NextGen Scientific Publication
American International Journal of Computer Science and Technology
Volume 7 Issue 2, Pg. No. 177-186, AIJCST-V7I2P114, 2025
https://doi.org/10.63282/3117-5481/AIJCST-V7I2P114

Original Article

Hybrid Deep Learning Framework for Predicting CRISPR-Cas9 Off-Target Activities Using Experimentally Validated Mismatch and Indel Datasets

***Dr. Hemanshi Dhaduk**
Monroe University, King Graduate School, New Rochelle, NY, USA.

Article History:

Received: 02.03.2025

Revised: 28.03.2025

Accepted: 10.04.2025

Published: 17.04.2025

Abstract:

The CRISPR-Cas9 system has been discovered as a potential biological gene-editing method. However, there is still a considerable chance of unintended consequences. Models for off-target prediction have been developed recently thanks to advances in deep learning, although these models often fall short of fully using sequence pair information. The challenge of "off-target," however, hampers the swift development of gene editing technologies. In this paper. This study presents a successful work on classification with ML and DL models, considering a structured dataset and utilizing multiple performance measures. The proposed models Logistic Regression (LR) and Recurrent Neural Network (RNN) were compared with the several baseline models: ANN, Naïve Bayes (NB) and SVM. The data was preprocessed for reliable training and testing by data cleaning and encoding, normalization and balanced splitting. The experimental findings show that the suggested models work better than the current techniques: LR achieves an accuracy of 97%, and RNN achieves the best accuracy of 97.32%, precision of 98%, recall of 97.2%, and F1 score of 97.9%. The outcomes show that the models proposed here have the potential to improve classification and overall prediction performance.

Keywords:

CRISPR-Cas9, Off-Target Prediction, Deep Learning (DL), Lightweight Model, Hyperparameter.

1. Introduction

The adaptive immune systems of bacteria and archaea were used to create the CRISPR-Cas9 system, a revolutionary genome editing tool. Nowadays, it is extensively used in agricultural biotechnology, gene therapy, disease modeling, and functional genomics [1][2]. The Cas9 nuclease and CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) sequences work together in this system, which uses a single guide RNA (sgRNA) to locate and attach to certain DNA sites. The Cas9 enzyme may specifically target genes for deletion, insertion, or repair by introducing double-strand breaks (DSBs) at the identified site. Improving the safety and reliability of genome editing requires accurate prediction of off-target effects, which pose significant challenges for therapeutic and clinical applications.

To address off-target effects, computational prediction methods have been broadly categorized into heuristic, machine learning, and DL approaches [3][4]. Early heuristic techniques are based on empirical criteria that take into account variables including the number, location, and distribution of mismatches between sgRNA and possible off-target sites. Programs such as MIT and CCTop score off-target sequences based on specific criteria, which is done for PAM sequences and mismatches. Likewise, CROP-IT uses sequence properties and features along with structural properties for prediction. These rule-based approaches, however, are limited by the lack of complexity in the scoring systems and generalizability due to the inability to fully capture the complex interactions governing the specificity of CRISPR-Cas9 [5][6][7].

Copyright by the Author(s). This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-NC-ND 4.0) License (https://creativecommons.org/licenses/by-sa/4.0/)

Data-driven approaches have made significant advances in off-target prediction by combining experimentally validated datasets and various sequence-related features. These methods take into account the nucleotide composition, the GC content, the PAM context, and the structural properties to better identify potential off-target sites [8][9]. Despite such advances, these approaches are usually based on fixed criteria and may struggle to correctly reflect the complexity of genomic relationships, particularly in the context of high-dimensional, large-scale biological data. This constraint makes it hard for them to extrapolate to other datasets and other experimental conditions.

To address these challenges, existing computational models are increasingly making use of ML and DL methods to improve off-target prediction of CRISPR-Cas9. Machine learning techniques rely on statistical or data-based techniques that learn patterns from experimentally verified data sets, thus enhancing prediction performance compared to rule-based techniques [10]. Building upon this, deep learning approaches further advance the field by automatically extracting complex and hierarchical representations from raw sequence data, enabling more accurate identification of off-target sites, including those involving mismatches and insertions/deletions (indels) [11]. The integration of ML and DL not only improves predictive accuracy and scalability but also reduces dependence on manual feature engineering, making these approaches highly effective for handling large-scale genomic data. Consequently, ML and DL play a vital role in developing robust, efficient, and generalizable frameworks for CRISPR-Cas9 off-target activity prediction (ML, DL).

