***ABSTRACT:***

***Modern cyber security relies heavily on intrusion detection in network traffic to quickly identify and mitigate security threats. Conventional intrusion detection systems frequently depend on lone, stand-alone models that could find it difficult to change with the ways that network attacks are evolving. In order to overcome this difficulty, we provide an ensemble-based strategy that improves threat detection effectiveness and accuracy by utilizing the strength of several intrusion detection models. Increasing security measures against data breaches and network invasions is essential due to the ever-growing usage of the Internet and networks. As intrusions are frequently hidden within of valid network packets, firewalls have a difficult time identifying and stopping them. Furthermore, the majority of network monitoring systems and algorithms find it increasingly difficult to handle the sheer volume of network traffic. Various intrusion detection strategies have been proposed in response to these issues, with machine learning techniques emerging as a possible route for handling these situations. This paper introduces an Intrusion Detection System (IDS) that leverages stacking ensemble learning. The core ensemble comprises three fundamental machine learning models: k-nearest-neighbours, Decision Tree, and Random Forest. To enhance classification performance, the proposed system combines a total of seven machine learning algorithms with pre-processing methods. Stacking ensembles improve system performance greatly by combining the results of these basic models with a meta-model represented by the Logistic Regression algorithm. The UNSW-NB15 dataset is used to assess the IDS's efficacy. With an astounding 96.16% accuracy rate in the training phase and an even greater 97.95% accuracy rate in the testing phase, the suggested IDS performs remarkably well. Additionally, the precision scores are remarkable, with training scores of 97.78% and testing scores of 98.40%. These outcomes highlight the system's capacity to detect and stop network intrusions, showing considerable gains in a number of assessment parameters.***

***Keywords: Decision tree; k-nearest-neighbour, random forest; Intrusion detection system (IDS), stacking***

**INTRODUCTION**

The Internet's explosive growth and the widespread use of networks in the modern digital age have completely changed how we interact, collaborate, and transact business. Unprecedented levels of ease and connectedness have been brought about by this technological revolution, but it has also created a formidable obstacle in the form of the growing threat of data breaches and network attacks.

These intrusions carried out via network packets; pose a continuous and changing threat to the confidentiality and safety of our digital infrastructure. To combat these threats, one of the primary challenges lies in the deceptive similarity between regular network traffic and intrusion attempts. Intruders often exploit a high volume of typical data to camouflage their malicious activities, rendering it exceedingly challenging for conventional security measures, including firewalls, to differentiate between benign and malicious actors.

Furthermore, classic network monitoring systems and rule-based algorithms face major problems from the sheer volume and complexity of network traffic in today's interconnected world. It's getting harder to find anomalies and possible dangers in this deluge of data, which calls for more advanced and flexible solutions.

The discipline of intrusion detection has seen a growth of creative ideas and methodology in response to this urgent requirement. Among these, machine learning approaches have attracted a lot of interest due to their potential to improve intrusion detection systems' capabilities. Machine learning has potential in enhancing the precision and effectiveness of detecting and addressing security events by utilizing algorithms and data-driven analyses.

Network security has become critical in an era where communication technologies and the Internet are developing quickly. Businesses all over the world are forced to make large investments in protecting their critical operations and sensitive data. They use a variety of security measures, like as intrusion detection systems (IDS), firewalls, and antivirus software, to accomplish this, all with the goal of protecting the integrity and confidentiality of their networks and related assets. Among these, intrusion detection systems (IDS) are essential because they actively detect anomalous network activity and rapidly alert network administrators to it [1].

In the last ten years, network security has seen a significant rise in the use of machine learning techniques. These methods have become quite popular because of their exceptional capacity to extract relevant features from network data and then apply learnt patterns to produce precise classifications [2]. Deep Learning (DL)-based intrusion detection systems (IDS) are particularly noteworthy because of their deep neural network topologies, which enable them to independently extract complex features from unprocessed data and provide a powerful defense against new threats [3]. Ensemble learning stands as another potent machine learning paradigm. It involves the construction and thoughtful integration of multiple models, which can encompass classifiers or domain experts, to address specific computational intelligence challenges. Ensemble learning is used in many different fields, and its major goals are to reduce the risk resulting from inadvertently choosing a weak model and to improve model performance (e.g., classification, prediction, function approximation).

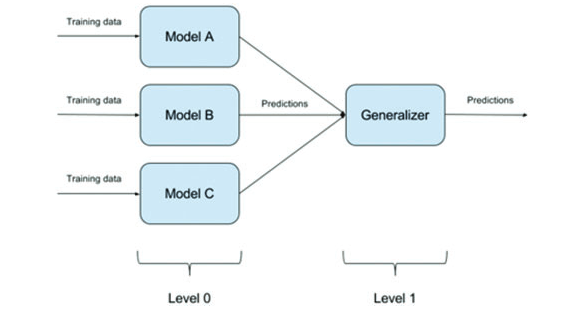


Fig1: Ensemble classifiers basic architecture

In this study, a novel intrusion detection system (IDS) is presented that utilizes a wide range of individual experts to thoroughly examine requests and responses in network traffic. The primary objective of this system is to monitor network activity from its inception within the network and track it until it exhibits signs of potential harm or engages in prohibited actions. This approach to achieving classifier diversity is distinctive: it involves the use of distinct training parameters for each classifier. Through this technology, classifiers can define a wide spectrum of decision boundaries. By amalgamating the outputs of these diverse classifiers, the system minimizes overall errors and significantly enhances accuracy when compared to other deep learning (DL) and artificial intelligence (AI) systems.

