VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2024, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 12, Number 1, January-March 2024 pp: 25-35

An Intelligent Model for Predictions of PIWI-Interacting RNAs and Their Functions

Anam Umera , Sajid Mahmood , Usman Inayat *

School of Systems and Technology, University of Management and Technology C-II Jahor Town, Lahore, Pakistan

Keywords:
PIWI-interacting RNAs, Intelligent Model for PiRNAs, Prediction of PiRNAs, Machine learning based model, Functions of PIWI-interacting RNAs.

Journal Info:
Submitted:
June 15, 2022
Accepted:
September 16, 2022
Published:
September 20, 2022

Abstract

Small non-coding RNAs known as "piwi-interacting RNAs" (piRNAs) are essential for squelching transposable elements in animal germlines, preserving genomic integrity and fertility. The creation of new drugs and the identification of various tumor types are linked to the PiRNA molecules. Additionally, it is related to controlling transcription of genes, squelching transposons, and preserving genomic stability. The discovery of piRNAs and their functionality has grown to be a significant research topic in bioinformatics because of the crucial influence that piRNAs play in biological processes. The 2L-piRNA-ML predictor is a strong two-layer predictor that is suggested in this research to enhance the prediction of PiRNA and their functionality. The suggested model uses Quadratic Discriminant Analysis Classifier, Linear Discriminant Analysis, Passive Aggressive Classifier, Extra Tree Classifier, Logistic Regression, Random Forest, Ridge Classifier CV for classification. It also employs DNC and TNC for extraction of features. The suggested model is created using a two-layer construction strategy. The 1st layer makes a prediction about a given sequence whether it is PiRNA or not, and the 2nd layer makes a prediction about a given PiRNA sequence whether it is having the function of instructing target mRNA deadenylation or not. Proposed model achieved 95.65% accuracy at the first layer and 92.30% accuracy at the second layer.

***Correspondence author email address:** UsmanInayat (usman.inayat@umt.edu.pk)
DOI: 10.21015/vtse.v12i1.1720

1 Introduction

This document shows the suggested format and appearance of a manuscript prepared for KJCISjournals. Accepted papers will be professionally typeset. This template is intended to be a tool toimprove manuscript clarity for the reviewers. The final layout of the typeset paper will not matchthistemplate layout.

Almost all non-coding RNAs (ncRNAs) have drawn a great of interest for their involvement in cellular processes and disorder pathogenicity since the discovery that just 2–3 % of the mammalian chromosome is translated and ultimately converted into proteins [1]. All regulating ncRNAs can be roughly split into smaller and large ncRNAs depending on the molecule size [2].

In prokaryotic cells, PIWI-interacting RNA (PiRNA) is a documented class of short non-coding RNA molecules having a polymer that is 24 – 31 nucleotides long. PiRNAs carry out a wide range of genetic and biological tasks, such as gene inactivation, controlling the activation of genes, preserving, and forming genetic material, and synthesis of specific protein [3]. Recent studies indicate that piRNAs are widely transcribed in a variety of physiological cell types and are associated with a wide range of clinical conditions besides those that have been documented in the germ. PiRNAs, for instance, have been found to explain inordinately in many diseases [4]. PiRNAs have also demonstrated promise as prognostic indicators for a variety of tumors [5]. The dysregulation of piRNAs in cancer suggests that they may be potential targets for cancer therapy [6]. The most prevalent malignancy and the leading reason for tumor related deaths in women is breast tumor. PiRNAs perform a significant effect in breast tumor and may be used as diagnostics and treatment approaches since excessive transcription of piRNAs is seen in tumors and is linked to tumors cancer cells proliferation and spread [7].

Additionally, there is an indication that piRNA-mediated genomic mechanisms contribute to cancer. In breast tumor cells, piR-651 was shown to be substantially amplified. One of the biggest reasons of cancer-related deaths worldwide is lung cancer [8]. A possible diagnosis sign and source of treatment strategies in lung cancer, excessive piRNA transcription is linked to growth and spread. In laboratory and in mammalian lung cancer progress might be inhibited by piR-55490, which was also inversely correlated with patients' survival. While other research has established that in lung cancer cells piR-651 is enhanced, piR-L-163 was the piRNA that was most frequently down regulated in long cell lung cancer when opposed to the equivalent non tumor lung cells [9].

The most well-studied piRNAs in gastric cancer are piR-651 and piR-823, which may be useful indicators with excellent sensitivity and specificity for identifying gastric cancer and differentiating across various kinds of gastric cancer [10]. PiRNA also plays an important role in maintaining genome stability and controlling transposable element activity in various organisms, including humans [11]. PiRNAs also connected to several pathologic diseases which affect the heart. Moreover, the presence of piRNA in cardiovascular precursor cells raises the possibility that piRNA contributes to the survival and development of cardiac myocyte as well as the processes of heart recovery. Additionally, it has been discovered that the perivascular heart contains piRNAs. PiRNA is also expressed significantly in patients with heart frustration blood liposome. PiRNAs could possibly have a role in the indicators of heart disease [12]. PiRNAs are inquisitively also controllers of pancreas beta neuron function, according to a previous study.