1.1. Problem Statement

The issue of CRISPR-Cas9 off-target effects is a serious problem as the off-target effects are still not predicted accurately due to the occurrence of mismatches and sgRNA-target DNA sequences, as well as, data imbalance problems. Current computational tools are generally ineffective in quantifying complicated sequence interactions, resulting in limited accuracy and generalizability of computational predictions of off-target effects.

1.2. Motivation and Contribution

The rationale behind the conduct of this study is that there is an urgent necessity to enhance accuracy and reliability of predicting off-target effects of CRISPR-Cas9 because unwanted edits can restrict its safe use in clinical and therapeutic fields. The known heuristic and conventional methods frequently cannot cover the complex sequence variants, particularly ones with the mismatch and insertions/deletions (indels). Furthermore, difficulties in data imbalance and generalizations are other issues that decrease prediction performance. Thus, the given work is driven by the need to create a sophisticated framework based on the application of ML and DL methods that would allow modeling forthcoming complex patterns and improving accuracy of off-target prediction. The following are some of the main contributions in this research:

- Effectively trained and evaluated the CRISPR/Cas9 Off-Target Dataset of 217,733 sequence pairs.
- Implementing overall data preprocessing techniques like data cleaning, encoding, normalization and data balancing for improving data quality.
- Developed an efficient classification model using Logistic Regression (LR) and Recurrent Neural Network (RNN).
- Used different evaluation metrics such as accuracy, precision, recall, F1score and ROC to ensure reliable and comprehensive evaluation of the performance.

This study provides strong *justification* by addressing the need for accurate and efficient classification models using a well-structured dataset and comprehensive preprocessing techniques. The *novelty* is the melding of both LR and RNN models, so that a fair comparison between the traditional and DL approaches can be made. Furthermore, the proposed framework shows enhanced performance and reliability, proving its effectiveness in comparison to existing techniques.

1.3. Organization of the Paper

The paper organized as follows: Section II evaluate pertinent literature, Section III entails the description of the dataset, pre-processing operations, and model development, Section IV present the experimental results and comparison analysis, and Section V conclude the research with some critical insights as well as recommendations to future research.

2. Literature Review

To aid and enhance this study, a comprehensive review and analysis of the most relevant research studies on the prediction of off-target activity of CRISPR-Cas9 has been conducted. The summarized recent works, including the proposed models, datasets and the major findings with issues involved, are summarized in Table I.

G. Zhang et al. (2024) introduced Logistic Regression (LR) model for off-target activity prediction with performance value of 0.7687 (ROC_AUC), with stable and consistent classification results. The model is efficient to compute and has high interpretability, thus appropriate for use as a baseline model approach. Its performance is, however, less compared to sophisticated models like CRISPR-SGRU, which is able to capture complex sequence patterns[12].

Y. Yang et al. (2023) proposed Logistic Regression (LR) model with ROC_AUC is 0.7687, which is showing good classification baseline performance. However, LR offers stable and interpretable predictions, even if it is less successful than advanced models like HNN (ROC_AUC = 0.8731). Moreover, LR reliably achieves high accuracy even on unbalanced datasets, which makes it a good reference point for evaluating more sophisticated methods [13]. Similarly, A. Imani et al. (2023) proposed Logistic Regression (LR) model with a ROC_AUC of 0.7687, which shows the model's reliable baseline performance in classification. While its ROC_AUC is less than that of complicated models like HNN (0.8731), LR offers steady and understandable predictions. Moreover, LR gives good performance on any unbalanced dataset, making it a good baseline for measuring the performance of more sophisticated models [14].

K. Sanyour and W. Kim (2022) examined the needs of the CRISPR/Cas9 process to find optimal gRNAs to guide Cas9 to on-targets and steer clear of its off-targets. In order to anticipate the cleavage effectiveness of gRNA, many ML methods are crucial in the search for efficient gRNAs for the CRISPR/Cas9. Here, want to provide an overview and thorough study of many ML and DL techniques that are successful in forecasting the on-target actions of CRISPR gRNA sequences. Combining DL with other ML methods yields superior results in hybrid techniques, according to the findings of the comparison [15].