The main goal of this research is to create an intrusion detection system that is both efficient and effective in identifying security breaches in the intricate world of network traffic. The following succinctly describes the research objectives:

1. Performance Assessment of Machine Learning Algorithms: The study aims to evaluate how well different machine learning algorithms perform in terms of intrusion detection. In order to determine these algorithms' advantages and disadvantages in relation to network security, the study will compare and contrast them.

2. Leveraging Stacking Ensemble Learning: The research makes use of the potent stacking ensemble learning technique, building upon the foundation of many machine learning algorithms. This method is used to increase the IDS's overall robustness and performance. The system attempts to gain increased accuracy and adaptability in spotting security threats by merging the results of numerous distinct models.

3. Assessment Employing UNSW-NB15 Dataset: With the UNSW-NB15 dataset, the suggested IDS is thoroughly assessed to confirm its efficacy. This well-known benchmark dataset provides a wide range of incursion types and network traffic scenarios. The assessment seeks to demonstrate the system's capacity to identify and address security flaws in actual network traffic.

4. Evaluation through Comparison: Comparing the suggested intrusion detection system's performance against that of current intrusion detection systems is a crucial component of the research. The purpose of this comparison analysis is to demonstrate the improvements and benefits that the new approach has over traditional IDS solutions. We shall explore the details of the suggested IDS in the following sections of this study, explaining the approaches used for classifier diversity and stacking ensemble learning. The evaluation's empirical findings, which were obtained using the UNSW-NB15 dataset, will be discussed and will provide insight into the system's improved capacity to identify security breaches in intricate network traffic situations. In the end, our research advances the continuous search for network security solutions that are resistant to changing threats and efficient.

**2. Related Work**

Jim Anderson first proposed the idea of intrusion detection systems (IDS) in 1980 [4], which signalled the start of the path toward improved network security. Since then, a number of intrusion detection systems (IDS) have been developed and matured to meet the changing requirements of network security, contributing to the substantial expansion of the intrusion detection sector [5]. Combining the words "intrusion" and "detection systems," an IDS is an acronym that is essential to the protection of network environments and computer systems. To put it simply, a "intrusion" is any unapproved access to computer systems or internal network data that is intended to jeopardize its availability, confidentiality, or integrity [6]. Being a watchful defender against illegal activity is the detection system's job within an intrusion detection system (IDS). Consequently, an IDS operates as a vigilant security tool, consistently monitoring both host and network activities. Its primary objective is to detect any signs of suspicious behaviour that deviates from established security standards and threatens the network's availability, confidentiality, and integrity. Typically, AN ID is connected to a network adapter configured using port mirroring technology to enable it to effectively carry out its function, as illustrated in Figure 1. The intrusion detection system (IDS) promptly alerts the network administrator upon detecting malicious activities.

In IDS, Deep Learning and Machine Learning Researchers and practitioners have investigated the integration of machine learning (ML) and deep learning (DL) technologies within intrusion detection systems (IDS) in response to the always changing threat landscape. To extract valuable insights from large datasets, machine learning and deep learning are both essential. Network security has embraced these technologies more quickly in the last ten years due to the widespread use of powerful Graphics Processing Units (GPUs) [7, 8]. When it comes to analyzing and predicting network traffic, ML and DL approaches excel. Interestingly, although DL-based IDS uses its deep neural network architecture to automatically learn complex features directly from raw data, ML-based IDS frequently depends on engineered features to extract meaningful information from network traffic data [2][3].

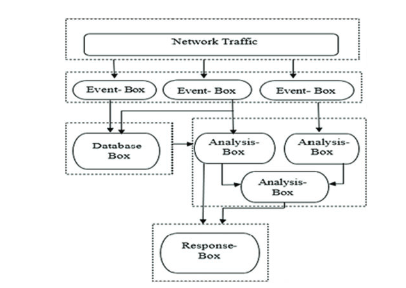
As shown in Fig. 2, the architecture of IDS is usually made up of four main functional modules: Event-Boxes, Database-Boxes, Analysis-Boxes, and Response-Boxes. As sensors, Event-Boxes keep an eye on the system all the time, gathering pertinent data for further examination. The Database-Boxes serve as the repository for the gathered data, guaranteeing that the data is accessible for processing. The fundamental processing module is represented by the Analysis-Boxes, wherein dangerous conduct is detected through a close examination of the gathered events. The most important stage is reacting quickly to any malicious activity that is identified; the Response-Boxes are responsible for this. This visual representation is depicted in Figure 2, serving as an illustration of the IDS framework.