For recognizing piRNA and non-piRNA sequences, numerous computational models have been proposed. To predict piRNA, a sequential computational model called "piRNA-CNN" based on convolutional neural network is proposed in [13]. DNC and TNC are used for feature extraction. In [14–16] proposed a powerful predictor called 2L-piRNA.

It is a two-layer ensemble model, with the 1st layer used to determine whether a query RNA molecule is a PiRNA or non-PiRNA, and the 2nd layer used to determine whether a piRNA can direct the deadenylation of a target mRNA or not. For classification, SVM classifier, CNN models are used and pseudo dinucleotide composition (k=2) for feature extraction. "2S-piRCNN", a two-stage deep-learning classifier that makes use of a CNN, is proposed in [17]. In [18] "2L-piRNAPred" model is proposed. 2L-piRNAPred is SVM-based predictor. 2-layer merged Scheme for point out piRNAs in the 1st layer and identifying if piRNA have the task of directing selected mRNA deadenylation in the 2nd layer. In [19] proposed "piRNN" for identifying piRNA. Convolutional neural network classifiers were used, each of which had been trained using datasets from four organisms. Each sequence was represented by a matrix of k-mer frequencies. In [20] proposed "2L-piRNADNN" model, which is based on a deep neural network (DNN), to identify piRNAs. Using conventional learning techniques, the 2L-piRNADNN classifier automatically extracts important features

from RNA sequences. In addition, the suggested model creates a feature vector using the di-nucleotide auto covariance (DAC) approach and six physic - chemical characteristics. First, using a benchmark dataset, the accuracy of 2L-piRNADNN model is matched to that of well used classifier techniques. Second, its accuracy is evaluated against current cutting-edge computer frameworks. The results of experiment demonstrate that, with accuracy levels of 91.81% and 84.52% at the first and second levels, correspondingly, the suggested approach outperformed the current predictors.

In [21] "2L-piRNADNN" model proposed. To reduce computing complexities via parallel processing, the DNN model employing the Spark computing platform is proposed. The RNA sequences are converted into a feature vector of numeric values by the suggested model's use of the dinucleotide auto covariance approach. In [22] created a quick, reliable, and effective deep learning technique called piRDA for locating the correlations between piRNAs and diseases. Without using any features engineering, the suggested architecture takes the most important and generic information from the unprocessed sequences expressed in a piRNA disease pair. K-fold cross-validation is used to assess the effectiveness of the suggested approach piRDA. In contrast to community methods, the piRDA greatly outperforms them all in terms of quality assessment criteria for the detection of piRNA disease connections. In this [23] study, they introduce a unique technique called piRNAdetect for accurate computerized prediction of piRNAs in genomic sequences. In the suggested approach, they first categorize piRNA sequences in the training dataset that have comparable sequence patterns, and then they extract useful prediction characteristics by using n-gram classifiers. piRNAdetect model achieved 84.40% accuracy.

In [14], they provide an integrated strategy for piRNA prediction that considers a range of genetic and epigenetic traits that can be utilized to describe these molecules. They examined and gathered a sizable a variety of piRNA characteristics that have been empirically verified in numerous species. In an object-oriented framework that uses a Various Kernel Learning technique, these properties are expressed by several kernels. The developed tool, known as IpiRId, outperforms all other tools with prediction findings that reach more than 90% accuracy for the three examined species. Additionally, their method enables researchers to examine the applicability of each specified trait in a particular species.

The proposed method can also be modified to anticipate different types of ncRNAs because it is modular and easily expandable. IpiRId model performance in term of Acc, Sp, Sn is 93.66%, 96.58%, 90.74% respectively. For classification, SVM classifier, CNN models are used and pseudo dinucleotide composition (k=2) for feature extraction. "2S-piRCNN", a two-stage deep-learning classifier that makes use of a CNN, is proposed in [24]. The 2lpiRNApred a combined algorithm with two layers' approach. In the first layer it recognizes piRNAs and assesses if they are involved in the second layer process of inducing target mRNA deadenylation. To make the attributes' dimensions smaller, a new feature extraction approach based on Luca fuzzy entropy and Gaussian participation function (LFE-GM) was presented. Two types of classifiers—Sparse Recognition Algorithm (SRA) and Five attribute detection techniques using the Support Vector Machine with Mahalanobis Proximity Dependent Rotational Basis Process: Extended serial similarity pseudo dinucleotide composition, k-mer, general series similarity pseudo dinucleotide composition, Standardized Moreau Broto auto-correlation, and Geary auto-correlation—were combined to create the unified classifier method with two layers, The outcomes show that 2lpiRNApred outperforms six other available prediction methods by a wide margin. 2lpiRNApred model performance in term of Acc, Sp, Sn, MCC is 88.72%, 85.54%, 91.89%, 0.775 respectively.

The roles of piRNAs in the regulatory and post transcriptional control of transposons and genome are currently the subject of [25] study. They created piRBase, a database aiding piRNA operational investigation, to gather and evaluate these data. Since its debut in 2014, piRBase has merged 264 data sets from 21 different species and accumulated 173 million piR-

NAs. The most recent piRBase version (V2.0, 2018) put more of an emphasis on the thorough identification of piRNA sequences and the growing quantity of piRNAs. The piRNA receptors and disorder piRNA information was also included in the piRBase version V2.0.

