

# Enhancing Data Transmission Security in Cloud Using Machine Learning

Mrs. K. Suma<sup>1</sup>, Jaligam Kranthi Kumar<sup>2</sup>, Immaneni Anisha<sup>3</sup>, Gurram Pranava Sessa Sai<sup>4</sup>

<sup>1</sup>Assistant Professor, Department of CSM, CMR Institute of Technology, Hyderabad, Telangana India

<sup>2</sup>Student, Department of CSM, CMR Institute of Technology, Hyderabad, Telangana, India

<sup>3</sup>Student, Department of CSM, CMR Institute of Technology, Hyderabad, Telangana, India

<sup>4</sup>Student, Department of CSM, CMR Institute of Technology, Hyderabad, Telangana, India

## ABSTRACT

The popularity of cloud computing is on the rise due to its flexibility and scalability. Nonetheless, since cloud computing involves transfer of data over a network, the security threat posed by cyber criminals targeting the data being transferred over a network is significant. Conventional approaches aimed at ensuring data security through encryption techniques are effective. Unfortunately, they are not dynamic enough to cope with emerging methods used by cyber criminals to exploit vulnerabilities in a network. Besides, the processes involved in encryption consume computational resources leading to performance degradation. This paper introduces a machine learning approach for improving the security of data transfer within a cloud environment. This framework incorporates the utilization of two algorithms including K-Nearest Neighbors (KNN) and Artificial Neural Networks (ANN). Both approaches will be combined in an ANN-KNN architecture. Artificial Bee Colony (ABC) optimization will also be applied. The proposed framework will undergo training using the NSL-KDD dataset. The results obtained from the experiment show that ANN-ABC model yields an accuracy of more than 99%. This is significantly better than other models. The ANN-KNN architecture incorporating ABC optimization will detect any security threat to cloud-based networks in real-time.

**The keywords to consider include:** Cloud Security, Machine Learning, ANN, KNN, ABC Optimization, Intrusion Detection.

## I. INTRODUCTION

Generally, the idea behind network security has been that of trusted and untrusted network. Organizations have traditionally depended on perimeter-based type of security systems to safeguard their trusted network from any kind of attack from untrusted networks. However, with today's technology, this theory might no longer be valid in most cases. With advancements in technology platforms like Cloud Computing, Remote Access, and multiple tenant infrastructures, organizations' internal networks have practically ceased to become totally trusted. Though business firms have benefited greatly from cloud computing, moving user information to the cloud server remains one of the most daunting challenges.

With cloud computing, it has become possible for businesses to manage their application and data usage in terms of using their computing resources over the internet. There are many other functions which can be performed by cloud platforms, such as managing

large volumes of data, multiple access of data by people or processes, real-time services sharing, and processing and distribution work being done by multiple users or processes. Hence, it can be concluded that the cloud platform is very essential for every organization. Machine Learning (ML) has been identified as a powerful solution in this regard, since ML algorithms help to detect any abnormal behavior and potential threats to computer security. Machine Learning models analyze vast amounts of network traffic to recognize abnormal data patterns.

The article describes a novel ML-based technique that uses the KNN, ANN, and ABC algorithms in an integrated way to protect Cloud Data Transmission Security using such technologies. The architecture of the solution will ensure reliability, scalability, and adaptability of the threats' detection with minimum false positive.

The adoption of cloud computing, big data analytics, IoT, and distributed applications has largely altered the design of computing infrastructure. Leveraging the cloud enables users to leverage limitless volumes of computing storage, computing power, and computing software resources e.g. not having to spend millions of dollars in setting up an advanced physical infrastructure of enterprise computers to do this. Due to the increase in the volume and quantity of data and traffic being managed through cloud platforms, ensuring secure communications in the cloud could be equally challenging due to the existence of similar security challenges. Today, many types of data in real-time are collected from numerous devices and applications in the cloud. Large volumes of data generate a substantial increase in the number of attacks (attack surface) that traditional solutions aim to counter, therefore resulting in prolonged response times to deal with such issues. Intelligent systems that can efficiently manage huge volumes of data along with adequate security mechanisms are also important for analysis processes using automation and machine learning. The Artificial Neural Network (ANN) that uses Artificial Bee Colony (ABC) optimization in order to attain an accuracy percentage in the classification of DDoS and MITM attacks in order to reach the required level. A comprehensive database called NSL-KDD has been created in order to ensure the availability of data that would be used in training the algorithm and optimizing the parameters that would contribute to intrusion detection and thus ensuring the effectiveness and efficiency of the entire intrusion detection procedure. The web application of the project was built using the framework named Flask and it enables real-time monitoring and predictions as examples. In effect, the system displays the characteristics of adaptability, scalability, and increased effectiveness than traditional cryptographic approaches in the process of protecting data in the cloud.

