

**A Project report on**

**STUDENT PERFORMANCE IN TRAININGS**

A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the  
academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

Submitted by

Ch. SRIDHAM  
(20H51A05K4)

G. POOJITHA  
(20H51A0512)

IMITAZ AHMAD WANI  
(20H51A05Q1)

Under the esteemed guidance of

Ms. G. Srividya  
(Assistant Professor)



**Department of Computer Science and Engineering**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(UGC Autonomous)

\*Approved by AICTE \*Affiliated to JNTUH \*NAAC Accredited with A<sup>+</sup> Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

**2020- 2024**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



**CERTIFICATE**

This is to certify that the Major Project Phase I report entitled “**A BI-OBJECTIVE HYPERHEURISTIC SUPPORT VECTOR MACHINES FOR BIG DATA CYBER SECURITY**” being submitted by **CH. SRIDHAM (20H51A05K4), G. POOJITHA (20H51A0512), IMTAZ AHMAD WANI (20H51A05Q1)** in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Ms.G.Srividya**  
Assistant Professor  
Dept. of CSE

**Dr. Siva Skandha Sanagala**  
Associate Professor and HOD  
Dept. of CSE

## ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Dr. Vijaya Kumar Koppula , Professor of CSE & DEAN** , Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala**, Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana**, Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

Ch. Sridham      20H51A05K4  
G. Poojitha      20H51A0512  
Imtiaz Ahmad Wani      20H51A05Q1

## TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	LIST OF FIGURES	
	ABSTRACT	
<b>1</b>	<b>INTRODUCTION</b>	
	1.1 Problem Statement	1
	1.2 Research Objective	2
<b>2</b>	<b>BACKGROUND WORK</b>	
	2.1. Kernel based SVM for Cyber Security	
	2.1.1. Merits	6
	2.1.2. Demerits	7
	2.1.3. Challenges	
	2.2. Ensemble Learning for Cyber Security	
	2.2.1. Merits	8
	2.2.2. Demerits	8
	2.2.3. Challenges	9
	2.3. Deep Learning for Cyber Security	
	2.3.1. Merits	10
	2.3.2. Demerits	11
	2.3.3. Challenges	13
<b>3</b>	<b>RESULTS AND DISCUSSION</b>	
	4.1. uml diagrams	16
	4.2. Flowchart	17
<b>4</b>	<b>CONCLUSION</b>	18
	5.1 Conclusion	
	<b>REFERENCES</b>	<b>20</b>

**List of Figures**

**FIGURE**

<b>NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
1.1	ump diagram 1	3
1.2	uml diagram 2	3
1.3	uml diagram 3	4

## **ABSTRACT**

Cyber security in the context of big data is known to be a critical problem and presents a great challenge to the research community. Machine learning algorithms have been suggested as candidates for handling big data security problems. Among these algorithms, support vector machines (SVMs) have achieved remarkable success on various classification problems. However, to establish an effective SVM, the user needs to define the proper SVM configuration in advance, which is a challenging task that requires expert knowledge and a large amount of manual effort for trial and error. In this paper, we formulate the SVM configuration process as a bi-objective optimization problem in which accuracy and model complexity are considered as two conflicting objectives. We propose a novel hyper-heuristic framework for bi-objective optimization that is independent of the problem domain. This is the first time that a hyper-heuristic has been developed for this problem. The proposed hyper-heuristic framework consists of a high-level strategy and low-level heuristics. The high-level strategy uses the search performance to control the selection of which low-level heuristic should be used to generate a new SVM configuration. The low-level heuristics each use different rules to effectively explore the SVM configuration search space. To address bi-objective optimization, the proposed framework adaptively integrates the strengths of decomposition- and Pareto-based approaches to approximate the Pareto set of SVM configurations. The effectiveness of the proposed framework has been evaluated on two cyber security problems: Microsoft malware big data classification and anomaly intrusion detection. The obtained results demonstrate that the proposed framework is very effective, if not superior, compared with its counterparts and other algorithms.

