

C. Evaluation Metrics

To evaluate the effectiveness of each feature selection method, we use is a we used the following metrics:

- *Accuracy*: The classification accuracy of a predictive model using the selected features. Accuracy measures the overall correctness of the model by determining the ratio of correctly classified instances to the total instances.
- *Precision*: The proportion of true positives among all positive predictions. It measures the accuracy of positive predictions.
- *Recall*: The proportion of true positives out of all actual positive instances. It measures how well the model captures all the actual positive instances.
- *F-measure (F1 Score)*: A measure of model performance that considers both precision and recall. It is particularly useful when there is an uneven class distribution.
- *AUC*: A metric used to evaluate the performance of a classification model, especially for imbalanced data.

The outcomes of these metrics depend not only on feature selection but also on the choice of predictors. In this study, we employ a random forest model with 100 decision trees as the predictor.

D. Experimental Results

We experimented on six datasets and reported *Accuracy*, *precision*, *recall*, and the macro *F1-score* in Tables I, II and III. From the experimental results, we can draw the following conclusions: 1) *Multi-agent Reinforcement Learning performs the best.* MARFLS performs the best in term of most comprehensive metrics such as AUC, accuracy and F1-score, which demonstrates its ability to identify the most relevant features for cancer prognosis while maintaining or improving model performance. Specifically, if we considering the AUC, the overall performance ranking of these algorithms on the 6 datasets (BRCA-ER, BRCA-HER2, BRCA-PR, BRCA-TN, LUAD, and OV) is MARLFS $>$ RFE $>$ LASSO/RF/PCA; if considering accuracy, the ranking is MARLFS $>$ LASSO/RFE/RF $>$ PCA; if considering F1-score, the ranking is MARLFS $>$ LASSO/RFE/PCA $>$ RF.

2) *RFE performs well among all the methods, second only to MARLFS.* RFE recursively removing the least important features in order to identify the important features. It performs well on the datasets. However RFE is computation-instensive

TABLE I: Experimental results on breast cancer dataset (BRCA). The table presents the classification results for four breast cancer subtypes, using the genes selected by each method.

Dataset	Method	AUC	Accuracy	Precision	Recall	F1-score
BRCA(ER)	*LASSO*	0.8833	0.9277	0.9161	0.8886	0.9012
	RFE	0.8614	0.9036	0.8725	0.8725	0.8725
	RF	0.8329	0.8916	0.8705	0.8329	0.8492
	PCA	0.8644	0.8916	0.8536	0.8644	0.8588
	MARLFS	**0.8886**	**0.9640**	**0.9800**	**0.8833**	**0.9163**
BRCA(HER2)	*LASSO*	0.7083	**0.9156**	**0.9551**	0.7083	**0.7706**
	RFE	**0.7197**	0.8675	0.7311	**0.7148**	0.7224
	RF	0.6250	0.8916	0.9438	0.6250	0.6702
	PCA	0.6180	0.8795	0.8180	0.6180	0.6542
	MARLFS	0.6346	0.8927	0.9264	0.6346	0.6681
BRCA(PR)	*LASSO*	0.8300	0.8554	0.8634	0.8300	0.8410
	RFE	**0.8801**	0.8915	0.8927	**0.8768**	**0.8834**
	RF	0.8045	0.8313	0.8354	0.8045	0.8145
	PCA	0.8357	0.8554	0.8558	0.8358	0.8434
	MARLFS	0.8340	**0.8926**	**0.9117**	0.7840	0.8157
BRCA(TN)	*LASSO*	0.7767	0.8554	0.7999	0.7767	0.7872
	RFE	0.8332	0.8916	0.8504	0.8372	0.8435
	RF	0.7845	0.8675	0.8228	0.7845	0.8009
	PCA	0.7582	0.8434	0.7840	0.7504	0.7647
	MARLFS	**0.9000**	**0.9757**	**0.9871**	**0.9000**	**0.9267**

TABLE II: Experimental results on lung cancer dataset (LUAD).

Method	AUC	Accuracy	Precision	Recall	F1-score
LASSO	0.5000	**0.9231**	0.4615	0.5000	0.4800
RFE	0.5000	**0.9231**	0.4615	0.5000	0.4800
RF	0.5000	**0.9231**	0.4615	0.5000	0.4800
PCA	0.5000	0.8846	**0.6449**	**0.7083**	0.6681
MARLFS	**0.5583**	0.8867	0.6433	0.7000	**0.6687**

TABLE III: Experimental results on the ovarian cancer dataset (OV).

Method	AUC	Accuracy	Precision	Recall	F1-score
LASSO	0.5992	0.6087	**0.6923**	0.5992	0.5965
RFE	0.5992	0.6087	0.5962	0.5992	0.5965
RF	0.5595	0.6087	0.5735	0.5595	0.5548
PCA	0.6706	0.6957	0.6792	0.6706	**0.6734**
MARLFS	**0.7194**	**0.7000**	0.4970	**0.7000**	0.5794

due to its recursive execution.

3) *LASSO also performs well among all the methods, ranking just behind RFE.* For our datasets with high dimension for which the number of features far exceeds the number of samples, LASSO's can effectively prevent overfitting and improve the model's generalization ability.

4) *Both random forest and PCA are not good at handling the gene expression datasets.* Random forest works lightly better than PCA but worse than others. Random Forest provides intrinsic measures of feature importance during training, and does not require explicit feature selection. As a dimensionality reduction technique which preserves most of the variance information through principal components, PCA may result in the loss of some specific features that are crucial for classification. In our experiments, one tricky problem is the selection of the number of principal components, which has a significant impact on the performance.