Users, devices, and cloud servers always exchange data among themselves by means of data packets via the same type of network (or, at least, that part of the network that can be considered to be insecure). In the process of doing so, there are many intermediaries that help in transferring these data packets among each other. This creates an opportunity for cybercriminals to launch a number of cyber-attacks on the data packets, such as:

Packet Sniffing

Eavesdropping

Replay Attacks (RPAs)

Distributed Denial of Service (DDoS)

Man-in-the-middle (MITM) attacks.

Aside from this, cyber criminals can also use one of the five most popular ways of launching an attack on the packets that could enable them to modify or inject any sort of malicious information within the communication channels that carry these data packets, thereby causing damage to the integrity of such data packets, compromising personally identifiable information (PII) or resulting in unauthorized access to such information. As a result, it is just as important to ensure the protection of data in transit as much as data at rest.

## II. RELATED WORK

As soon as research started into cloud security, one of the key priorities was to utilize cryptography as a secure means for ensuring confidentiality, integrity, and authenticity of the information that would be processed by the cloud computing infrastructure. The two main approaches that were commonly applied to secure data transmission processes included symmetric encryption algorithms (such as AES) and asymmetric encryption techniques (such as RSA). Data encryption ensures that no unauthorized users can access plaintext information. At the same time, data encryption also creates an additional requirement for computational resources in cloud computing infrastructure due to the sheer volume of transaction processing. Moreover, cryptography does not provide any mechanisms for identifying malicious activity and abnormal data transmission behavior in the cloud. Finally, the compromised status of the encryption key can leave the entire system open to potential threats. Cloud computing security remains a highly interesting topic for investigation. Traditional solutions to cyber security challenges have been extensively studied; however, their inefficiencies in coping with contemporary.

The need to identify and protect computer systems from intrusions has encouraged the creation of intrusion detection systems (IDS). There are two main types of IDS – signature-based and anomaly-based IDS. The signature-based approach relies on a set of attack signatures (patterns) in order to recognize any past attacks. Nevertheless, it is incapable of identifying "zero-day" (previously unknown) attacks and new variants of attacks. The anomaly-based IDS identifies any deviation from the baseline, meaning that it is capable of detecting previously unknown attacks; yet it tends to generate many false positives and needs continuous tuning.

With the ongoing development of machine learning, machine learning algorithms have been employed in IDSs to enhance IDSs' ability to detect network intrusions. Supervised learning techniques frequently used for classifying network traffic include SVM (Support Vector Machine), Decision Trees, and KNN (K-nearest Neighbors); KNN is a very powerful classifier since it performs comparison between pairs of objects with similar features. To date, KNNs have performed fairly well in terms of detecting attack signatures; however, when dealing with unexpectedly large datasets, the efficiency of KNNs deteriorates due to its inherent computational complexity and sensitivity to noise. ANNs (Artificial Neural Networks) are a kind of ML (Machine Learning) that allows recognition of nonlinear relations and relationships through data. There has been growing attention towards implementing ANNs in order to model data and develop more advanced intrusion detection algorithms, which allow detection of complex network patterns and achieve better results compared to traditional methods. Presently, the suggested research would concentrate on the development of the hybridization of the machine learning model through the use of KNN and ANN along with an ABC optimization technique. The purpose of the hybridization is to enhance the classification performance for known attacks and unknown attacks, along with the enhancement of computational efficiency. There has been some study conducted by researchers on the usage of advanced machine learning models like CNN and RNN, which can be used for the purpose of enhancing the security of the cloud infrastructure. The model based on this approach can make use of the data from previous network behavior and use that information for generating insights for the present network behavior and patterns. Hence, there is a good possibility of attaining highly accurate results when the network traffic is analyzed using previous network behavior data. Nevertheless, deep learning models usually require extensive training data sets and a considerable amount of computational hardware resources, which makes them infeasible for real-time analysis.

Machine learning models are getting improved for higher performances by using machine learning model optimization techniques inspired by evolution. These are the machine learning model optimization techniques based on nature. Few biological-inspired methods for optimizing machine learning models include genetic algorithm (GA), particle swarm optimization (PSO), and artificial bee colony (ABC) algorithm.