Fig 2: Architecture for IDS

IDSs that are Host-Based (HIDS) and Network-Based (NIDS). Event-Boxes in IDS can be categorized into two primary types based on their data sources: Host-Based IDS (HIDS) and Network-Based IDS (NIDS). HIDS primarily concentrates on monitoring system calls and process identifiers, while also assessing the behaviour of the host system. In contrast, NIDS is oriented towards network events, examining factors such as traffic volume, protocols, IP addresses, and service ports [1].

**2.1 IDS based on signatures and Anomaly-Based IDS**

The vital function of intrusion detection is carried out by the Analysis-Boxes in an intrusion detection system (IDS), which are divided into two primary types: Anomaly-Based IDS and Signature-Based IDS (SIDS). In order to detect intrusions, SIDS compares collected data to a database of known attack signatures, identifying patterns that correspond to known attack characteristics [9]. This method has limits in detecting new or previously unknown attacks, but it is quite effective in identifying established threats. Nonetheless, a noteworthy benefit is its notably lower false-positive rate.

**2.2 IDS Based on Anomalies (AIDS)**

The methodology used by Anomaly-Based IDS (AIDS) differs from that of Signature-Based IDS (SIDS). Its main goal is to determine a baseline threshold and comprehend the normal behaviour of a system. An anomaly alarm is set off when an observation departs from this consistent trend [10]. As AIDS focuses on finding unexpected occurrences, it is especially useful for identifying new or previously undetected attacks.

**2.3 Stacked Group Education in IDS**

Researchers have investigated the merging of different machine learning methods through Stacked Ensemble Learning as a means of efficiently detecting network-based assaults. For example, a stacked ensemble method was assessed and found to be more successful in classifying network-based attacks than other popular machine learning algorithms like Support Vector Machines (SVM), Random Forest, Classification and Regression Trees (CART), and Artificial Neural Networks (ANN).

In a similar spirit, researchers used the stacking classifier method to apply multiple learning algorithms to the UNSW-NB15 dataset. Additionally, they used a mixed strategy for feature selection that combined SVM and Lasso regression. The outcomes demonstrated that Lasso regression obtained an R2 score assessment of 59%. Furthermore, Gao et al. experimented with various training ratios and generated many decision trees while examining the NSL-KDD dataset. The creation of the Multi Tree detection technique, which combines adaptive voting across several classifier algorithms, marked the end of their investigation. The strategy was to raise the detection accuracy from 84.1% to 85.2%.

The authors of a different study used a variety of classifiers, such as Random Forests, Decision Trees, K-Nearest Neighbours (KNN), and Deep Neural Networks (DNN). The value of ensemble learning in intrusion detection was demonstrated by the improved detection accuracy that resulted from the adoption of an adaptive voting system.

**3.Data Acquisition**

To assess and appraise our proposed approach, we employed the UNSW-NB15 dataset, obtained from ACCS, which is acknowledged as a modern benchmark dataset for Network Intrusion Detection Systems (NIDS). This dataset comprises a substantial 2.5 million records, encompassing 45 distinctive features. To streamline our analysis, we performed some modifications on the original UNSW-NB15 dataset, reducing the feature set from 45 to 43. These features encompass both flow-based and packet-based attributes.

The features within the dataset can be methodically categorized into four clear groups: content, fundamental, flow, and time-based characteristics. These subsets of features offer valuable insights into the patterns of network traffic and behaviour.

Regarding the data volume, we carefully selected a subset of 257,673 data instances from the UNSW-NB15 dataset for our analysis. This subset was then further partitioned into two essential segments: training data instances, consisting of 175,341 records, and testing data instances, which include 82,332 records.

This streamlined dataset configuration allowed us to efficiently train and test our Intrusion Detection System (IDS) while preserving the most pertinent characteristics of the UNSW-NB15 dataset.

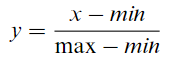
**4. Outline of the Suggested System for Ensemble Learning**

Ensemble learning is a prominent machine learning approach that combines the advantages of individual classifiers. This fusion leads to the development of a unified classification model that demonstrates notably improved performance in both classification and prediction tasks. When compared to single, standalone classifiers, ensemble learning has demonstrated consistently better performance in the particular setting of intrusion detection systems (IDS). Developing a strong network intrusion detection system by utilizing ensemble classifiers' enormous potential is the main goal of this research.

To apply our suggested technique, datasets that are mainly intended for training must be carefully pre-processed. As the basis for our work, we have employed the UNSW-NB15 dataset, which is a large collection of 45 unique attributes. These features consist of four nominal qualities and forty-one numerical attributes, the latter of which were converted to numerical values in order to aid in the process of analysis and model building.

In our pre-processing workflow, we have identified data normalization as a critical step. This necessity arises from the inherent variability in dimensions and units utilized during data collection. To ensure uniformity and equitable impact on the learning process, it becomes essential to standardize and align disparate feature values to a consistent scale. In this research, we have opted for the Max-Min normalization algorithm as the preferred method to accomplish this pivotal data adjustment.