W. Xiang et al. (2021) allows the model to fully use the knowledge about sgRNA-DNA base pair matching. In comparison to DeepCRISPR, ROC-AUC and PR-AUC on the Mismatch data set rose by 0.6% and 41.9%, respectively. The ROC-AUC was 1.5% higher and the PR-AUC was 133.4% higher in comparison between CRISPR-Net and new Indels dataset test scenario [16]. Likewise, R. Niu et al. (2021) suggested CRISPR-Cas9 technology has substantially improved gene editing, yet off-target consequences continue to be a key problem. On the GUIDE_II dataset, the model AttnToMismatch_CNN achieved strong performance with an accuracy of 0.961 and error rate of 0.071, highlighting its effectiveness in off-target prediction [17]. Q. Liu et al. (2020) integrated the GloVe model with the sequence characteristics of sgRNA-DNA pairings, and used the taught word vector matrix to construct a DL model with five hidden layers that included biLSTM and CNN. Applying performance verification to the DeepCrispr data set, they discovered that auROC in the "leave-one-sgRNA-out" cross-validation could reach 0.957 and auPRC in the same environment 0.429 (with Pearson and spearman values of 0.495 and 0.151, respectively) [18]. While current models work well, there are still some research gaps in CRISPR-Cas9 off-target prediction. Common methods, such as LR, fail to model complex sequence dependence, and DL models can be expensive to compute and need large data sets. Furthermore, the majority of studies fail to consider data imbalance and generalization across multiple datasets. Thus, a balanced framework that integrates efficiency, accuracy and robustness for a reliable prediction of off-target is needed.

Table 1. Recent Studies on CRISPR-Cas9 Off-Target Prediction Using Machine Learning Techniques

Author (Year)	Model / Approach	Dataset / Task	Accuracy / ROC-AUC	Key Findings	Limitations & Future Work
G. Zhang et al. (2024)	Logistic Regression (LR)	Off-target prediction	ROC-AUC: 0.7687	Simple, interpretable, computationally efficient baseline model	Limited ability to capture complex sequence patterns; future work includes integrating deep learning or hybrid models for improved accuracy
Y. Yang et al. (2023)	Logistic Regression (LR)	Classification (imbalanced data)	ROC-AUC: 0.7687	Stable and consistent performance; effective for imbalanced datasets	Lower performance than advanced models (e.g., HNN); future work can focus on ensemble or deep models to improve predictive capability
A. Imani et al.	Word2Vec + CNN (Hybrid)	sgRNA classification	AUROC: 99.61%	High performance using embedding +	May require large training data and high computational

(2023)				CNN; outperforms CNNCRISPR	cost; future work can explore lightweight architectures and generalization across datasets
K. Sanyour & W. Kim (2022)	Hybrid ML + DL (Survey)	gRNA efficiency prediction	—	Hybrid approaches show strong performance in CRISPR prediction tasks	Lack of unified benchmark dataset; future work should focus on standardized evaluation and real-world validation
W. Xiang et al. (2021)	Deep Learning (sgRNA-DNA matching)	Mismatch & Indels datasets	ROC-AUC: +0.6%, +1.5%	Significant PR-AUC improvement; effective feature utilization	Model complexity and interpretability issues; future work can focus on explainable AI and reducing computational overhead
R. Niu et al. (2021)	AttnTo Mismatch_CNN	GUIDE_II dataset	Accuracy: 0.961	High accuracy with low error rate; strong prediction capability	Limited testing on diverse datasets; future work includes cross-dataset validation and robustness improvement
Q. Liu et al. (2020)	GloVe + BiLSTM + CNN	DeepCRISPR dataset	AUROC: 0.957	Effective sequence embedding and hybrid deep learning model	Lower PR-AUC compared to newer models; future work can enhance imbalance handling and feature representation

3. Research Methodology

In this study leverage a CRISPR-Cas9 off-target dataset, which represents paired sequences labeled, which followed by data preprocessing tasks, including cleaning, working with missing values, one-hot encoding, and normalization. A stratified method is then utilized to divide the processed dataset into training and testing data. RNN and LR are the two ML models implemented to learn patterns and conduct binary classification. Lastly, the models' performance is measured using acc, prec, rec, F1, and the ROC curve as measures of predictive effectiveness. The suggested flowchart for CRISPR-Cas9 prediction is shown in Fig. 1. Activities That Are Off-Target.