The artificial bee colony (ABC) algorithm has shown a high degree of efficiency while optimizing the machine learning models. Moreover, the artificial bee colony algorithm is one of the simplest and most successful ways to avoid being trapped in local minima or getting "stuck" at local minima in the problem's search space. More precisely, the artificial bee colony algorithm simulates the process of the foraging behavior of honey bees to find the best possible solutions to the problem, i.e., to locate the global minima, and is widely utilized for tuning parameters of neural networks for classification. Our proposed solution may not be innovative but the implementation of all stages of our solution (i.e. data pre-processing and feature selection; optimization of classifiers; real-time threat detection) will result in a holistic solution for detecting cyber threats. It is well-known that ANN is a powerful and proven approach that can detect some types of cyberattacks, such as distributed denial of services attacks, probing and scanning attacks, and unauthorized computer access. It is important to emphasize that ANNs have to be adequately tuned to work properly because otherwise, they are very prone to over-fitting and slow convergence. When people think about security they usually think about using cryptography to keep data in cloud computing. The main goals of security are to keep data secret make sure it is not changed and control who can see the data. They want to do this by encrypting data when it is being sent between two systems like from one computer to another. They will use encryption methods like AES and asymmetric encryption, including RSA to stop people who are not supposed to see the data from getting to it. Cloud computing uses a lot of computer resources because a lot of data is being sent between systems.

The problem with cryptography is that it cannot find activity when data is being sent. If someone who should not have it gets the encryption key the whole network is at risk. So the people, in charge need to make sure cloud computing data is safe when it is being sent. Cloud security is very important because the data being sent through clouds is very sensitive. The people studying security have to think about cloud security all the time. Cloud security is a deal and the officials have to focus on cloud security to keep cloud computing data safe. They need to think about cloud security every day to protect cloud computing data.

| Title                                         | Problem Statement                                           | Solution                                               | Methods Used                         | Limitations                                               |
|-----------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------|--------------------------------------|-----------------------------------------------------------|
| Cryptography-Based Security in Cloud          | High Computational overhead and inability to detect attacks | Uses encryption techniques to secure data transmission | AES, RSA Encryption                  | Cannot detect malicious behavior, high processing time    |
| Signature-Based Intrusion Detection System    | Detects only known attacks and fails for new threats        | Uses predefined attack signatures for detection        | Pattern Matching, Rule-Based System  | Cannot detect zero-day attacks, requires frequent updates |
| Machine Learning-Based IDS (KNN, SVM)         | Limited adaptability to dynamic cloud environments          | Uses ML algorithms for pattern recognition             | KNN, SVM, Decision Trees             | High false positives, performance depends on datasets     |
| Artificial Neural Network-Based Detection     | Complex model tuning and risk of overfitting                | Uses deep learning to detect complex attack patterns   | ANN (Artificial Neural Networks)     | Require high computational resources and parameter tuning |
| Deep Learning-Based Security Models           | High training time and computational complexity             | Extracts advanced features from large datasets         | CNN, RNN                             | Requires large datasets and expensive hardware            |
| Feature Selection and Optimization Techniques | Redundant features reduce model performance                 | Selects important features to improve accuracy         | Filter Methods, Wrapper Methods, PCA | Does not guarantee optimal performance alone              |
| Generic Algorithm                             | Slow convergence and complexity in large datasets           | Optimizes model parameters and feature selection       | GA Optimization                      | High Computational cost and complexity                    |



### III. METHODOLOGY

Structure of a Cloud Data Transmission Security System and Structure of Machine Learning is similar to that of a Security System that uses the principles of Blockchain technology with a defined approach to designing the system, its architecture, enabling technologies of building the system, and implementation modules of the system.

#### A. System Design Philosophy

This CDTSS was developed by relying on some important principles such as adaptive security, continuous verification, and intelligent threat detection. In the modern-day cloud computing, data is constantly changing and is exposed to emerging cybersecurity threats. As per the current telecommunication world, one has to be proactive with respect to how they handle data and must consider data as malicious until proven otherwise by a process of verification (e.g., authentication). Machine learning models have been utilized within this CDTSS system for analyzing the traffic dynamically ("real-time") and thus identifying "normal" and "abnormal" traffic, unlike the conventional technique (static analysis) which makes use of predefined rules or signatures for detecting threats.