Each feature is separately scaled and translated as part of the Max-Min normalization process to fit into a preset range, usually ranging from zero to one. This change helps improve our ensemble learning model's convergence and overall performance while also lessening the effect of feature magnitude variations.-

---(1)

**5. Architecture of the Proposed Approach**

This study's primary goal is to identify network intrusions with accuracy. In order to accomplish this, our suggested method is divided into two basic levels, each of which is essential to the intrusion detection procedure.

Layer of Base-Learner: The Base-Learner layer, which is the initial layer, forms the basis of our intrusion detection system. We carefully choose three base binary classifiers in this layer, each of which is intended to detect intrusion patterns:

1. Random Forest Classifier of handling complicated issues, this classifier makes use of an ensemble of decision trees. Prediction accuracy is increased by using numerous decision trees on various dataset subsets and averaging their outcomes.

2. Decision Tree Classifier algorithm is highly effective at extracting useful data from large datasets. It makes value predictions using a test dataset to gauge its accuracy and a training dataset to build a decision tree.

3. K-Nearest Neighbours (K-NN) Classifier: Based on the available data, K-NN determines the probability that a given data point belongs to the closest group.

The Combining-Module layer receives the combined output of these base classifiers, which enhances the system's overall intrusion detection capabilities.

Layer of Combining Modules: As an aggregator, the Combining-Module layer balances the output produced by each of the separate basic classifiers. By significantly improving prediction accuracy, this aggregation technique makes the system more robust and trustworthy in detecting network intrusions.

Stacked generalization is harnessed to direct the outcomes from the initial ensemble layer towards a meta-classifier. In this context, the meta-classifier employs regression to forecast the probability of a specific class or event, be it pass/fail or win/lose.

In the domain of ensemble learning, three fundamental techniques are typically classified: bagging, boosting, and stacking. Of these, bagging stands as the most widespread method for forecasting test results. It entails extensive training of models using misclassified data to boost overall performance.

In our research, stacking, also known as stacked generalization, takes precedence as the preferred ensemble technique. Stacking is primarily utilized to elevate classification performance by amalgamating insights from multiple classifiers, as illustrated in Figure 3.

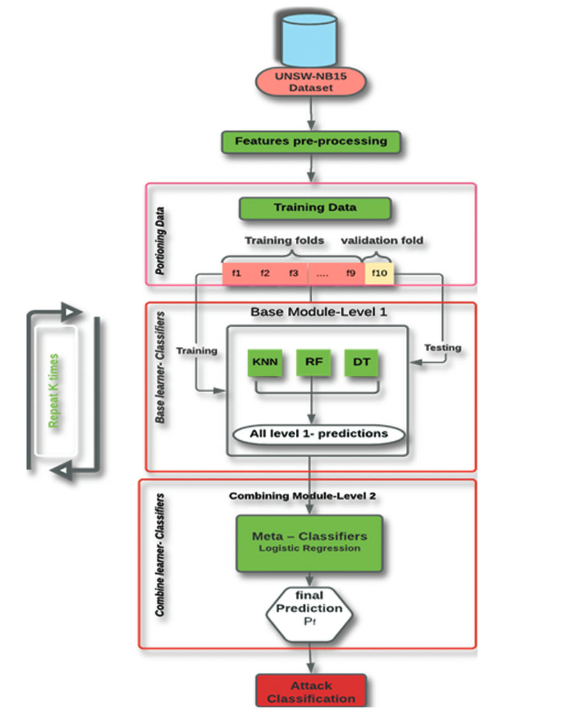
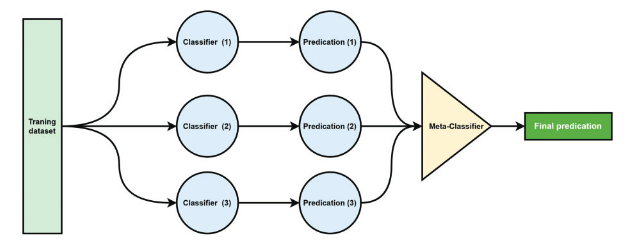


Fig3: IDS based on stacking

Stacking Ensemble Learning represents a departure from the traditional bagging and boosting methods, as it introduces a two-tiered approach comprising level 0 for base learners and level 1 for the meta-learner. In this distinctive setup, diverse classification models are trained on the training dataset at the initial level. The outputs generated by these base models contribute to the creation of a fresh dataset, which serves as input for the stacking learner.Fig 4: Stacking Classifiers

In summary, our proposed approach combines the strengths of multiple classifiers through a sophisticated ensemble learning framework. This two-layered architecture enhances the accuracy and reliability of intrusion detection by leveraging the diverse capabilities of the selected base classifiers and optimizing their outputs through the stack generalization process. Ultimately, this approach aims to provide a robust and effective solution for the detection of network intrusions.