In [26–28], they hypothesize that piRNAs may have a role in RSA pathogenicity by targeting genes that affect adherence and extrinsic matrix components, according to a bioinformatics study. These findings will act as the theoretical underpinning for piRNA-targeted treatment approaches for RSA. To enhance the prediction of piRNAs and their function using deep learning techniques, a two-layer predictor is proposed in this [29] study. The suggested approach uses a variety of feature extraction techniques to consider the physical and chemical characteristics of the biological sequences while extracting features. The k-fold cross-validation approach is used to thoroughly assess the suggested strategy output. The evaluation's findings indicate that the suggested model outperformed the current models, with accuracy gains of 7.59 and 2.81 percent at layer 1st and layer 2nd, correspondingly. The suggested paradigm is believed to be useful for precision medicine and cancer diagnostics.

In [30], they address current developments in our knowledge of piRNA function as well as prospective testing and treatment uses of piRNAs in a variety of digestion malignancies. In order to predict connections between piRNAs and disorders using information-retrieving technology, they present a unique predictor dubbed iPiDA-LTR. According to research findings, iPiDA-LTR shows promise in detecting disorders linked to both previously identified and newly discovered piRNAs. The biogenesis, effective roles, and new developments of piRNA/PIWI complicated proteins in human disorders have been discussed in [31]. A gold standard piRNA collection for six species is presented to facilitate more productive piRNA research. The details on piRNA clusters, piRNA variations, splicing junction piRNAs, and piRNA profiles of expression from other datasets are presented in [32]. In [33], a framework has been proposed in which the actions of ancient PIWI proteins were independent of cell regulation and extensively expressed. The authors have discovered organisms that are precisely linked to Drosophila melanogaster have lost both the PIWI gene Ago3 and the Yb gene [34]. They demonstrate that even in the absence of Yb, the RNA remains available to produce transposon antisense piRNAs in large quantities in the soma.

The 2L-piRNA-ML predictor is a strong two-layer predictor that is suggested in this research to enhance the prediction of PiRNA and their functionality. The suggested model uses Quadratic Discriminant Analysis Classifier, Linear Discriminant Analysis, Ridge Classifier CV, Passive Aggressive Classifier, Extra Tree Classifier, Logistic Regression, Random Forest for classification. It also employs DNC and TNC for extraction of features. The suggested model is created using a two-layer construction strategy. The 1st layer makes a prediction about a given sequence whether it is PiRNA or not, and the 2nd layer makes a prediction about a given PiRNA sequence whether it is having the function of instructing target mRNA deadenylation or not. Proposed model overall accuracy is 95.65% at the first layer and 92.30% at the second layer.

To do so, this contribution is organized as follows: Section 2 represent material and methods. Section 3 represent the results and discussion of the proposed model. Finally, conclusion and future work is presented Section in 4.

2 Materials and Methods

The following seven stages are used to complete the proposed model as shown in Figure 1: (1) data acquisition and analysis,(2) feature extraction (DNC, TNC), (3) training the ML models (Step 1), (4) generating new features on the basis of theML model's outputs of step 1 and the DNC feature, and (5) training the ML classifier (Step 2), (6) generating new features on the basis of the ML model's outputs of Step 2, (7) training the ML classifiers (step 3). The detail of these stages is given below. Same model and stages are used for both first layer and second layer prediction.

2.1 Data Acquisition and Analysis

The same dataset used in [18] is applied. There are 709 piRNA samples having the function of instructing target mRNA deadenylation (denoted as S^{+}_{inst}), 709

Figure 1. Proposed 2-layer model.

piRNA samples without this function (denoted as $S^{+}_{non-inst}$), and 1418 non-piRNA samples. PiRNA samples denoted as S^{+} and non-PiRNA samples denoted as S^{-}. Consequently, the datasets for this investigation can be described as follows:

$$S = S^{+} \cup S^{-} \quad (1)$$

$$S^{+} = PiRNA \quad (2)$$

$$S^{-} = NonPiRNA \quad (3)$$

$$S^{+} = S^{+}_{inst} \cup S^{+}_{non-inst} \quad (4)$$

S represent first layer dataset and S+ represent second layer dataset.

2.2 Feature Extraction

The piRNA and non-PiRNA sequence dataset is used to extract features, and two feature encoding techniques are used (i.e., DNC, TNC). In this research, iLearn web server is used for feature extraction. DNC is used to extract 16 features and TNC to extract 64 features in total.