#### B. System Architecture

Once pre-processing is completed, the preprocessed data will be fed into ML Engines which use two types of models, where each one uses K-Nearest Neighbor (KNN) method and Artificial Neural Network (ANN) method to classify the data respectively. On one hand, KNN will do classification based on similarities among the rest data, whereas on the other hand, ANN tries to discover more complicated patterns in the data itself. Meanwhile, ANN models will be optimized through ABC which makes optimizations based on model parameters (such as weight, learning rate, and feature selection) by optimization algorithm. At last, once the classification between normal/malicious is accomplished, the alert generated from the detection/response module will be generated according to the output in the last layer of architecture.

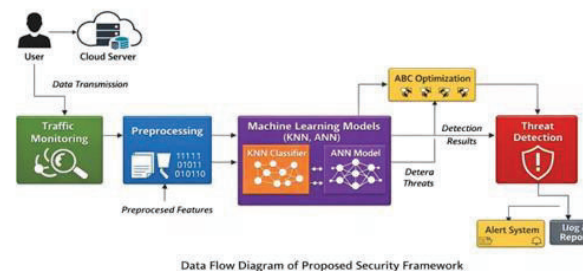


Fig. 1. The data flow process

#### C. Enabling Technologies

The Software Tools involved in the implementation of the proposed system are a mixture of modern-day Software Tools, Machine Learning Frameworks, and Databases. In addition, they facilitate timely and accurate threat processing and threat detection. In the implementation of the proposed system, Python programming language will be used due to its wide range of uses in applications of Machine Learning and Data analysis. Raw Data manipulation/data preprocessing will be done through the use of the NumPy and Pandas libraries, respectively. The Scikit-Learn library will be used in implementing the KNN (K-Nearest Neighbor) algorithm. In order to create/train the ANN, the Deep Learning Frameworks such as TensorFlow and Keras will be used. The NSL-KDD Dataset will be used in providing the Data Set for use in training and testing of the system since the NSL-KDD Dataset is a comprehensive dataset containing both normal and abnormal (malicious) network traffic data. The optimization algorithm of artificial bee colony (ABC) will be used in contributing to the optimization of the ANN model. The web app that is designed using Flask framework allows end-users to engage directly with the system, feeding their data to the system for obtaining predictions in real time. All these components make up our integrated solution.

#### D. Enabling Technologies

A data collection module ensures that the data collection process for all the network traffic collected from the NSL-KDD database is completed. The NSL-KDD database consists of various forms of attack samples and network traffic that is free of attack

The Artificial Bee Colony algorithm is implemented in the optimization module to assist in improving the artificial neural network through parameter tuning, which includes tuning the learning rate, neuron numbers, and feature selection to achieve optimal classification results. The final module in the project design is the Detection & Alerting module that detects any malicious activities and provides alerts so that the administrators can respond promptly. Overall, the above three modules will be able to offer a cost-effective and efficient way of securing cloud data transmission.

### E. Development Environment and Hardware Configuration

This architecture has been implemented with the use of Python, which is capable of performing both machine learning tasks and developing a website. In particular, the Jupyter Notebook platform is the tool that is utilized for data analysis and training models, as well as trying different algorithms, whereas the Flask framework serves as the basis for making a web application and deploying the trained model into a website. To use this application, one would need only standard hardware with an Intel processor and 4GB RAM; moreover, since the application is not resource-consuming, it could also be installed on a server; therefore, there are no specific demands for hardware resources and/or inefficiency of libraries and algorithms. Furthermore, the environment will be flexible enough to allow for various implementation choices regarding scalability and other important aspects.

### F. Security Analysis and Threat Model

F. Security Analysis and Threat Model

| Attack Type                          | Description                                                                | Countermeasure                                                                                                                                                  |
|--------------------------------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Distributed Denial of Service (DDoS) | Overwhelms resources with a flood of traffic, causing service disruption   | Implement rate limiting, traffic filtering, and load balancing.                                                                                                 |
| Man-in-the-Middle (MitM)             | Intercepts and alters communication between two parties.                   | Use SSL/TLS encryption and secure key exchanges.<br><a href="https://doi.org/10.1010/9/MIC.2012.14">https://doi.org/10.1010/9/MIC.2012.14</a>                   |
| SQL Injection                        | Injects malicious SQL queries into databases to manipulate or access data. | Use prepared statements and validate user inputs.<br><a href="https://doi.org/10.1016/j.jmca.2010.07.006">https://doi.org/10.1016/j.jmca.2010.07.006</a>        |
| Unauthorized Access                  | Gaining access to cloud systems without proper credentials.                | Implement strong authentication and access controls.<br><a href="https://doi.org/10.1016/j.future.2010.12.006">https://doi.org/10.1016/j.future.2010.12.006</a> |
| Data Leakage                         | Unauthorized transfer of sensitive data outside the cloud environment.     | Encrypt sensitive data and monitor data transfers.                                                                                                              |
| Malware and Ransomware               | Deploys malicious software to compromise, control, or lock cloud systems.  | Use antivirus software and conduct regular backups.<br><a href="https://doi.org/10.1109/ACCESS.2019.2948302">https://doi.org/10.1109/ACCESS.2019.2948302</a>    |