# **CHAPTER 1**

## **INTRODUCTION**

# CHAPTER 1

## INTRODUCTION

### 1.1. Problem Statement

The rapid growth of big data in the field of cyber-security has posed significant challenges in effectively detecting and preventing cyber threats. Support Vector Machines (SVM) have emerged as a popular technique for addressing cyber-security challenges. However, the traditional SVM approach often struggles to handle large-scale datasets efficiently, leading to reduced accuracy and increased computational time. Moreover, cyber-security professionals face the challenge of balancing two conflicting objectives: maximizing the detection rate of cyber threats and minimizing false positives. These objectives often require different parameter settings, making it challenging to find an optimal solution that satisfies both objectives simultaneously. To address these issues, this project aims to develop a Bi-objective Hyper-heuristic approach for Support Vector Machines in the context of big data cyber-security. The project seeks to investigate and design an algorithm that can automatically select the most suitable combination of hyper parameters and feature selection techniques to optimize both the detection rate and minimize false positives. The project will involve exploring and developing novel hyper-heuristic algorithms that intelligently search through a large space of potential solutions, including different combinations of SVM hyper parameters, feature selection methods, and other pre-processing techniques. The proposed approach will leverage the power of big data analytics to handle large-scale datasets efficiently, improving the accuracy of cyber threat detection while reducing computational time. By developing a bi-objective hyper-heuristic approach, this project aims to provide cyber-security professionals with a tool that can effectively and efficiently address the challenges of big data cyber-security. The resulting approach will help improve the decision-making process by enabling security analysts to focus on genuine cyber threats while minimizing false positives, ultimately enhancing the overall cyber-security posture of organizations in the face of rapidly evolving cyber threats.



## 1.2. Research Objective

The primary objective of this project is to develop a novel Bi-objective Hyper-heuristic approach for Support Vector Machines (SVM) in the domain of big data cyber-security. The project aims to address the challenges associated with handling large-scale datasets efficiently while optimizing the detection rate of cyber threats and minimizing false positives simultaneously. One objective is to investigate and design an algorithm that can automatically select the most suitable combination of hyper parameters and feature selection techniques for SVM. By leveraging hyperheuristics, the algorithm will intelligently search through a vast space of potential solutions, including different combinations of SVM hyper parameters, feature selection methods, and other pre-processing techniques. This objective will contribute to enhancing the accuracy of cyber threat detection by finding an optimal configuration that improves the performance of SVM on big data cyber-security datasets. Another objective is to develop an approach that can handle the computational demands of big data analytics efficiently. Traditional SVM approaches often struggle with scalability, leading to increased computational time and resource requirements. By employing techniques specifically tailored for big data, such as distributed computing and parallel processing, the project aims to improve the efficiency of SVM-based cyber-security systems. This objective will ensure that the developed approach can handle the growing volume and complexity of cyber-security datasets without compromising performance. In summary, the project's objectives include developing a bi-objective hyper-heuristic approach for SVM that can automatically select optimal hyper parameter and feature selection configurations, improving the scalability and efficiency of SVM-based cyber-security systems for big data, and effectively balancing the detection rate of cyber threats with the minimization of false positives. By accomplishing these objectives, the project aims to contribute to the advancement of cyber-security practices in the era of big data, ultimately enhancing the overall cyber-security posture of organizations and protecting against evolving cyber threats.