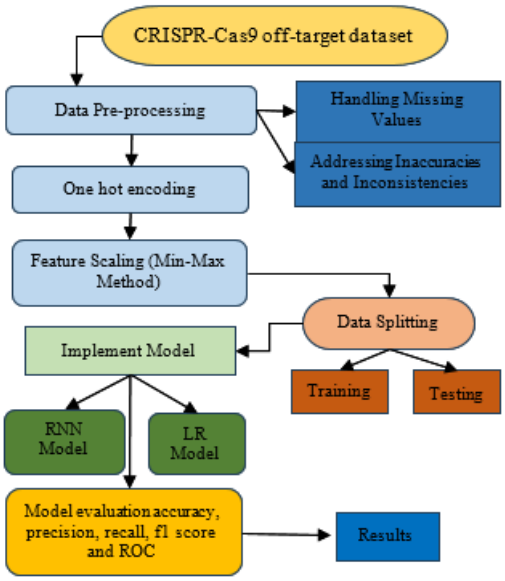


Figure 1. Proposed flowchart CRISPR-Cas9 Off-Target Activities prediction using machine learning

This section includes a detailed description of each of the steps in the suggested methodology:

3.1. Data Gathering and Analysis

This study employs on CRISPR-Cas9 off-target dataset for ML model construction and prediction. The dataset comprises 217,733 gRNA–DNA sequence pairs derived from nine guide RNAs, with each sequence being 23 base pairs long and classified as "1" for on-target and "0" for off-target. Data visualizations like bar graphs and heatmaps are used to investigate distribution patterns and feature correlations, and it is extensively utilized for assessing on-target efficiency and off-target consequences.

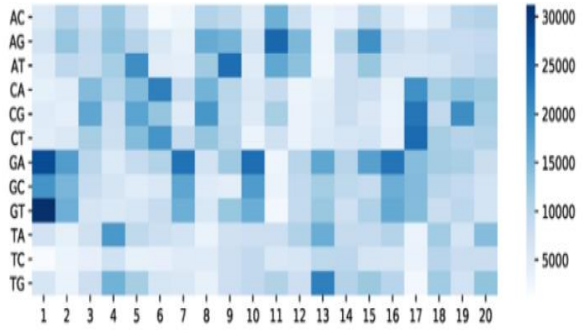


Figure 2. Heatmap of CRISPR/Cas9 off-target dataset

Fig. 2 shows the correlation trends in CRISPR/Cas9 off-target data, with correlations observed between various features. It aids in the determination of both strong positive and negative correlations, which are applicable to the feature selection and the optimization of the model.

In Fig. 3 shows the 15 highest features that are most influential in predicting the model on the test dataset. The features have varying effects, and some have a pronounced negative influence on others. Through this analysis, it is possible to understand the importance of features and enhance the interpretability of the model.

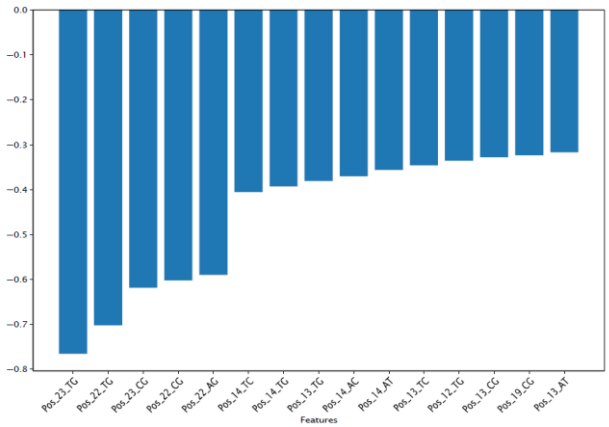


Figure 3. Top 15 features contributing most to the model's overall predictions

3.2. Data Pre-Processing

The dataset used in data preparation included concatenation, data cleaning, and feature engineering using the CRISPR/Cas9 off-target dataset. The preprocessing processes consisted of work with the missing values, the mistakes and inconsistencies, the data encoding, labeling and normalization. Preprocessing processes are summarized as follows:

- **Handling Missing Values:** This is the process of identifying partial data records and using suitable methods like imputation or deletion in order to safeguard the quality and integrity of the data.
- **Addressing Inaccuracies and Inconsistencies:** This step is aimed at detecting and rectifying errors, duplicate records, and inconsistency of data to achieve consistency and reliability in the entire dataset.

3.3. One hot encoding

This operation transforms categorical or sequence data to a numeric representation using one-hot encoding to enable it to be used in model training. The categories are represented in the form of a distinct binary vector, where no confusion can occur among various features.

3.4. Feature Scaling (Min-Max Method)

The dataset was normalized using min–max scaling technique to restrict values within range of 0 to 1. This method was used to improve the performance of the classifier and lessen the effect of outliers. The normalization process follows the mathematical expression given in Equation (1):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where, X stands for feature's initial value, $X'$is Normalized Value, X_{min} is feature's lowest possible value, and X_{max} is highest value of the same.

3.5. Data Splitting

The data is stratified by splitting it into training and testing datasets (70:30). This method guarantees that the class proportions in both sets match those of the dataset and maintains data consistency during model evaluation and training.