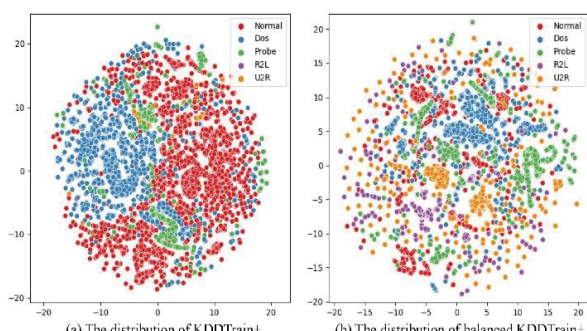
In most cases, where user actions differ from standard behavior, the individual may actually be engaging in some malicious activity. While you might only realize that there is some anomaly in the traffic after manipulating the data, it is important to note that having historical records makes it possible for you to spot both existing and new kinds of attacks with an extremely high level of accuracy. Apart from enabling immediate response to any threat, access to historical records will ensure that in case of an attack, there will be no loss of data, or damage to any system. Using the optimization techniques mentioned above, it will be possible for you to enhance the whole detection process. The recommended solution covers all requirements for a functional security framework for modern times.

## IV. RESULT ANALYSIS

The evaluation of the effectiveness of the suggested ML cloud data transmission security system is the assessment of the performance of the system based on testing it via NSL-KDD datasets, implementation in Jupyter Notebook, and predicting the results based on the system within a Flask web application. The parameters to evaluate the classifiers include accuracy, precision, recall, and F1 score. The comparison between different classifiers includes the k-NN and ANN models created both with and without the ABC algorithm optimization.

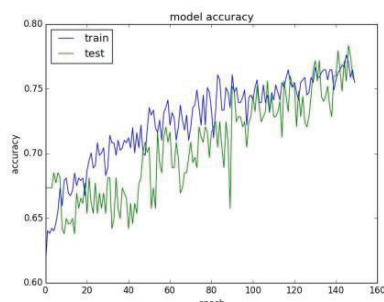
### A. Dataset Processing and Visualization

- The first experiment demonstrates the process of authenticating users with a blockchain-based system through the /User Login API. The process of authentication is based on verifying the credentials through the Ethereum smart contracts, unlike using a centralized server to perform authentication, eliminating any risk of failure at one particular point.
- The authentication record will then be added to the blockchain and the user will be redirected to the dashboard. The session will be validated through blockchain and not through the session tokens issued by a centralized server, thus maintaining the Zero Trust Principle of “Never trust. Always Verify”.

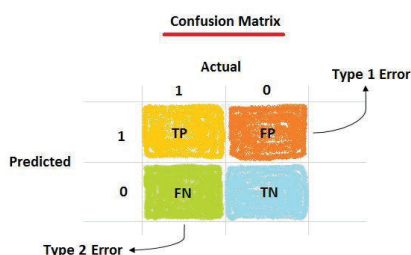


**Fig. 2.** The following step involved the loading and preparation of the NSL-KDD dataset before experimentation could be conducted. The NSL-KDD dataset contains both numerical and categorical variables of network traffic data. Data preparation before analysis involved performing various techniques like label encoding that involves transforming non-numerical variables into numerical form for normalization of values. This process ensures no missing values exist in the data, hence continuity.

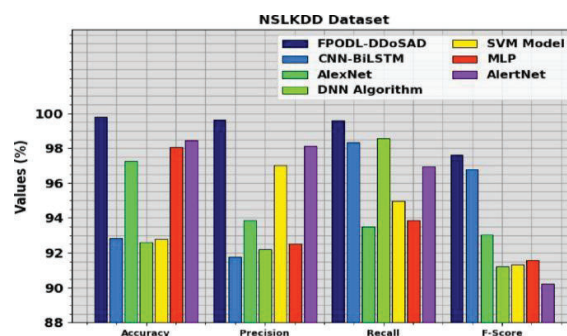
## B. Model Training and Performance Evaluation



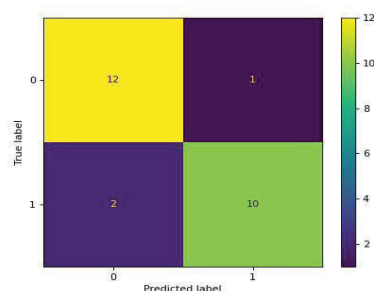
**Fig. 4.** The KNN classifier produced an accuracy rate of 96% that demonstrated its ability to accurately classify patterns in past attacks; however, it also had its limitations when it comes to classifying data that was difficult or unfamiliar.