**6. Experimental Results and Analysis**

We offer a thorough presentation and analysis of the experimental results from our suggested Intrusion Detection System (IDS), which is based on machine learning techniques and approaches, in this part. Using the UNSW-NB15 dataset, a well-known benchmark for network intrusion detection, we deployed seven different algorithms to the suggested IDS system in order to thoroughly evaluate its efficacy. A range of assessment criteria and performance metrics were used to evaluate the operation of our IDS system. These measurements are crucial reference points for determining how well the system detects network breaches

We employed the following essential evaluation and performance metrics to assess the effectiveness of our proposed Intrusion Detection System (IDS):

1. **Accuracy:** Accuracy is the ratio of correctly classified instances to total instances, which indicates the overall correctness of the IDS. Better accuracy is indicated by higher numbers, which range from 0 to 1.

(2)

1. **Precision:** Precision assesses the system's accuracy in detecting intrusions among the occurrences it classifies as positive. It represents the percentage of real positives out of all cases that are classified as positive.

(3)

1. **Recall:** Also known as sensitivity or true positive rate, recall evaluates the ability of the intrusion detection system to find all pertinent incursions. It measures the percentage of real positives among all real positive cases that have been found.

(4)

1. **AUC (Area Under the Receiver Operating Characteristic Curve):**The metric known as AUC, or Area Under the Receiver Operating Characteristic Curve, evaluates the overall effectiveness of the intrusion detection system by taking into account the trade-off between the true positive rate and the false positive rate. It offers a thorough evaluation of the discriminating power of the system.

(5)

1. **F1-Measure:**The harmonic mean of recall and precision is the F1-Measure. It provides a fair assessment of the IDS's effectiveness, especially in the case of unbalanced datasets.

(6)

1. **Mean Squared Error (MSE):**The Mean Squared Error (MSE) is a metric that assesses the degree of accuracy in the IDS's predictions by measuring the average squared difference between the actual and anticipated values.

(7)

By employing these diverse assessment criteria, we aim to offer a thorough and well-rounded evaluation of the proposed IDS approach. These metrics collectively shed light on the system's accuracy, precision, recall, discrimination power, and ability to handle imbalanced datasets.

**7. Performance Evaluation of ML Algorithms**

This evaluation's main goal was to determine how accurate our suggested approach was, using the UNSW-NB15 dataset in particular and utilizing the entire feature space with 42 characteristics for binary classification.

We used a 10-fold cross-validation approach; more precisely stratified 10-fold cross-validation, to guarantee a strong evaluation. The dataset is split into ten subsets for this process; nine of these subsets are used to train the classifiers, and the tenth subset is used for testing.

Table 3 presents a summary of the findings of this thorough evaluation that includes important performance measures This information encompasses the metrics such as accuracy, precision, recall, F1-Score, AUC, and MSE values for the recommended machine learning algorithms as applied to the training dataset. The examination yielded some noteworthy results, such as: Random Forest (R.F.) is the best performer with the greatest accuracy score of 96.12%; Linear SVM is the best performer with the highest precision score of 99.79%, while it has a lower recall score of 91.21%.

• Although Naïve Bayes (N.B.) has an impressive 93.41% recall score, its accuracy is marginally lower.

At 96.38%, Decision Tree (D.T.) exhibits the highest recall value. In addition, D.T. has less variance in recall and accuracy than the N.B. model, which raises its F1 score.

• Random Forest (R.F.) performs exceptionally well, with the lowest Mean Squared Error (MSE) of 0.0388 and the highest F1 score of 97.18%.

**8. Ensemble Learning for Enhanced Performance**

The performance of our intrusion detection system received a substantial boost through the adoption of the stacking ensemble technique. This approach involves the fusion of multiple base models (level-0 models) alongside a meta-model (level-1 model) responsible for aggregating predictions generated by these base models.

In our ensemble configuration, we harnessed the potential of Random Forest, Decision Tree, and KNN as the foundation models. Concurrently, Logistic Regression (L.R.) assumed the pivotal role of the meta-model. This ensemble methodology yielded remarkable results, notably achieving an accuracy rate of 96.16%, marking a significant enhancement in our intrusion detection system's performance.

To summarize, our analysis demonstrates the significant potential of the suggested machine learning techniques for identifying network breaches. The results emphasize the accuracy-related advantages of Random Forest (R.F.) and the drawbacks and advantages of alternative classifiers. Additionally, our intrusion detection system's accuracy is further improved by using the stacking ensemble method, highlighting the value of ensemble learning in improving performance.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Method** | Accuracy | Precision | Recall | F1-score | AUC | MSE |
| SVM(RBF) | 93.60 | 99.63 | 91.69 | 95.49 | 95.36 | 0.0640 |
| Random forest(RF) | 96.12 | 97.98 | 96.38 | 97.18 | 95.96 | 0.0388 |
| Logistic regression(LR) | 93.40 | 99.13 | 91.83 | 95.34 | 94.79 | 0.0660 |
| Linear SVM | 93.31 | 99.79 | 91.21 | 95.31 | 95.33 | 0.0669 |
| Naïve Bayes(NB) | 81.38 | 78.16 | 93.41 | 85.10 | 79.44 | 0.1862 |
| Decision tree(DT) | 95.00 | 96.27 | 96.38 | 96.33 | 94.23 | 0.0500 |
| KNN | 93.76 | 95.90 | 94.98 | 95.44 | 93.02 | 0.0624 |
| Stacking ensemble | 96.16 | 97.78 | 96.62 | 97.20 | 95.88 | 0.0384 |