3.6. Implemented Model

In this section discuss the RNN and RF modes in discuss below:

3.6.1. Recurrent Neural Network Model

The RNN model can handle sequential data with ease, as its units are designed to retain previous states. This memory attribute was used to calculate present state. RNN units thus take in two inputs: the previous output and the current input. The mechanics underlying RNN units may be expressed as Equation (2).

$$y(t) = f\big(x(t) + y(t-1)\big) \quad (2)$$

Where, $x(t)$ is current input, $y(t)$is current output, and $y(t-1)$ is past output. A function denoted as (x) represents this connection. The nonlinearity introduced by the basic hyperbolic tangent function (x) in RNN is due to its unstable gradients, which is a characteristic of RNN. The application of the rectified linear unit function or any other non-saturating activation function leads to arbitrary big gradients. Fig. 4 shows the model's architecture, consisting of two recurrent layers and a dense layer.

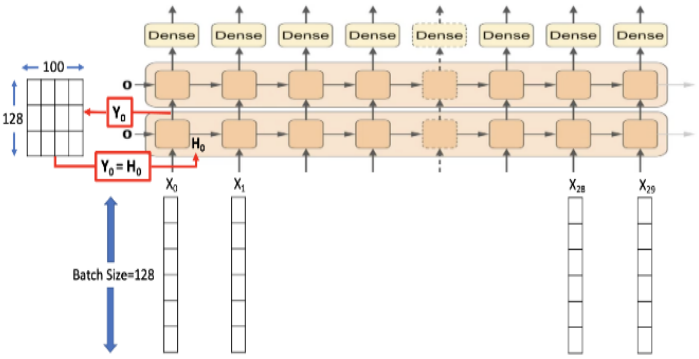


Figure 4. RNN model structure

The RNN model employs a Bidirectional LSTM architecture with 2 layers and 128 hidden units, processing one-hot encoded input sequences of length 23. The model operates with a batch size of 32 and captures temporal dependencies across time steps while maintaining state information between iterations. Binary classification is performed via a dense layer with ReLU and Sigmoid activation; overfitting is prevented by employing a DropoutRate of 0.3. The model is trained for 50 epochs with AdamOptimizer, lossfunction being BinaryCross-Entropy, and LearningRate as 0.001. Xavier initialization and gradient clipping (1.0) are effective strategies to overcome issues such as long training times and batch correlation in stateful RNNs.

3.6.2. Logistic Regression (LR) Model

The proposed LR model is a supervised ML method for binary classification problems. It predicts the likelihood that a sample belongs to a certain class by applying a sigmoid function to a linear combination of input data. The model is simple and efficient, and it yields interpretable results based on the learned coefficients.

$$P(y=1|x)\frac{1}{1+e^{-\left(\beta_0+\beta_1 x_1+\beta_2 x_{2\pm---\mp}\beta_n\beta_n\right)}} \quad (3)$$

In this Equation (3), P(y=1|x) represents the predicted probability, where $x_1, x_2, \ldots\ldots\ldots x_n$ are input features and $\beta_0, \beta_1, \ldots\ldots\ldots \beta_n$ are model parameters. These parameters are tuned during training to reduce the number of misclassifications and enhance prediction accuracy.

$$\hat{y} = I(P(y=1|x) \geq 0.5) \quad (4)$$

This Equation (4) requires an indicator function $I(\cdot)$, which is equal to 1 when the condition is satisfied, and 0 when it is not. Represents the decision rule of Logistic Regression in a more compact and mathematically elegant way.

The Logistic Regression (LR) model employs L2 regularization with a strength (C) of 1.0, and uses the solver 'liblinear' for efficient binary classification. It was trained for a maximum of 100 iterations with a tolerance of 0.0001, and class weights were set to 'balanced' to deal with data imbalance. The model normalizes the input features and outputs predictions with the sigmoid function, using a threshold value of 0.5.

4. Results and Discussion

The cloud GPU that came with Google Colab was used for all of the testing. A 2.3 GHz CPU, 12 GB GPU, 12.72 GB RAM, and 107 GB drive made up the hardware setup used for the investigation. The techniques and model used in this study were implemented using Python 3.6, TensorFlow 2.3, and sklearn.