## UML DIAGRAMS

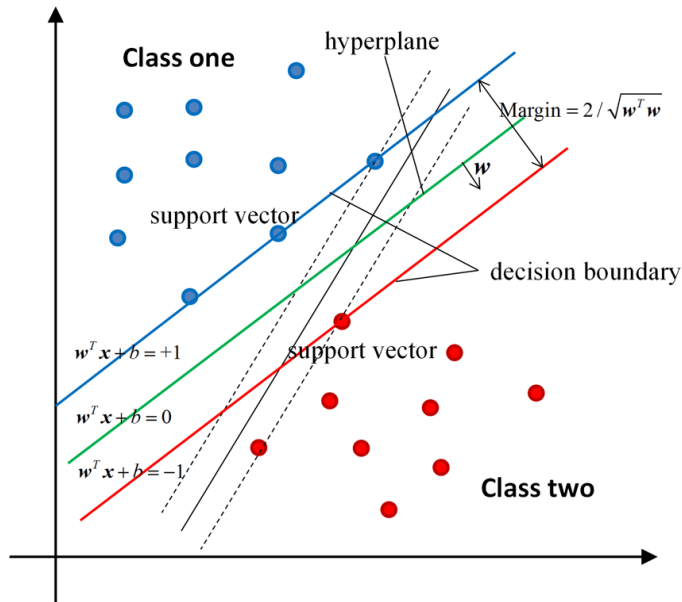


Fig – 1.1 UML DIAGRAM

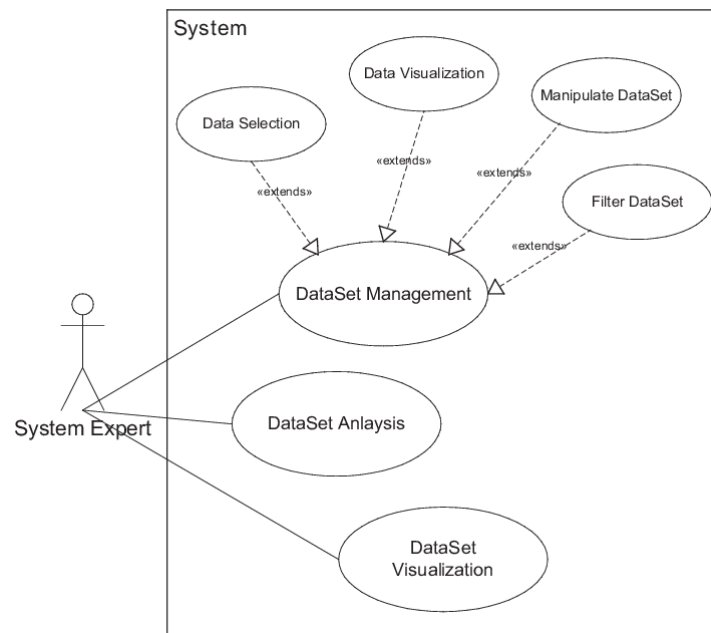


Fig-1.2 UML DIAGRAM

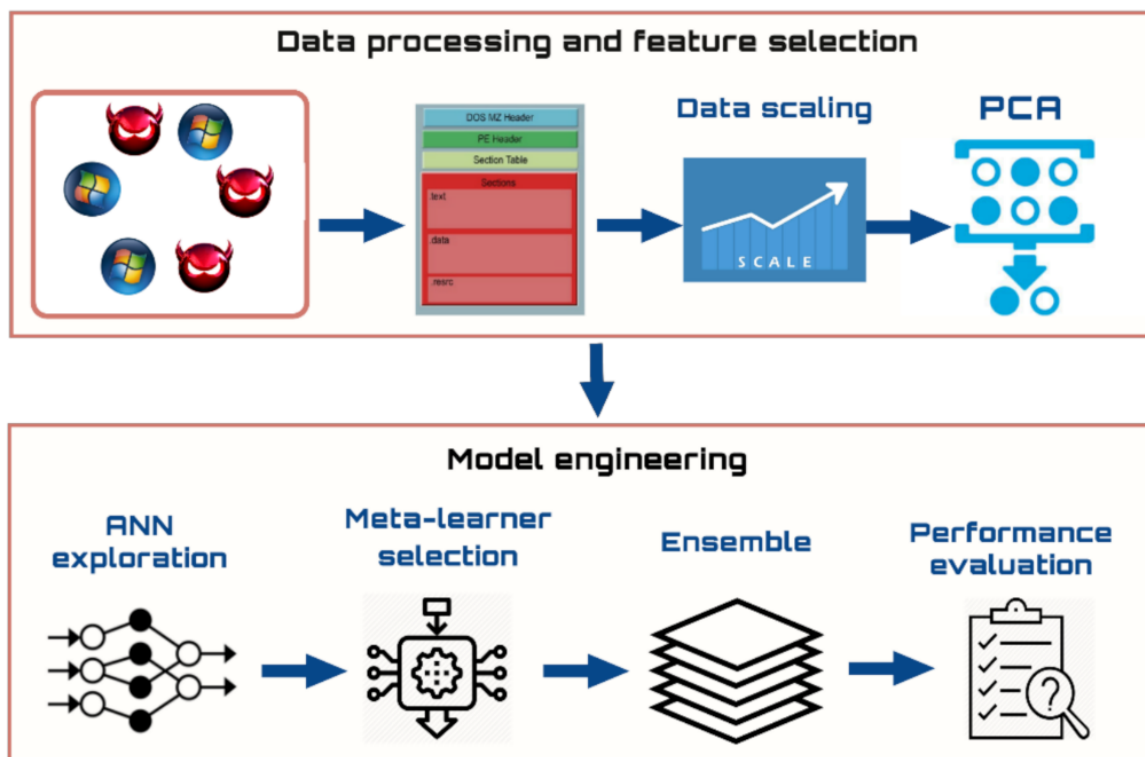


Fig-1.3 UML DIAGRAM

# **CHAPTER 2**

## **BACKGROUND**

### **WORK**

## **CHAPTER 2**

### **BACKGROUND WORK**

#### **EXISTING SYSTEM**

##### **2.1 : Kernel-Based SVM for Cyber-Security**

Deep learning is an advanced model of traditional machine learning. This has the capability to extract optimal feature representation from raw input samples. This has been applied towards various use cases in cyber security such as intrusion detection, malware classification, android malware detection, spam and phishing detection and binary analysis. This paper outlines the survey of all the works related to deep learning based solutions for various cyber security use cases. Keywords: Deep learning, intrusion detection, malware detection, Android malware detection, spam & phishing detection, traffic analysis, binary analysis Cyber security involves protective key data and devices from cyber threats. It's a vital part of corporations that collect and maintain large databases of client data, social platforms wherever personal information were submitted and also the government organizations wherever secret, political and defense information comes into measure. It helps in protecting against vulnerable attacks that possess threat to special data, might or not across numerous applications, networks and devices. With the quantity of individuals accessing the information online which is increasing daily and also the threats to the data are increasing, with the cost of online crimes calculable in billions. Cyber security is that the set of technologies and processes designed to shield computers, networks, programs, and data from attack, unauthorized access, change, or destruction. These systems are composed of network security and host security systems, every of those has a minimum firewall, antivirus computer code, associated