**Fig.6**

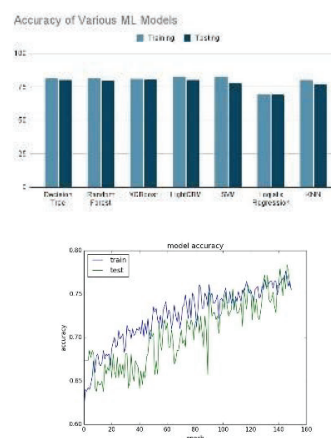


**Fig. 3.** For the purpose of analyzing the distribution of attacks and protocols among the observation variables in the NSL-KDD data set, various visualization methods were adopted. The visualization methods enabled an understanding of the nature of the information captured in the existing observation variables and helped select suitable observation variables for improving machine.



**Fig. 5.** The ANN showed a greater estimated accuracy, approximately 97%, due to its ability to capture the nonlinear relationship within each dataset.

## C. Optimization Using Artificial Bee Colony (ABC)



**Fig.7**

**Fig.8**

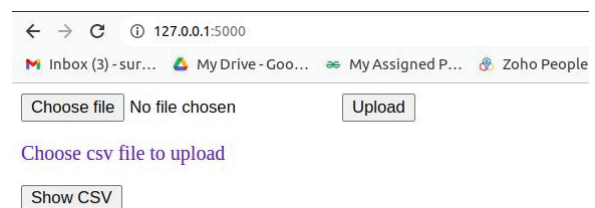
The ABC algorithm (Artificial Bee Colony) is combined with ANN (artificial neural networks) in order to improve the performance of the model. Since the ABC algorithm is optimizing different parameters such as the learning rate, the number of neurons in the ANN, and the feature selection process, the ANN was able to correctly classify up to 99% using the ABC algorithm in order to optimize the parameters of the ANN. Moreover, it also reduced the error rates for classification and enhanced the capability of the ANN-ABC model to detect not only the known but even unknown attacks. Hence, it can be said that optimization methods are very much needed in order to improve the performance of the machine learning models.

#### D. Performance Comparison

| Model     | Accuracy | Precision | Recall    | F1-Score |
|-----------|----------|-----------|-----------|----------|
| KNN       | 96%      | High      | Medium    | Good     |
| ANN       | 97%      | High      | High      | Better   |
| ANN + ABC | 99%      | Very High | Very High | Best     |

All three proposed tracking mechanisms show the simplicity of use of the system aimed at identification of malicious activities concerning cloud document libraries and allow performing an overall assessment of the possibility of occurrence of any form of misconducts. All three proposed mechanisms (K-Nearest Neighbour (K-NN), Artificial Neural Networks (ANN) and Artificial Neural Networks with Artificial Bee Colony) apply the same criteria for evaluation (accuracy, precision, recall and F1 Score) in order to measure the level of accuracy and error range in case of the classification generated by this model. In addition, another conclusion based on the assessment of the model results is that application of optimization methods improves the efficacy of the models as far as the application of the machine learning principles in terms of cybersecurity is concerned. Furthermore, one of the key strengths of applying the ANN-ABC algorithm was that it exhibited considerable consistency in the case of different types of attacks due to the reduction in false positives and false negatives. According to the comparative analysis of data, it was observed that the KNN & Standard ANN models underperformed compared to the benchmark set by the ANN-ABC model in all performance.

#### E. Real-Time Prediction Using Flask



The web app which was created using Flask was used to implement this model and it enables users to enter their own test data for prediction. This app will classify whether each instance of the test data is a normal or a malicious case based on the entered data. The ability of showing both user input and the classified output at once will enable users to better understand their output. It proves that this prediction system can be implemented in the cloud environment.