4.1. Evaluation Metrics

The suggested design was evaluated using many metrics. Accuracy is defined as the ratio of correct predictions made by the trained model to the total number of cases in the dataset. The accuracy of a model may be evaluated by dividing the number of properly projected positive cases by the ratio of correctly expected positive cases. Recall ratio of those events that had been correctly predicted as positive to all those that should have proved positive. A balanced metric that considers both recall and precision is the F1 score, which provides information about the model's accuracy and completeness. In addition, the model's performance was evaluated using ROC and precision-recall (PR) curves. The model's performance was evaluated using the area under the PR curve (PRAUC) and the area under the AUROC. For off-target prediction and other situations where the labels are significantly skewed, AUROC is a better metric for evaluating predictive performance. These metrics are derived in Equations (5-8):

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \quad (5)$$

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

$$Recall = \frac{\text{TP}}{TP+FN} \quad (7)$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (8)$$

In this context, TP is the sum of all correctly predicted off-targets, FP is the number of non-off-targets that were assumed to be off-targets, FN is number of off-targets that were not off-targets, and TN is number of non-off-targets that were correctly anticipated.

4.2. Results

This section summarizes the experiment design and the performance of the suggested models at training and testing stages, where it shows effective processing and reliable results. Table II shows the performance of CRISPR-Cas9 off-target prediction classification performance of RNN and LR models. RNN model has a slightly higher accuracy (97.32) compared to the LR model (97) with an equal precision of 98. Besides, the RNN model has a higher recall (97.2), as well as F1-score (97.9), which implies more balanced and effective predictive performance.

Table 2. Performance Results of Proposed Models using CRISPR/Cas9 Off-Target Dataset

Matrix	Recurrent Neural Network (RNN)	Logistic Regression (LR)
Accuracy	97.32	97
Precision	98	98
Recall	97.2	96
F1-score	97.9	97

In Fig. 5 illustrates both the Validation Accuracy and Validation Loss of proposed RNN model across training epochs. During the initial epochs, the validation accuracy exhibits a rapid increase, rapidly approaching near-perfect performance. Subsequently, it stabilizes with very minor fluctuations, consistent with strong convergence. Simultaneously, the validation loss decreases sharply in the early stages and gradually levels off at a very low value. This trend demonstrates effective learning, stable model performance, and minimal overfitting throughout the training process.

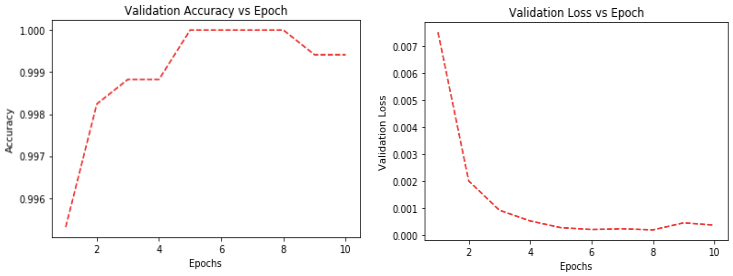


Figure 5. Accuracy and Loss curve for proposed RNN Model

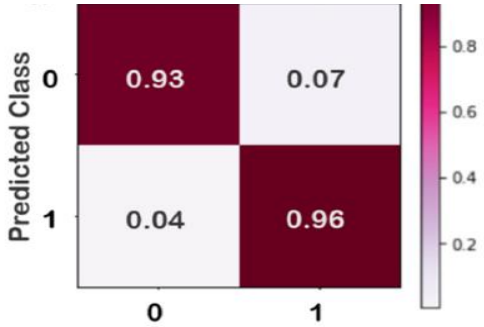


Figure 6. Confusion Matrix for the LR Model

Fig. 6 shows high predictive accuracy in both classes, and the true positive rate of Class 1 is 0.96, whereas the true negative rate is 0.93. The overwhelming diagonal values are properly classified, whereas the off-diagonal values contain a few errors, with 0.07 FP and 0.04 FN. In general, the model performs well, with a good balance of sensitivity and specificity and extremely low misclassification rates.

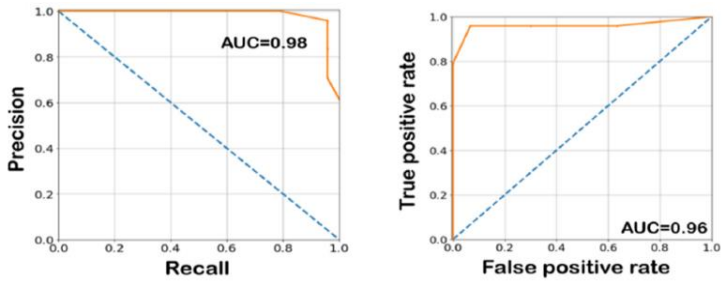


Figure 7. PR and ROC Curve for LR model

Fig. 7 shows LR model's (classification) performance on ROC and Precision-Recall curves. The PR curve has an AUC of 0.98, which indicates that it has a very good tradeoff between precision and recall, and the false positives are small. Similarly, the ROC curve has an AUC of 0.96, indicating strong discrimination between classes and much better performance than a random classifier.