Note that none of the methods show good AUC in Table II due to the extremely unbalanced dataset in LUAD.

V. Discussion

The results indicate that RL presents a powerful alternative to traditional feature selection techniques. By modeling feature selection as a sequential decision-making problem, the RL approach effectively explores complex feature interactions, leading to the development of more accurate models. This capability is advantageous in high-dimensional gene expression datasets, where the relationships between genes are biologically relevant.

Additionally, the RL algorithm reduces the number of features needed while maintaining accurate predictions, highlighting its efficiency. This reduction is important in biomedical research, where identifying a smaller number of more biological meaningful features is essential for constructing robust predictive models. Such models can significantly enhance clinical decision-making, as they focus on the most relevant genomic features that influence disease outcomes.

Importantly, the selected genes from large-scale cancer patients gene expression data, can serve as promising targets for drug development. By identifying key genes that play critical roles in cancer progression or response to treatment, researchers can prioritize these targets for further investigation. This targeted approach can simplify the drug discovery process and open up new strategies for personalized medicine.

VI. Conclusion

In this study, we introduced a reinforcement learning-based approach for feature selection, demonstrating its effectiveness on large-scale gene expression datasets from breast cancer, lung adenocarcinoma, and ovarian cancer. Our reinforcement learning algorithm performed comparably to or better than traditional feature selection methods, such as LASSO, RFE, Random Forest, and PCA, in terms of accuracy, precision, recall, F1 score, and AUC. These findings suggest that reinforcement learning has the potential to enhance feature selection in large-scale genomic data for cancer studies, potentially leading to improved prognostic models and more personalized treatment strategies. Future work will explore the application of this reinforcement learning-based approach to additional cancer types and its integration into clinical pipelines.

References

[1] Z. Wang, M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics," *Nature reviews genetics*, vol. 10, no. 1, 2009.

[2] W. Zhang, J.-W. Chang, L. Lin, K. Minn, B. Wu, J. Chien, J. Yong, H. Zheng, and R. Kuang, "Network-based isoform quantification with RNA-seq data for cancer transcriptome analysis," *PLoS computational biology*, vol. 11, no. 12, p. e1004465, 2015.

[3] B. S. Shastry, "SNP alleles in human disease and evolution," *Journal of human genetics*, vol. 47, no. 11, pp. 561–566, 2002.

[4] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 58, no. 1, pp. 267–288, 1996.

[5] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine learning*, vol. 46, pp. 389–422, 2002.

[6] E. Azim, D. Wang, K. Liu, W. Zhang, and Y. Fu, "Feature interaction aware automated data representation transformation," in *Proceedings of the 2024 SIAM International Conference on Data Mining (SDM)*. SIAM, 2024, pp. 878–886.

[7] K. Liu, Y. Fu, P. Wang, L. Wu, R. Bo, and X. Li, "Automating feature subspace exploration via multi-agent reinforcement learning," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 207–215.

[8] Brigham & Women's Hospital & Harvard Medical School, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 490, no. 7418, pp. 61–70, 2012.

[9] Cancer Genome Atlas Research Network and others, "Comprehensive molecular profiling of lung adenocarcinoma," *Nature*, vol. 511, no. 7511, p. 543, 2014.

[10] Cancer Genome Atlas Research Network, "Integrated genomic analyses of ovarian carcinoma," *Nature*, vol. 474, no. 7353, p. 609, 2011.

[11] W. Zhang, N. Johnson, B. Wu, and R. Kuang, "Signed network propagation for detecting differential gene expressions and dna copy number variations," in *Proceedings of the ACM conference on bioinformatics, computational biology and biomedicine*, 2012, pp. 337–344.

[12] K. Tadist, S. Najah, N. S. Nikolov, F. Mrabti, and A. Zahi, "Feature selection methods and genomic big data: a systematic review," *Journal of Big Data*, vol. 6, no. 1, pp. 1–24, 2019.

[13] M. Onesime, Z. Yang, and Q. Dai, "Genomic island prediction via chi-square test and random forest algorithm," *Computational and Mathematical Methods in Medicine*, vol. 2021, no. 1, p. 9969751, 2021.

[14] J. G. Thomas, J. M. Olson, S. J. Tapscott, and L. P. Zhao, "An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles," *Genome Research*, vol. 11, no. 7, pp. 1227–1236, 2001.

[15] I. Jain, V. K. Jain, and R. Jain, "Correlation feature selection based improved-binary particle swarm optimization for gene selection and cancer classification," *Applied Soft Computing*, vol. 62, 2018.

[16] Y. Han, L. Huang, and F. Zhou, "A dynamic recursive feature elimination framework (drfe) to further refine a set of omic biomarkers," *Bioinformatics*, vol. 37, no. 15, pp. 2183–2189, 2021.

[17] S. Sayed, M. Nassef, A. Badr, and I. Farag, "A nested genetic algorithm for feature selection in high-dimensional cancer microarray datasets," *Expert Systems with Applications*, vol. 121, pp. 233–243, 2019.

[18] Y. Saeys, I. Inza, and P. Larranaga, "A review of feature selection techniques in bioinformatics," *bioinformatics*, vol. 23, no. 19, 2007.

[19] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[20] M. J. Goldman, B. Craft, M. Hastie, K. Repečka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A. N. Brooks *et al.*, "Visualizing and interpreting cancer genomics data via the xena platform," *Nature biotechnology*, vol. 38, no. 6, pp. 675–678, 2020.

[21] J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson *et al.*, "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal," *Science signaling*, vol. 6, no. 269, pp. pl1–pl1, 2013.

[22] "Scikit-learn," https://scikit-learn.org/.