- **Di-Nucleotide Composition (DNC)**
  Feature extraction approach that uses a set of 2 sequential nucleotides to represent an RNA sequence is called Di-Nucleotide Composition (DNC). The probability of each set, denoted by N1 N2 for the first set, N2 N3 for the second set, and so on, is calculated [12]. 16 features are provided by the di-nucleotide composition. It's described as: $D(a,b) = \frac{N(ab)}{(n-1)}$, $a,b \in A,C,G,T(U)$ where N(ab) is the number of Di-Nucleotide described by nucleic acid types a and b. For example, $ab \in AA, AC, ...L, GT, TT$ sequences.
- **Tri-Nucleotide Composition (TNC)**
  Another feature extraction approach is called Tri-Nucleotide Composition (TNC), which uses three pairs of sequential nucleotides to describe an RNA sequence. Each pair's probability is computed. For instance, the first set in an RNA sequence is N1 N2 N3, the second set is N2 N3 N4, and so on, producing a $4x4x4 = 64D$ matching features vector [10]. It's outlined as: $D(a,b,c) = \frac{N(abc)}{(n-2)}$, $a,b,c \in A,C,G,T(U)$ where N(abc) is the number of Tri-Nucleotide described by nucleic acid types $a, bandc$. For example, $abc \in AAA, AAC, AAG, \ldots, TTT$ sequences.

2.3 Train Machine Learning Classifiers

We merge DNC and TNS features and make a dataset. Before experiment, we check the sub features contribution in overall dataset. So, we select best top 10 sub features from dataset using Random Forest Regressor. Figure 2 presents the important sub features. All sub features range is 0.00 to 0.05, but CG sub feature range is 0.35. So, we normalize the CG sub feature with 0.15.

The model, as shown in Figure 3 is divided in three steps. In Step 1, we use dataset and apply four Machine learning classifiers: Quadratic Discriminant Analysis Classifier (QDA), Linear Discriminant Analysis (LDA), Passive Aggressive Classifier (PAC), Ridge Classifier CV (RCCV). In Step 2 to make a new dataset, we merge the results that are obtained from the 1st step classifiers with DNC feature. And seven Machine learning classifiers: Quadratic Discriminant Analysis Classifier (QDA), Linear Discriminant Analysis (LDA),

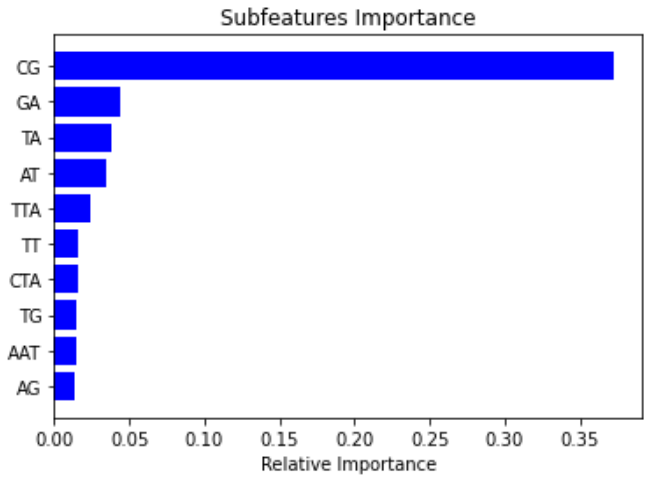

Figure 2. Top 10 important Sub-features with their relative importance.

Passive Aggressive Classifier (PAC), Extra Tree Classifier (ETC), Logistic Regression (LR), Random Forest (RF), Ridge Classifier CV (RCCV) are used. Now the results of second step are consider as a dataset and used in Step 3. In Step 3, we again used Logistic Regression classifier. The result of step 3 is considered final result.

The proposed model is developed on Jupyter notebook by using python programming language. Hold-out Method is used to split the dataset into two sets, the Training set and Test set. We split the 1st layer dataset into 20% for testing and 80% for training and 2nd layer dataset is split into 30% for testing and 70% for training.

3 Results and Discussion

In this section, the experimental results of the proposed model for both the first and second layers are discussed. As Figure 3 shows model is divided into three steps, so in this section the results of these three steps are provided.

3.1 1st Step Results of the Proposed Model

Table 1. shows the results of classifiers for both first layer and second layer.

Table 1. 1st step results of the proposed model.

Classifiers	1st Layer (ACC) %	2nd Layer (ACC) %
Quadratic Discriminant Analysis	85.73	68.0
Linear Discriminant Analysis	88.59	75.52
Ridge Classifier CV	86.80	76.0
Passive Aggressive	86.27	63.52

Figure 4 and Figure 5 represent the results of first step for both 1st and 2nd layers. These figures shows that Linear Discriminant Analysis classifier perform best for 1st layer and Ridge Classifier CV classifier for second layer.

3.2 2nd Step Results of the Proposed Model

Table 4 shows the results of classifiers for both first layer and second layer.

Table 2. 2nd step results of the proposed model.

Classifiers	1st Layer (ACC) %	2nd Layer (ACC) %
Quadratic Discriminant Analysis	92.03	73.43
Linear Discriminant Analysis	95.57	81.25
Passive Aggressive	92.03	80.46
Extra Trees	92.92	84.37
Logistic Regression	93.80	82.81
Random Forest	93.80	82.81
Ridge Classifier CV	93.80	82.81

Figure 6 and Figure 7 represent the 2nd step results of the proposed model for both 1st and 2nd layers. These figures shows that Linear Discriminant Analysis classifier also in 2nd step perform best for 1st layer. But in 2nd step Extra Trees perform best than Ridge Classifier CV classifier for second layer.