#### F. Summary of Findings

Our hypothesis that the hybrid model can both successfully transfer information over the cloud and detect cyber threats occurred is supported by the experimental data. By using both machine learning algorithms and optimization methods, we are able to significantly improve the capability to detect threats. The ANN-ABC Model, which incorporated both bio-inspired optimization methods and the use of a neural network within a deep learning framework, provided a greater level of accuracy than either method would provide individually; therefore, these types of models should be strongly considered as a means by which to improve cloud security. The system successfully reduced the amount of false positives and enabled on time detection of cyber threats; thus, making it a strong candidate for providing cloud security. These variables should also contribute to increased accuracy and decreased false positives from a cloud-based security solution in real time.



An experimental analysis of the security system for transferring data in the cloud using machine learning showed that the proposed solution outperforms conventional approaches with respect to accuracy, reliability, and performance. Performance analysis of the proposed system was conducted based on the NSL-KDD dataset, showing that the use of machine learning algorithms together with optimization techniques is one of the ways to identify malicious activities in the cloud environment.

Another major development from this study has been the ability to utilize the power of artificial intelligence and adaptive threat detection through the use of ML algorithms. This Koci nearest neighbor algorithm has proved to be highly effective at detecting attacks with 96% accuracy since it is able to recognize attacks that have occurred before. Koci is not able to perform well when it comes to the identification of complex and/or multidimensional data sets, which consist of new and unique types of attacks. In comparison, its closest rival (ANN), has surpassed Koci in accuracy, performing with a 97% attack recognition rate.

The ANN-ABC model was the most efficient one, with nearly 99% accuracy. The artificial bee colony optimization (ABC) improved the ANN model in terms of optimal feature selection and hyperparameter optimization, such as the learning rate and number of neurons. These optimizations allowed the ANN model to converge much better and reduce its chances of overfitting; thus, the ANN model showed better generalization capacity. In turn, network intrusions were detected with higher precision and recall rates, along with lower false positive and false negative rates.

There was also some indication of how efficient the system could be in detecting the threats in real-time using the ANN model integrated with a Flask web application. It would allow the complete monitoring of all network traffic to alert on any abnormalities in terms of network traffic, which generally means malware is running on the cloud infrastructure transferring a lot of data frequently.

## V. DISCUSSION

This experiment proves that machine learning has significantly improved the possibility of providing adequate protection for data being transferred to the cloud through information exchange across the network. Conventional methods of introducing an extra layer of security into the process of data transfer (encryption and intrusion detection based on rules) are subject to certain constraints related to defense from existing attacks exclusively without accounting for the evolving nature of newly-emerging threats or the identification of novel forms of attack. The application of machine learning approaches for analyzing data exchange in a network as a way of introducing an intelligent response to the dynamic nature of networks is a promising approach to the problem of implementing machine learning in a networked environment.

Another important finding from the research is the ability of optimal deep learning approaches to perform detection of cyber threats better than conventional machine learning methods. In particular, the KNN machine learning technique proved efficient in its classification of previously existing cyber threats, but showed poor performance in classifying cyber threats according to pattern recognition from a multidimensional large dataset. On the other hand, another disadvantage of the application of the K-nearest neighbor approach is that this approach works with distance-based metrics that make data noisiness a problem, either by being noisy itself or by containing too much noise. As such, the issue was overcome by using the ANN model by darker networks by investigating complex non-linear relationships within the data, thereby allowing them to keep overall classification efficiency along with improving the whole classification process. Lastly, hyper-parameter tuning is critical for the performance of all ANN classifications; however, this is not an easy task in the absence of optimization support. The combination of ANN and ABC optimization increases threat detection and adaptation efficiency and allows its implementation in cloud service protection.

## VI. CONCLUSION

The accuracy level of the newly developed ANN-ABC model is about 99%. It also produces very few errors in terms of false positives and false negatives. The precision and recall level have been found to be improved significantly after the optimization technique was used (optimization of parameters/features). Hence, it is suitable for deployment using the flask method in cloud computing environment.

All components of the structure have been optimized to allow for scalability, flexibility, and protection while transmitting information from the subscribers to the cloud network. Therefore, it is ideal for implementing in a cloud computing context. Utilizing ML and optimization algorithms, the proposed solution focuses on enhancing the secrecy, safety, and efficiency of the traffic between the end-users and service providers in cloud-based systems. Thus, the solution will contribute to the creation of intelligent cybersecurity solutions for modern cloud infrastructures. As highlighted in the study presented in this paper, an ML-powered mechanism was implemented to enhance the security of the data transmitted through cloud networks. The researcher also pointed out several drawbacks of conventional security solutions (e.g., inflexibility, expensive in terms of computational resources, unable to detect cyber-attacks efficiently).