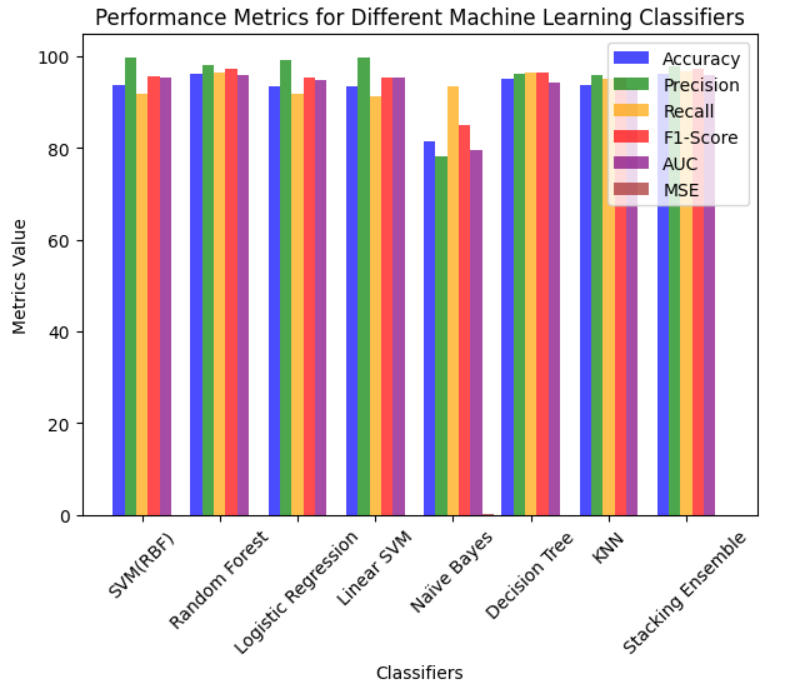


Fig 5: Comparison of ML classifiers on the trained dataset

Table 4 presents the test dataset results, showcasing the performance metrics of the recommended ML algorithms, including In terms of accuracy, precision, recall, F1-Score, AUC, and MSE, Random Forest (R.F.) clearly distinguishes itself with outstanding scores, attaining a 98.52% accuracy and an impressive 97.72% precision. 98.12% recall, 97.96% F1-score, and an impressive 98.54%

In the AUC rankings, Decision Trees (D.T.) closely follow as the runner-up, demonstrating impressive performance with 96.74% accuracy, 96.98% precision, 97.11% recall, and a robust F1-score of 96.70%.

Taking the third position, K-Nearest Neighbours (KNN) exhibits noteworthy figures, including 93.33% accuracy, 95.56% precision, 92.17% recall, and a commendable F1-score of 93.46%.

In terms of Mean Squared Error (MSE), which measures prediction accuracy, Random Forest (R.F.) demonstrates the best performance with an MSE of 0.0206. Decision Trees (D.T.) follow with an MSE of 0.0326, and K-Nearest Neighbours (KNN) exhibits an MSE of 0.0667.

Additionally, the stacking ensemble approach attains an accuracy of 97.95%, showcasing its effectiveness in combining multiple classifiers to enhance overall performance.

For a more visual representation of the performance comparison among the ML classifiers on the test dataset, please refer to Fig. 6.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Method | Accuracy | Precision | Recall | F1-score | AUC | MSE |
| SVM(RBF) | 93.10 | 92.65 | 95.01 | 93.81 | 92.89 | 0.0690 |
| Random forest(RF) | 97.94 | 98.52 | 97.72 | 98.12 | 97.96 | 0.0206 |
| Logistic regression(LR) | 90.70 | 91.54 | 91.58 | 91.56 | 90.61 | 0.0930 |
| Linear SVM | 91.04 | 92.70 | 90.89 | 91.78 | 91.06 | 0.0896 |
| Naïve Bayes(NB) | 76.43 | 86.28 | 68.02 | 76.07 | 77.38 | 0.2357 |
| Decision tree(DT) | 96.74 | 96.98 | 97.11 | 97.04 | 96.70 | 0.0326 |
| KNN | 93.33 | 95.56 | 92.17 | 93.84 | 93.46 | 0.0667 |
| Stacking ensemble | 97.95 | 98.40 | 97.87 | 98.13 | 97.96 | 0.0205 |