an intrusion detection system (IDS). This survey summarizes the importance of cyber security using Deep learning techniques (DL). Deep learning technique is been employed by researchers in recent days. Deep learning may be used along side the prevailing automation ways like rule and heuristics based mostly and machine learning techniques. This study helps is understand the advantage of deep learning algorithms to classify and correlate malicious activities that perceived from the varied sources like DNS, email, URLs etc. Not like ancient machine learning approaches, deep learning algorithms don't follow any feature engineering and have illustration ways. They will extract best options by themselves.

Still, further domain level options got to outline for deep learning ways in information science tasks. The cyber security events thought-about during this study are enclosed by texts. To convert text to real valued vectors, numerous linguistic communication process and text mining ways are incorporated along with deep learning. Deep learning is a prominent algorithm employed in several cyber security areas. Considering several traditional methods and machine learning methods deep learning algorithms considered as a robust way to solve problems. From this study it is clear that most of the deep learning algorithms comes up with better accuracy rate, which will be helpful in building an real time application for analyzing malicious activities over network.

### Merits:

Accuracy: Kernel-based SVMs are known for their high accuracy in classifying cyber threats.

Effective Feature Mapping: They can effectively map input patterns into higher-dimensional spaces, capturing complex relationships in the data.

Flexible Kernel Selection: Kernel functions can be chosen based on the characteristics of the data, allowing for adaptability.

### Demerits:

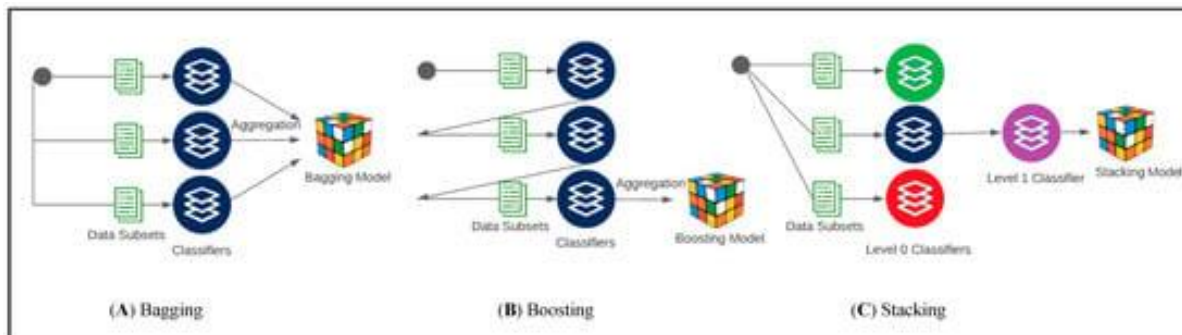
Computationally Intensive: Kernel-based SVMs can be computationally intensive, especially with large datasets.