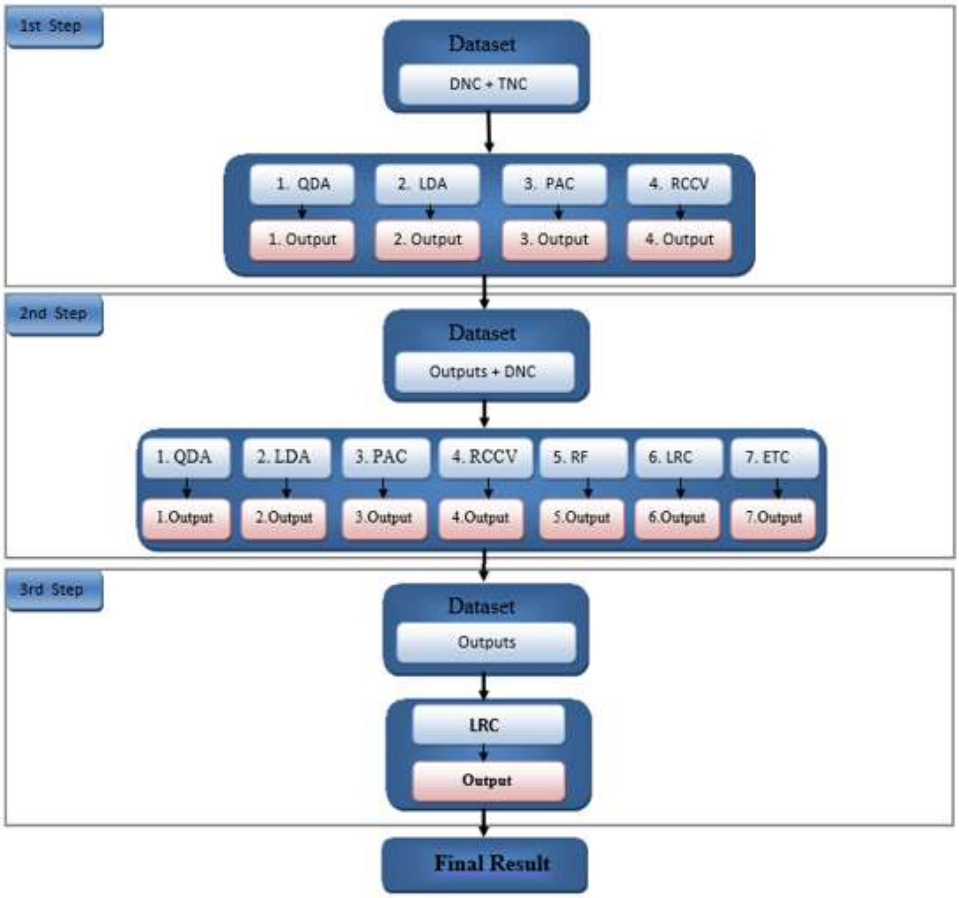

Figure 3. Structure of the Proposed 3-Layer Model.

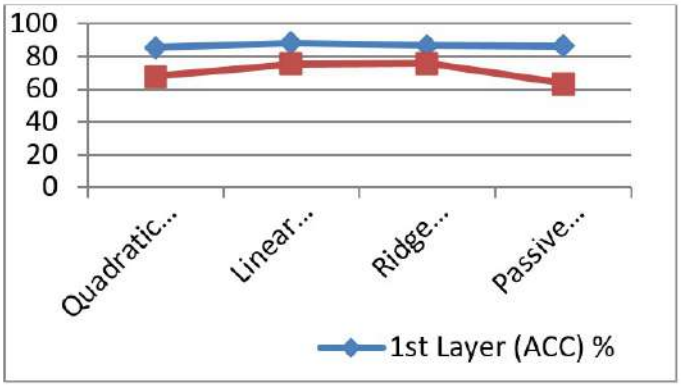

Figure 4. 1st step results of the proposed model.

3.3 3rd Step Results of the Proposed Model

Table 3 shows the results of classifiers for both first layer and second layer. 3rd step results for both 1st layer, and 2nd layer is considered the final results.

Table 3. Results from the 3rd step of the proposed model.

Classifiers	1st Layer (ACC) %	2nd Layer (ACC) %
Logistic Regression classifier	95.65	92.30

In step 3 we use Logistic Regression classifier. Figure 8 and Figure 9 Shows the 3rd step results of the proposed model for both 1st and 2nd layers.

Our findings demonstrate significant gains in prediction speed, specificity, and accuracy over current approaches. The model's ability to accurately predict piRNA structures across a broad variety of piRNA sequences highlights its potential impact on the fields of molecular biology and bioinformatics.