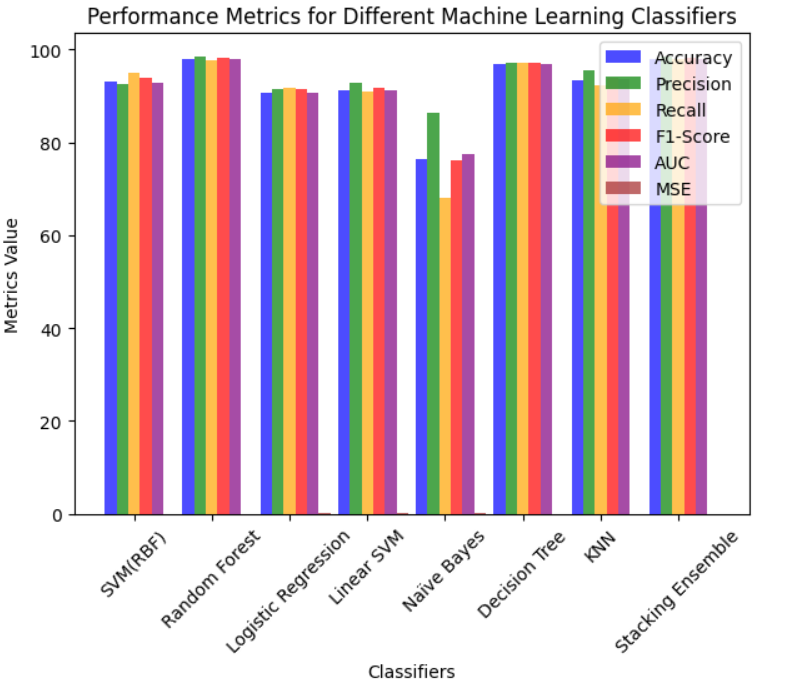


Fig 6: Comparison of ML classifiers on the test dataset

In summary, this study adeptly tackled the crucial challenge of selecting the optimal classifier for precise classification tasks, with a specific emphasis on intrusion detection in network data. Through extensive experimentation and comparative analysis, we evaluated a variety of classifiers, including Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Trees, and Naïve Bayes. The research underscores the substantial influence of ensemble learning approaches, demonstrating their capacity to mitigate the risks associated with suboptimal classifier choices when compared to the exclusive reliance on a single classifier.

**9. Conclusion and Future Directions**

In summary, this study successfully tackled the issue of classifier selection for precise classification tasks, with a specific focus on intrusion detection within network data. Our thorough analysis encompassed a range of classifiers, including Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Trees, and Naïve Bayes. The main findings from our research underscore the clear benefits of embracing an ensemble approach, where multiple classifiers work together in synergy. This approach substantially minimizes the chances of suboptimal decisions compared to sole reliance on a single classifier.

A significant breakthrough in this research is the utilization of the stacking ensemble method. Through the integration of base models like Random Forest (R.F.), Decision Tree (D.T.), k-Nearest Neighbours (KNN), and the incorporation of a meta-model represented by Logistic Regression (L.R.), we attained a remarkable accuracy rate of 97.95% in the testing phase. This outcome unmistakably underscores the enormous potential of ensemble learning in elevating the performance of Intrusion Detection Systems (IDS). It underscores the power of combining diverse models to enhance accuracy and reliability in the domain of intrusion detection within network security.

**10. References**

[1] R. A. Disha and S. Waheed, “Performance analysis of machine learning models for intrusion detection  
system using gini impurity-based weighted random forest (GIWRF) feature selection technique,” Cyberse-  
curity, vol. 5, no. 1, pp. 1–22, 2022. <https://doi.org/10.1186/s42400-021-00103-8>

[2] Aruna, R., Kushwah, V.S., Praveen, S.P. et al. Coalescing novel QoS routing with fault tolerance for improving QoS parameters in wireless Ad-Hoc network using craft protocol. Wireless Netw (2023). <https://doi.org/10.1007/s11276-023-03515-1>

[3] B. Dong and X. Wang, “Comparison deep learning method to traditional methods using for network  
intrusion detection,” in 8th IEEE Int. Conf. on Communication Software and Networks (ICCSN), Beijing,  
China, pp. 581–585, 2016.

[4] H. Debar, M. Dacier and A. Wespi, “Towards a taxonomy of intrusion-detection systems,” Computer  
Networks, vol. 31, no. 8, pp. 805–822, 1999. <https://doi.org/10.1016/S1389-1286(98)00017-6>

[5] Reddy, N.V.R.S., Chitteti, C., Yesupadam, S., Desanamukula, V.S., Vellela, S.S., Bommagani, N.J. (2023). Enhanced speckle noise reduction in breast cancer ultrasound imagery using a hybrid deep learning model. Ingénierie des Systèmes d’Information, Vol. 28, No. 4, pp. 1063-1071. <https://doi.org/10.18280/isi.280426>

[6] Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07)

[7] Vellela, S.S., Balamanigandan, R. An intelligent sleep-awake energy management system for wireless sensor network. Peer-to-Peer Netw. Appl. (2023). <https://doi.org/10.1007/s12083-023-01558-x>

[8] Vellela, S. S., &Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.

[9] Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2022). A New Multi-Level Semi-Supervised Learning Approach for Network Intrusion Detection System Based on the ‘GOA’. Journal of Interconnection Networks, 2143047.