Manual Hyper-Parameter Tuning: Selecting the appropriate kernel and hyper-parameters often requires expert knowledge and can be time-consuming.

Limited Scalability: Handling big data can be challenging due to memory and processing constraints.

## 2.2: Ensemble Learning for Cyber-Security

Ensemble learning is one of the most well-known ML techniques used in the field of cyber-attack classification and detection. These techniques are usually constructed using multiple ML classifiers to solve the same problems and combine the results with one of the voting techniques.



## Conclusions and Future Work

Wireless Sensor Networks, a key underlying technology of the Internet of Things, are prone to several types of cyber-attacks that could compromise availability, privacy, control, and reliability. One of the effective methods for detecting and mitigating cyber-attacks on WSNs is Machine Learning. Classifier ensembles have been successfully used to detect cyber-attacks in various scenarios. We proposed a new ensemble-based approach, Weighted Score Selector, which uses a pool of conventional supervised ML techniques: Support Vector Machine, Gaussian Naive Bayes, Decision Tree, K-Nearest Neighbor, Random Forest, and LightGBM. The proposed approach's performance was compared to three classical ensemble techniques that integrate the decisions of multiple machine learning models to detect cyber-attacks on WSNs. Performance comparisons were conducted using the following metrics: accuracy, probability of false alarm, probability of detection, probability of misdetection, memory usage, processing time, and prediction time per sample.

Evaluation results indicated that the proposed ensemble approach yielded promising results in terms of probability of detection, probability of false positives, and probability of misdetection. Future work could entail expanding the presented approach to examine more evaluation metrics or additional heterogeneous classifiers using other datasets. This work can be accomplished by extending the pool of base classifiers to include other machine learning algorithms, such as unsupervised and reinforcement learning, and activating a subset of them to contribute to the detection process, depending on the learning conditions and application domain. The activated base classifiers will be selected to match the requirements of the environment where they will be applied.

### Merits:

**Improved Generalization:** Ensemble methods like Random Forest or AdaBoost can improve model generalization and reduce overfitting.

**-Flexibility:** These methods can accommodate a variety of base classifiers and feature selection techniques.

**Diverse Models:** Ensembles combine multiple models to address different aspects of cyber-security, enhancing overall accuracy.

### **Demerits:**

Complexity: Ensembles can be complex to configure and may require substantial computational resources.

Increased Training Time: Training multiple base classifiers can be time-consuming.

Interpretability: The combination of multiple models may make it challenging to interpret decisions.

## **2.3: Deep Learning for Cyber-Security**

### **INTRODUCTION**

Cyber security Internet is open source for web access like for the purpose of railway reservation, online banking, online fees submission etc. Security concern is the most threatening topic for users about their confidential information's storage. various security designing and algorithms has been designed to impose secure environment for user but still malicious activities, codes, algorithms, design are acting on web application to create abnormal behavior for web usage or to steal confidential, secure information for the intension of unauthorized access ,illegitimate access , access for destroying or altering the contents. Attacker's performs Site phishing, Dos attacks, pattern recognition for brute force attack etc, by using several hit and trial or input capturing methods, or by providing capturing codes or IP packets into web contents. Here in the proposed work a technique of detecting malicious Socket address (IP Address and port no.) has been presented, which detects and blocks if any suspicious cases are found and passes the contents to concern user. Here we use SVM technique for classification, detection and prediction of Blacklisted IP addresses and blacklisted port's addresses. The proposed algorithm provides accuracy of 96.99% and which is the best among the present systems. It is light weight system and easy to implement on existing applications. Internet looks like a web of unknown routes or path, numbers of methods or ways are available for accessing web application, its contents, internet application follows OSI model, where each layer has various Protocols, security mechanism, filtering, and encapsulation. But various route means various point for malicious injection by attacker to create injury in web contents. Cyber attacks [1] are actions that attempt to bypass security mechanisms of computer system. A number of cyber attack detection and classification methods have been introduced with different levels of success that is used as a countermeasure to preserve data integrity and system availability from attacks.