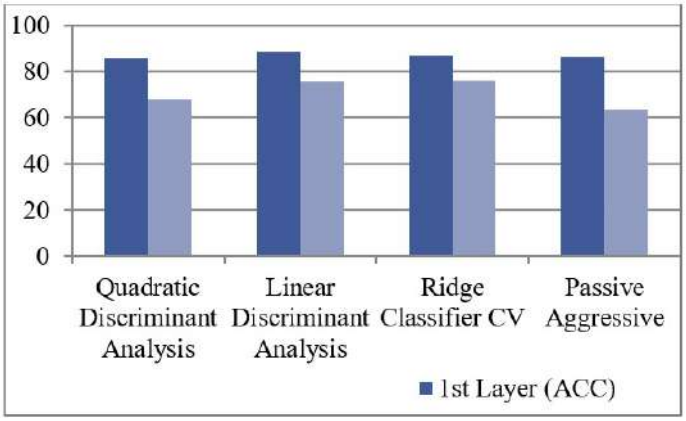

Figure 5. 1st step results using Bar chart.

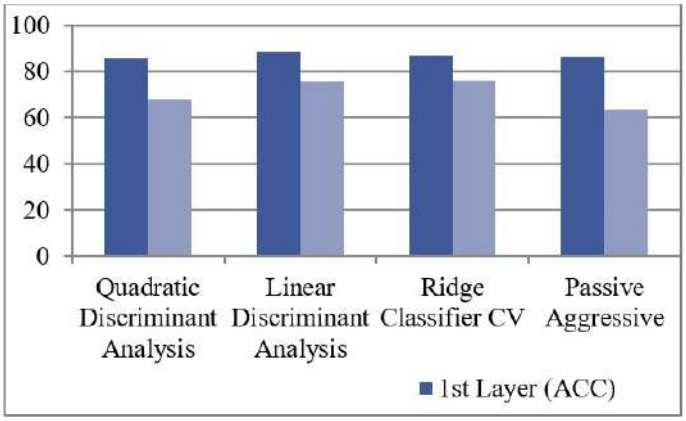

Figure 7. 2nd Step Results using Bar chart.

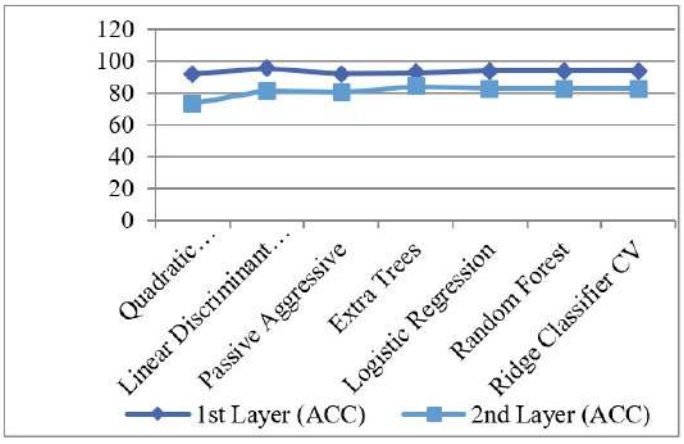

Figure 6. 2nd Step Results of the proposed model.

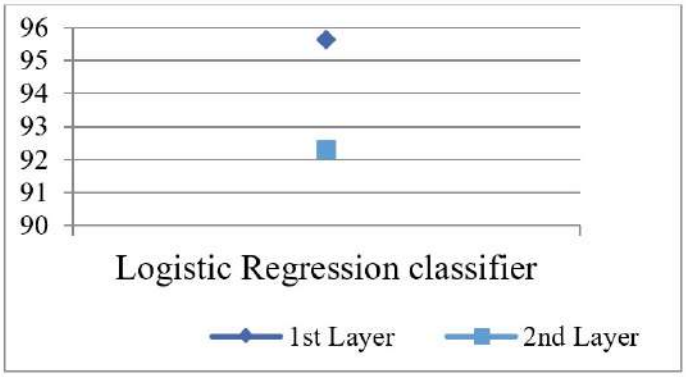

Figure 8. Results from the 3rd Step of the proposed model.

We use accuracy (ACC), sensitivity (Sn), Specificity (Sp), Matthews's correlation coefficient (MCC), Area under the curve (AUC), F1-Score, Confusion matrix and Precision metrics to check the model performance. Table 3. shows the results of both first and second layer. In step three of the proposed model logistic Regression Classifier is use for classification. So, the success rate of Logistic Regression Classifier for first-layer in terms of ACC, Sn, Sp, MCC, AUC, F1-Score, and Precision are 95.65%, 100.0%, 91.66%, 91.66%, 95.83%, 95.65%, 95.83%, respectively. Similarly, the success rate of Logistic Regression Classifier for second-layer in terms of ACC, Sn, Sp, MCC, AUC, F1-Score, and Precision are 92.30%, 94.73%, 90.0%, 84.73%, 92.36%, 92.30%, 92.36%, respectively.

Table 4. Comparison of the proposed 2L-piRNA-ML model with [18] benchmark model.