5. Comparative Analysis

A comparative analysis of proposed RNN and LR models against current models has been developed, as shown in Table III, to assess their performance. As pointed out in the table, classical algorithms like ANN, and Naive Bayes have a less accurate rate and varying precision, whereas SVM demonstrates comparatively balanced results. Conversely, the proposed LR model achieves much higher accuracy (97) and high precision (98%). In addition, the suggested RNN model outperforms all other procedures, achieving the highest accuracy (97.32%), recall (97.2%), and F1-score (97.9%), and exhibiting better predictive performance.

Table 3. Comparison of Different Machine Learning Models for CRISPR-Cas9 Off-Target Activities Prediction

Model	Accuracy	Precision	Recall	F1-score
ANN[19]	50.7	50.7	97.2	66.7
NB[20]	83	0.25	76.7	0.49
SVM[21]	87.2	88.3	88.1	88.2
Proposed LR	97	98	96	97
Proposed RNN	97.32	98	97.2	97.9

6. Discussion

The discussion shows that the proposed models are efficient at accurately predicting CRISPR-Cas9 off-target activity. The two models perform well, but the RNN model achieves slightly higher results because it can capture sequential dependencies in biological data. On the whole, the results prove the effectiveness and strength of the offered solution to enhance the accuracy of predictions.

7. Conclusion and Future Study

The off-target effects of CRISPR-Cas9 pose a serious obstacle to its widespread application. It is crucial to overcome this difficulty by developing off-target prediction systems that are both accurate and computationally efficient. The widely used genome-editing technique CRISPR/Cas9 has limited therapeutic utility due to off-target effects, particularly when used in regions with insertions and deletions (indels) and in unbalanced datasets. Despite numerous deep learning-based methods being proposed, some do not address these issues effectively, reducing prediction precision. The side-by-side comparison indicates that the traditional models, such as SVM, have 87.2% of accuracy, while ANN and Naive Bayes have lower performance and imbalanced accuracy and recall. The proposed Logistic Regression (LR) and Recurrent Neural Network (RNN) models are thus, very effective in comparison to these methods in all measures of evaluation. As it is important to mark, the RNN model has the highest acc of 97.32% and the F1score of 97.9% which proves the capability of the model to learn multifaceted sequences of data. Further research can be done in the future to improve model generalization and deal with bigger and more varied genomic data.

References

[1] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang, "Genome engineering using the CRISPR-Cas9 system," *Nat. Protoc.*, vol. 8, no. 11, pp. 2281–2308, Nov. 2013, doi: 10.1038/nprot.2013.143.

[2] Y. Mei, Y. Wang, H. Chen, Z. S. Sun, and X.-D. Ju, "Recent Progress in CRISPR/Cas9 Technology," *J. Genet. Genomics*, vol. 43, no. 2, pp. 63–75, Feb. 2016, doi: 10.1016/j.jgg.2016.01.001.

[3] T. Zhan, N. Rindtorff, J. Betge, M. P. Ebert, and M. Boutros, "CRISPR/Cas9 for cancer research and therapy," *Semin. Cancer Biol.*, vol. 55, pp. 106–119, Apr. 2019, doi: 10.1016/j.semcancer.2018.04.001.

[4] F. Jiang and J. A. Doudna, "CRISPR–Cas9 Structures and Mechanisms," *Annu. Rev. Biophys.*, vol. 46, no. 1, pp. 505–529, May 2017, doi: 10.1146/annurev-biophys-062215-010822.

[5] H. Wang, M. La Russa, and L. S. Qi, "CRISPR/Cas9 in Genome Editing and Beyond," *Annu. Rev. Biochem.*, vol. 85, no. 1, pp. 227–264, Jun. 2016, doi: 10.1146/annurev-biochem-060815-014607.

[6] S.-C. Liu *et al.*, "Target residence of Cas9-sgRNA influences DNA double-strand break repair pathway choices in CRISPR/Cas9 genome editing," *Genome Biol.*, vol. 23, no. 1, p. 165, Dec. 2022, doi: 10.1186/s13059-022-02736-5.