Cyber Security [2] concepts are studied for protecting web application from malicious data



A Bi-Objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security injection or from the occurrence of fraudulent scheme. Designs, tools, Algorithm and architectural testing etc. Here security for web application contents is presented by using Machine learning techniques. Frameworks are being designed for the purpose of by detecting, checking and blocking attack signatures and attack procedure and patterns. Attacker's work by noticing or by finding bypassing mechanism to access secure connections and designs. Every attack have their predefined levels, steps and pattern but they grows as the security implication increases, as security increases ,attack potential also increases. Every web-browser has their own deign pattern and algorithms like internet explorer, Opera, Google chrome, Netscape navigator, Mozilla Firefox, some of them works on HTTP and some on HTTPS. Https provides secure channel (Contains encryption, key exchange or algorithms schemes for packets) for web application and are less susceptible to attack but http (Packet is transferred from source to destination in original form without any encryption or key exchange) is susceptible to attack.

Our concept provides a secure application, based on classification of original and suspicious IP address and port address (socket) using SVM. Here dataset of different size is used for training and classification. Different parameters like Accuracy, detection time, training time, TPR, TNR, FPR, FNR and the graphical description shows the performance of our system. Our system shows the best performance result in accuracy which is 95.6% and best among the existing systems.

### **Merits:**

Feature Learning: Deep learning models can automatically learn relevant features from data, reducing the need for manual feature engineering.

High Accuracy: When properly trained, deep learning models can achieve state-of-the-art accuracy in cyber-threat detection.

Adaptability: They can adapt to evolving threats through retraining on new data.

### **Demerits:**

Data Requirements: Deep learning models often require large amounts of labeled data for training.

Computational Demands: Training deep neural networks can be computationally intensive, requiring powerful hardware.

**Black-Box Nature:** Deep learning models can be challenging to interpret and explain.

Researchers and practitioners in the field of big data cyber-security often need to carefully consider these factors when choosing an appropriate method for their specific use cases.

**Challenge kernel based svm for cyber security:**

**Scalability:** Adapting kernel-based SVMs for big data cyber-security is a significant challenge.

**Automatic Parameter Selection:** Developing automated methods for selecting the optimal kernel and hyper-parameters is necessary.

**Interpretability:** Ensuring that the decisions made by the model are interpretable and explainable in the context of cyber-security.

**Challenge Ensemble learning for cyber security:**

**Scalability:** Ensembles, like SVMs, can face scalability issues with large datasets.

**Model Selection:** Determining the optimal ensemble size and base classifiers is a challenge.

**Trade-offs in Diversity:** Achieving the right balance between diversity and accuracy in the ensemble can be tricky.

**Challenge Deep learning for cyber security:**

**Data Availability:** Obtaining labeled data for training deep models in cyber-security is a significant challenge.

**Model Complexity:** Finding the right architecture and hyper-parameters for deep networks is a non-trivial task.

**Interpretability:** Addressing the lack of interpretability in deep learning models is crucial for cyber-security applications

## **DISADVANTAGES OF EXISTING SYSTEM**

### **1. Limited Scalability:**

Traditional SVM-based systems often struggle to handle large-scale datasets efficiently. As big data continues to grow in the field of cyber-security, existing systems face limitations in terms of computational resources, memory requirements, and processing speed. This limitation hampers their ability to effectively analyse and detect cyber threats in real-time, hindering their practical application in big data cyber-security scenarios.

## **2. Lack of Automatic Hyper Parameter Selection:**

Existing SVM-based systems typically rely on manual selection of hyper parameters, such as the choice of kernel function, regularization parameter, and kernel-specific parameters. This manual tuning process is time-consuming, requires expert knowledge, and may not guarantee optimal performance across different datasets and cyber-security scenarios. The lack of automatic hyper parameter selection limits the efficiency and adaptability of existing systems.

## **3. Inability to Handle Conflicting Objectives:**

Cyber-security professionals often face the challenge of balancing multiple objectives, such as maximizing the detection rate of cyber threats while minimizing false positives. Existing systems typically optimize a single objective, such as accuracy or false positive rate, without considering the trade-offs between conflicting objectives. This limitation makes it difficult to find an optimal solution that satisfies both objectives simultaneously, leading to suboptimal decision-making and resource allocation.