Layers	Models	Acc (%)	Sp (%)	Sn (%)	Mcc (%)
1st Layer	**2L-piRNA-ML**	95.65	91.66	100.0	91.66
	2L-piRNAPred	89	87.5	90.4	0.779
2nd Layer	**2L-piRNA-ML**	92.30	90.0	94.73	84.73
	2L-piRNAPred	84.0	83.6	84.3	0.680

4 Conclusion

In this research, 2L-piRNA-ML strong two-layer predictor is proposed that enhances the prediction of

Figure 9. 3rd Step Results using Bar chart.

piRNA and their functionality. The suggested model uses Quadratic Discriminant Analysis Classifier, Linear Discriminant Analysis, Passive Aggressive Classifier, Extra Tree Classifier, Logistic Regression, Random Forest, Ridge Classifier CV for classification. It also employs DNC and TNC for extraction of features. The suggested model is created using a two-layer construction strategy. The 1st layer makes a prediction about a given sequence whether it is PiRNA or not, and the 2nd layer makes a prediction about a given PiRNA sequence whether it is having the function of instructing target mRNA deadenylation or not. Proposed model overall accuracy is 95.65% at the first layer and 92.30% at the second layer. In future, to make 2L-piRNA-ML accessible to general biologists, we would like to set up a web service. Our model's updates will all be published on the GitHub website. We are motivated to enhancing 2L-piRNA-MLin order to provide improved piRNA prediction for other species in the future. And we are also enhancing the model's predictive power to uncover piRNA roles in diseases for therapeutic advancements.

Author Contributions

Anam Umera: Dataset, Software , Writing- Original draft preparation.**Sajid Mahmood**: Methodology, Conceptualization. **Usman Inayat**: Visualization, Investigation , Validation, Writing- Reviewing and Editing.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

Funding Information

No funding was received for this study.

References

[1] J. Wang, S. Zhu, N. Meng, Y. He, R. Lu, and G.-R. Yan, "ncrna-encoded peptides or proteins and cancer," *Molecular Therapy*, vol. 27, no. 10, pp. 1718–1725, 2019.

[2] E. López-Jiménez and E. Andrés-León, "The implications of ncrnas in the development of human diseases," *Non-coding RNA*, vol. 7, no. 1, p. 17, 2021.

[3] K. Wang, T. Wang, X.-Q. Gao, X.-Z. Chen, F. Wang, and L.-Y. Zhou, "Emerging functions of piwi-interacting rnas in diseases," *Journal of cellular and molecular medicine*, vol. 25, no. 11, pp. 4893–4901, 2021.

[4] W. Weng, H. Li, and A. Goel, "Piwi-interacting rnas (pirnas) and cancer: Emerging biological concepts and potential clinical implications," *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1871, no. 1, pp. 160–169, 2019.

[5] Y. Yu, J. Xiao, and S. S. Hann, "The emerging roles of piwi-interacting rna in human cancers," *Cancer Management and Research*, pp. 5895–5909, 2019.

[6] A. Hashim, F. Rizzo, G. Marchese, M. Ravo, R. Tarallo, G. Nassa, G. Giurato, G. Santamaria, A. Cordella, C. Cantarella *et al.*, "Rna sequencing identifies specific piwi-interacting small non-coding rna expression patterns in breast cancer," *Oncotarget*, vol. 5, no. 20, p. 9901, 2014.

[7] S.-J. Zhang, J. Yao, B.-Z. Shen, G.-B. Li, S.-S. Kong, D.-D. Bi, S.-H. Pan, and B.-L. Cheng, "Role of piwi-interacting rna-651 in the carcinogenesis of non-small cell lung cancer," *Oncology letters*, vol. 15, no. 1, pp. 940–946, 2018.

[8] H. Fathizadeh and Z. Asemi, "Epigenetic roles of piwi proteins and pirnas in lung cancer," *Cell & Bioscience*, vol. 9, no. 1, p. 102, 2019.

[9] E. C. Smyth, M. Nilsson, H. I. Grabsch, N. C. van Grieken, and F. Lordick, "Gastric cancer," *The Lancet*, vol. 396, no. 10251, pp. 635–648, 2020.

[10] J. X. Feng and N. C. Riddle, "Epigenetics and genome stability," *Mammalian Genome*, vol. 31, no. 5, pp. 181–195, 2020.

[11] J. Yang, F. Xue, Y. Y LI, W. Liu, and S. Zhang, "Exosomal pirna sequencing reveals differences between heart failure and healthy patients." *European Review for Medical & Pharmacological Sciences*, vol. 22, no. 22, 2018.

[12] M. Tahir, M. Hayat, S. Khan, and K. to Chong, "Prediction of piwi-interacting rnas and their functions via convolutional neural network," *IEEE Access*, vol. 9, pp. 54 233–54 240, 2021.

[13] B. Liu, F. Yang, and K.-C. Chou, "2l-pirna: a two-layer ensemble classifier for identifying piwi-interacting rnas and their function," *Molecular Therapy-Nucleic Acids*, vol. 7, pp. 267–277, 2017.

[14] S. D. Ali, W. Alam, H. Tayara, and K. T. Chong, "Identification of functional pirnas using a convolutional neural network," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 19, no. 3, pp. 1661–1669, 2020.

[15] M. F. Zia, U. Inayat, W. Noor, V. Pangracious, and M. Benbouzid, "Locational detection of false data injection attack in smart grid based on multilabel machine learning classification methods," in *2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT)*. IEEE, 2023, pp. 1–5.