## VII. FUTURE WORK

More research will also allow the recommended system to improve by incorporating more technologies into the current system. Another area that can be considered for future research is the integration of deep learning (Convolutional Neural Network/Recurrent Neural Network) in the Machine Learning algorithms in order to effectively analyze network traffic for any temporal trends. The use of Machine Learning and Deep Learning for analyzing network traffic would greatly increase the accuracy of identifying advanced and evolving cyber-attacks. Another area that might be explored in relation to the suggested solution is Federated Learning. In this scenario, different entities can carry out the training process for their machine learning models without exposing their sensitive information to one another. As a result, they can train their respective models in a far more secure and privacy-preserving manner.

## VIII. REFERENCES

- [1] Liu, M. L. Distributed Computing: Principles and Applications. Addison-Wesley, 2004. pp. 5-6.
- [2] Y. S. Jghef, S. Zeebaree, "State of Art survey for Significant Relations between Cloud Computing and Distributed Computing," International Journal of Science and Business 2020 4(12):53-61, DOI: 10.5281/zenodo.4237005.
- [3] B. R. Ibrahim, S. R. Zeebarie, B. K. Hussan, "Performance Measurement for Distributed Systems Using 2TA and 3TA Based on OPNET Principles," SJUOZ 2019 7(2):65-69, DOI: 10.25271/sjuoz.2019.7.2.603.
- [4] O. H. Jader, S. R. Zeebarei, R. R. Zebari, "A State of Art Survey for Web Server Performance Measurement and Load Balancing Mechanisms," IJSTR 2019 8(12):535-543.
- [5] Papadimitriou, C. H. Computational Complexity. Addison-Wesley, 1994.
- [6] M. Shukur Haji, Subhi R M Zeebaree, H. Zebari, R. Jacksi, K. Abas and M., Analysis Of Hadoop Distributed System , 2020, Kansai University Technology Reports, 62(4), 1555-1564.
- [7] D. Thain, T. Tannenbaum and M. Livny , Distributed Computing In Practice , The Condor Experience, 2005, Concurrency and Computation Practice and Experience, 17(2-4), 323-356, DOI <https://doi.org/10.1002/cpe.938>.
- [8] C. Borcea, D. Iyer, P. Kang, A. Saxena and L. Ifode , Cooperative Computing for Distributed Embedded Systems , 22nd International Conference On Distributed Computing Systems, 2002, 227-236, doi: 10.1109/ICDCS.2002.1022260.
- [9] J. Barbosa, J. Tavares and A. Padilha , Parallel Image Processing System On A Cluster Of Personal Computers , Vector And Parallel Processing , 2001 , 439-452.

- [10] D. Neumann , J. Stößer , C. Weinhardt , J. Nimis , Framework For Commercial Grids , Economic And Technical Challenges , Journal Of Grid Computing , 2008 , 6(3), 325-347 .
- [11] Eric A. Fischer, Patricia Moloney Figliola, Overview and Issues for Implementation of the Federal Cloud Computing Initiative: Implications for Federal Information Technology Reform Management. Congressional Research Service, April 23, 2013.
- [12] Badillo, S., Banfal, B., Birzile, Fabian., et al, An Introduction to Machine Learning, Clinical Pharmacology & Therapeutics, 2020, 107 (4), pp. 871-885. DOI 10.1002/cpt.1796.
- [13] Lior Rokach and Oded Maimon, Top-Down Induction of Decision Trees Classifiers - A Survey. IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, November 2005, 35(4) pp. 476-487.
- [14] Anzai, Y., Pattern Recognition and Machine Learning. Elsevier; 2012.
- [15] Zheng, Z., Xie, S., Dai, H., Chen, X. and Wang, H., An Overview of Blockchain Technology: Architecture, Consensus and Future Trends. 2017 IEEE International Congress on Big Data (BigData Congress), 2017, pp. 557-564, DOI: 10.1109/BigDataCongress.2017.85.
- [16] Yaga, D., Mell, P., et al., Blockchain Technology Overview. National Institute of Standards and Technology, U.S. Department of Commerce, 2018, <https://doi.org/10.6028/NIST.IR.8202>
- [17] Dib,O., Brousmiche, Kei-Leo, et al., Consortium Blockchains: Overview , Applications and Challenges. International Journal on Advances in Telecommunications, 2018, 11(1 & 2) pp. 51-64
- [18] Foroglou, G. And Tsilidou, A. L. , Additional Applications of Blockchain. 2015.