[10] Vellela, S. S., Reddy, B. V., Chaitanya, K. K., &Rao, M. V. (2023, January). An Integrated Approach to Improve E-Healthcare System using Dynamic Cloud Computing Platform. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 776-782). IEEE.

[11] S Phani Praveen, RajeswariNakka, AnuradhaChokka, VenkataNagarajuThatha, SaiSrinivasVellela and UddagiriSirisha, “A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques” International Journal of Advanced Computer Science and Applications(IJACSA), 14(6), 2023.http://dx.doi.org/10.14569/IJACSA.2023.01406128

[12] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., &Swapna, D. (2022, November). An Adaptive Load Balancing Technique for Multi SDN Controllers. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1403-1409). IEEE.

[13] Vellela, S. S., BashaSk, K., &Yakubreddy, K. (2023). Cloud-hosted concept-hierarchy flex-based infringement checking system. International Advanced Research Journal in Science, Engineering and Technology, 10(3). Vellela, S. S., &Balamanigandan, R. (2023).

[14] Sk, K. B., Roja, D., Priya, S. S., Dalavi, L., Vellela, S. S., & Reddy, V. (2023, March). Coronary Heart Disease Prediction and Classification using Hybrid Machine Learning Algorithms. In 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA) (pp. 1-7). IEEE.

[15] VenkateswaraRao, M., Vellela, S., Reddy, V., Vullam, N., Sk, K. B., &Roja, D. (2023, March). Credit Investigation and Comprehensive Risk Management System based Big Data Analytics in Commercial Banking. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 2387-2391). IEEE.

[16] Vellela, S.S., Balamanigandan, R. Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. Multimed Tools Appl (2023). [https://doi.org/10.1007/s11042- 023- 15926-5](https://doi.org/10.1007/s11042-%20023-%2015926-5)

[17] Vullam, N., Vellela, S. S., Reddy, V., Rao, M. V., SK, K. B., &Roja, D. (2023, May). Multi-Agent Personalized Recommendation System in E-Commerce based on User. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1194-1199). IEEE.

[18] Rao, K. N., Gandhi, B. R., Rao, M. V., Javvadi, S., Vellela, S. S., & Basha, S. K. (2023, June). Prediction and Classification of Alzheimer’s Disease using Machine Learning Techniques in 3D MR Images. In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 85-90). IEEE.

[19] Rao, D. M. V., Vellela, S. S., Sk, K. B., & Dalavai, L. (2023). Stematic Review on Software Application Under-distributed Denial of Service Attacks for Group Website. Dogo Rangsang Research Journal, UGC Care Group I Journal, 13.

[20] Venkateswara Reddy, B., &KhaderBashaSk, R. D. Qos-Aware Video Streaming Based Admission Control And Scheduling For Video Transcoding In Cloud Computing. In International Conference on Automation, Computing and Renewable Systems (ICACRS 2022).

[21] Madhuri, A., Praveen, S. P., Kumar, D. L. S., Sindhura, S., &Vellela, S. S. (2021). Challenges and issues of data analytics in emerging scenarios for big data, cloud and image mining. Annals of the Romanian Society for Cell Biology, 412-423.

[22] Vellela, S. S., Balamanigandan, R., & Praveen, S. P. (2022). Strategic Survey on Security and Privacy Methods of Cloud Computing Environment. Journal of Next Generation Technology (ISSN: 2583- 021X), 2(1).

[23] Venkateswara Reddy, B., Vellela, S. S., Sk, K. B., Roja, D., Yakubreddy, K., & Rao, M. V. Conceptual Hierarchies for Efficient Query Results Navigation. International Journal of All Research Education and Scientific Methods (IJARESM), ISSN, 2455-6211.

[24] Sk, K. B., Vellela, S. S., Yakubreddy, K., & Rao, M. V. (2023). Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation. Khader Basha Sk, Venkateswara Reddy B, Sai Srinivas Vellela, Kancharakunt Yakub Reddy, M Venkateswara Rao, Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation, 10(3).

[25] Sk, K. B., & Vellela, S. S. (2019). Diamond Search by Using Block Matching Algorithm. DIAMOND SEARCH BY USING BLOCK MATCHING ALGORITHM", International Journal of Emerging Technologies and Innovative Research (www. jetir. org), ISSN, 2349-5162.

[26] S. S. Vellela, V. L. Reddy, R. D, G. R. Rao, K. B. Sk and K. K. Kumar, "A Cloud-Based Smart IoT Platform for Personalized Healthcare Data Gathering and Monitoring System," 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIANCON58793.2023.10270407.

[27] S. S. Priya, S. Srinivas Vellela, V. R. B, S. Javvadi, K. B. Sk and R. D, "Design And Implementation of An Integrated IOT Blockchain Framework for Drone Communication," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205659.

[28] N. Vullam, K. Yakubreddy, S. S. Vellela, K. Basha Sk, V. R. B and S. Santhi Priya, "Prediction And Analysis Using A Hybrid Model For Stock Market," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205638.