[16] U. Inayat, M. F. Zia, S. Mahmood, H. M. Khalid, and M. Benbouzid, "Learning-based methods for cyber attacks detection in iot systems: A survey on methods, analysis, and future prospects," *Electronics*, vol. 11, no. 9, p. 1502, 2022.

[17] T. Li, M. Gao, R. Song, Q. Yin, and Y. Chen, "Support vector machine classifier for accurate identification of pirna," *Applied Sciences*, vol. 8, no. 11, p. 2204, 2018.

[18] K. Wang, J. Hoeksema, and C. Liang, "pirnn: deep learning algorithm for pirna prediction," *PeerJ*, vol. 6, p. e5429, 2018.

[19] S. Khan, M. Khan, N. Iqbal, T. Hussain, S. A. Khan, and K.-C. Chou, "A two-level computation model based on deep learning algorithm for identification of pirna and their functions via chou's 5-steps rule," *International Journal of Peptide Research and Therapeutics*, vol. 26, pp. 795–809, 2020.

[20] S. Khan, M. Khan, N. Iqbal, M. Li, and D. M. Khan, "Spark-based parallel deep neural network model for classification of large scale rnas into pirnas and non-pirnas," *IEEE Access*, vol. 8, pp. 136 978–136 991, 2020.

[21] S. D. Ali, H. Tayara, and K. T. Chong, "Identification of pirna disease associations using deep learning," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 1208–1217, 2022.

[22] C.-C. Chen, X. Qian, and B.-J. Yoon, "Effective computational detection of pirnas using n-gram models and support vector machine," *BMC bioinformatics*, vol. 18, pp. 103–109, 2017.

[23] H. Huang, X. Yu, X. Han, J. Hao, J. Zhao, G. Bebek, S. Bao, R. A. Prayson, A. M. Khalil, E. Jankowsky *et al.*, "Piwil1 regulates glioma stem cell maintenance and glioblastoma progression," *Cell reports*, vol. 34, no. 1, 2021.

[24] Y. Zuo, Q. Zou, J. Lin, M. Jiang, and X. Liu, "2lpirnapred: a two-layered integrated algorithm for identifying pirnas and their functions based on lfe-gm feature selection," *RNA biology*, vol. 17, no. 6, pp. 892–902, 2020.

[25] J. Wu, X. Liu, H. Lu, H. Nie, Y. Tang, Y. Tang, G. Song, L. Zheng, and W. Qin, "Small rna sequencing revealed aberrant pirna expression profiles in deciduas of recurrent spontaneous abortion patients," *Biocell*, vol. 46, no. 4, p. 1013, 2022.

[26] S. Khan, M. Khan, N. Iqbal, M. A. A. Rahman, and M. K. A. Karim, "Deep-pirna: Bi-layered prediction model for piwi-interacting rna using discriminative features," *Comput. Mater. Contin*, vol. 72, pp. 2243–2258, 2022.

[27] M. Nasir and M. F. Zia, "Global maximum power point tracking algorithm for photovoltaic systems under partial shading conditions," in *2014 16th International Power Electronics and Motion Control Conference and Exposition*, 2014, pp. 667–672.

[28] M. F. Zia, M. Nasir, E. Elbouchikhi, M. Benbouzid, J. C. Vasquez, and J. M. Guerrero, "Energy management system for an islanded renewables-based dc microgrid," in *2020 2nd International Conference on Smart Power Internet Energy Systems (SPIES)*, 2020, pp. 543–547.

[29] A. Cai, Y. Hu, Z. Zhou, Q. Qi, Y. Wu, P. Dong, L. Chen, and F. Wang, "Piwi-interacting rnas (pirnas): Promising applications as emerging biomarkers for digestive system cancer," *Frontiers in Molecular Biosciences*, vol. 9, p. 848105, 2022.

[30] W. Zhang, J. Hou, and B. Liu, "ipida-ltr: Identifying piwi-interacting rna-disease associations based on learning to rank," *PLOS Computational Biology*, vol. 18, no. 8, p. e1010404, 2022.

[31] Z. Wu, X. Yu, S. Zhang, Y. He, and W. Guo, "Novel roles of piwi proteins and piwi-interacting rnas in human health and diseases," *Cell Communication and Signaling*, vol. 21, no. 1, p. 343, 2023.

[32] J. Wang, Y. Shi, H. Zhou, P. Zhang, T. Song, Z. Ying, H. Yu, Y. Li, Y. Zhao, X. Zeng *et al.*, "pirbase: integrating pirna annotation in all aspects," *Nucleic acids research*, vol. 50, no. D1, pp. D265–D272, 2022.

[33] S. Horjales, M. Li Calzi, M. Francia, A. Cayota, and M. Garcia-Silva, "pirna pathway evolution beyond gonad context: Perspectives from apicomplexa and trypanosomatids," *Frontiers in Genetics*, vol. 14, p. 1129194, 2023.

[34] S. Chary and R. Hayashi, "The absence of core pirna biogenesis factors does not impact efficient transposon silencing in drosophila," *Plos Biology*, vol. 21, no. 6, p. e3002099, 2023.