[7] A. Samach *et al.*, "CRISPR/Cas9-induced DNA breaks trigger crossover, chromosomal loss, and chromothripsis-like rearrangements," *Plant Cell*,

vol. 35, no. 11, pp. 3957–3972, Oct. 2023, doi: 10.1093/plcell/koad209.
[8] G. Zhang, Y. Luo, X. Dai, and Z. Dai, "Benchmarking deep learning methods for predicting CRISPR/Cas9 sgRNA on- and off-target activities," *Brief. Bioinform.*, vol. 24, no. 6, Sep. 2023, doi: 10.1093/bib/bbad333.
[9] F. Özden and P. Minary, "Learning to quantify uncertainty in off-target activity for CRISPR guide RNAs," *Nucleic Acids Res.*, vol. 52, no. 18, pp. e87–e87, Oct. 2024, doi: 10.1093/nar/gkae759.
[10] P. D. Hsu *et al.*, "DNA targeting specificity of RNA-guided Cas9 nucleases," *Nat. Biotechnol.*, vol. 31, no. 9, pp. 827–832, Sep. 2013, doi: 10.1038/nbt.2647.
[11] M. Stemmer, T. Thumberger, M. del Sol Keyer, J. Wittbrodt, and J. L. Mateo, "CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool," *PLoS One*, vol. 10, no. 4, p. e0124633, Apr. 2015, doi: 10.1371/journal.pone.0124633.
[12] G. Zhang, Y. Luo, H. Xie, and Z. Dai, "Crispr-SGRU: Prediction of CRISPR/Cas9 Off-Target Activities with Mismatches and Indels Using Stacked BiGRU," *Int. J. Mol. Sci.*, vol. 25, no. 20, p. 10945, Oct. 2024, doi: 10.3390/ijms252010945.
[13] Y. Yang, J. Li, Q. Zou, Y. Ruan, and H. Feng, "Prediction of CRISPR-Cas9 off-target activities with mismatches and indels based on hybrid neural network," *Comput. Struct. Biotechnol. J.*, vol. 21, pp. 5039–5048, 2023, doi: 10.1016/j.csbj.2023.10.018.
[14] A. Imani, J. Valiant, and A. A. Santoso Gunawan, "Deep Learning-based Approach on sgRNA off-target Prediction in CRISPR/Cas9," in *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, IEEE, Feb. 2023, pp. 440–444. doi: 10.1109/ICCoSITE57641.2023.10127682.
[15] K. Sanyour and W. Kim, "Machine Learning Algorithms in Gene Editing," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Dec. 2022, pp. 799–802. doi: 10.1109/BIBM55620.2022.9995135.
[16] W. Xiang, D. Chen, Y. Cui, and S. Peng, "H-VAE: A Hybrid Variational AutoEncoder with Data Augmentation in Predicting CRISPR/Cas9 Off-target," in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Dec. 2021, pp. 550–555. doi: 10.1109/BIBM52615.2021.9669570.
[17] R. Niu, J. Peng, Z. Zhang, and X. Shang, "R-CRISPR: A Deep Learning Network to Predict Off-Target Activities with Mismatch, Insertion and Deletion in CRISPR-Cas9 System," *Genes (Basel).*, vol. 12, no. 12, p. 1878, Nov. 2021, doi: 10.3390/genes12121878.
[18] Q. Liu, X. Cheng, G. Liu, B. Li, and X. Liu, "Deep learning improves the ability of sgRNA off-target propensity prediction," *BMC Bioinformatics*, vol. 21, no. 1, p. 51, Dec. 2020, doi: 10.1186/s12859-020-3395-z.
[19] A. Bhardwaj, P. Tomar, and V. Nain, "Machine Learning-Driven Prediction of CRISPR-Cas9 Off-Target Effects and Mechanistic Insights," *Eurobiotech J.*, vol. 8, no. 4, pp. 213–229, 2024, doi: 10.2478/ebtj-2024-0020.
[20] J. Charlier, R. Nadon, and V. Makarenkov, "Accurate deep learning off-target prediction with novel sgRNA-DNA sequence encoding in CRISPR-Cas9 gene editing," *Bioinformatics*, vol. 37, no. 16, pp. 2299–2307, 2021, doi: 10.1093/bioinformatics/btab112.
[21] J. Das, S. Kumar, D. C. Mishra, K. K. Chaturvedi, R. K. Paul, and A. Kairi, "Machine learning in the estimation of CRISPR-Cas9 cleavage sites for plant system," *Front. Genet.*, vol. 13, no. January, pp. 1–13, Jan. 2023, doi: 10.3389/fgene.2022.1085332.