## **4. Lack of Flexibility in Feature Selection:**

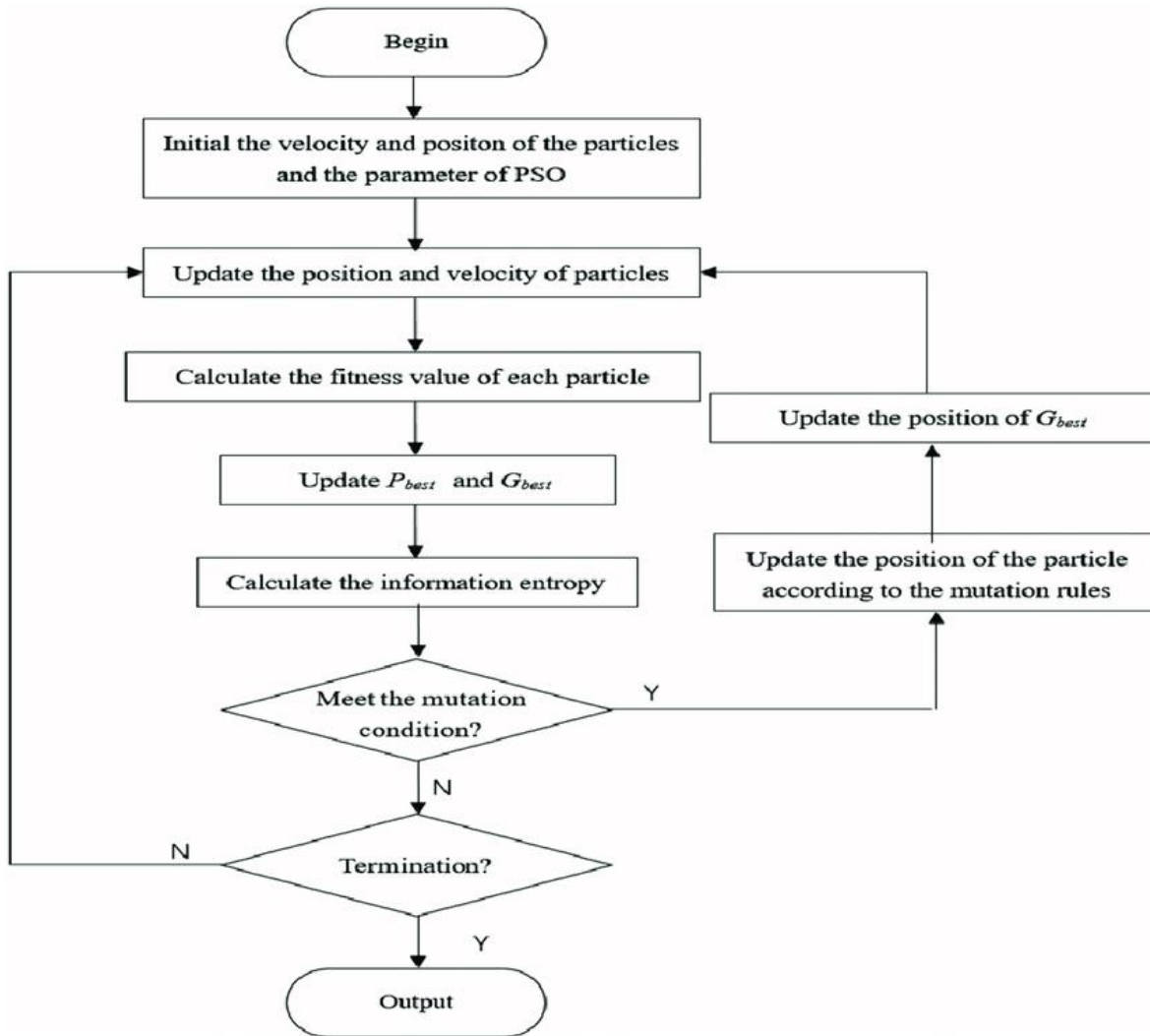
Feature selection plays a crucial role in SVM-based systems for cyber-security. However, existing systems often lack flexibility in incorporating different feature selection techniques tailored for big data. They may rely on manual feature engineering or use generic feature selection methods that do not consider the specific characteristics and requirements of cyber-security datasets. This limitation restricts the system's ability to effectively extract relevant features from big data and may lead to suboptimal performance.

## **5. Limited Adaptability to Evolving Cyber Threats:**

The field of cyber-security is constantly evolving, with new types of cyber threats and attack vectors emerging regularly. Existing SVM-based systems may struggle to adapt to these evolving threats due to their static nature. They may lack the ability to dynamically update and retrain the models based on the latest threat intelligence and changing data patterns. This limitation reduces their effectiveness in accurately detecting and preventing new and unknown cyber threats. A Bi-Objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security 12

## 6. Lack of Explainability:

Interpretability and explainability are crucial aspects in the field of cybersecurity. However, existing SVM-based systems may lack transparency in their decision-making process. They often provide limited insights into the reasons behind their classifications, making it challenging for security analysts to understand and validate the system's outputs. The lack of explainability can undermine trust in the system and hinder its adoption in critical cyber-security operations.



# **CHAPTER 3**

## **RESULTS AND DISCUSSION**

## **CHAPTER 3**

### **RESULTS AND DISCUSSION**

#### **Result and Disussion:**

**Accuracy and Detection Rates:** One of the key results when using Kernel-Based SVMs in cybersecurity is often a high degree of accuracy in threat detection. This means that the system is effective at correctly classifying malicious activities or threats, which is a critical factor in cybersecurity. Discussion should focus on how the choice of kernel and other hyperparameters influenced the accuracy.

**Trade-offs with Complexity:** In a cybersecurity setting, achieving high accuracy is essential, but it often comes at the cost of model complexity. This complexity can be reflected in the number of support vectors, which should be discussed. Researchers may explore the trade-off between a larger number of support vectors (which can lead to overfitting) and a smaller number of support vectors (potentially resulting in less accurate classification).

**Impact of Kernel Selection:** Different types of kernels, such as linear, radial basis function (RBF), or polynomial, have different impacts on the SVM's performance. The results should discuss how the choice of kernel function affected the model's accuracy and ability to generalize to different types of cyber threats.

#### **Scalability Challenges:**

When using Kernel-Based SVMs in cybersecurity, one common result might be scalability challenges, especially when dealing with large datasets. The discussion should address how the model's performance deteriorated as the dataset size increased and what measures were taken to mitigate these challenges.

**Hyperparameter Optimization:** The results should highlight the importance of selecting the right hyperparameters, such as the regularization parameter (C) and kernel-specific parameters. Researchers may discuss the methods used for hyperparameter optimization, as well as their impact on model performance.

**Comparison to Other Methods:** To provide a comprehensive discussion, it's common to compare Kernel-Based SVMs to other machine learning or deep learning methods used in cybersecurity. This comparison can highlight the advantages and disadvantages of SVMs in this specific context.

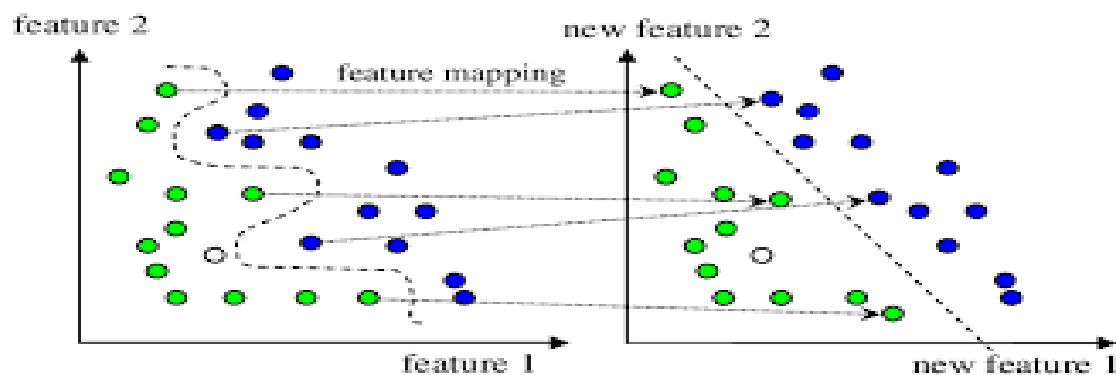
**Interpretability and Explainability:** Given the importance of interpretability in cybersecurity, the discussion should address how Kernel-Based SVMs can be made more interpretable.

Researchers may propose approaches to visualize or explain the decision boundaries of the SVM.

**Robustness and Resilience:** Kernel-Based SVMs should be evaluated for their resilience to different types of attacks or evasion attempts. The results and discussion should emphasize the model's ability to detect new and evolving cyber threats.

**Real-world Applications:** If available, researchers should discuss the real-world applications of Kernel-Based SVMs in cybersecurity. This can include case studies where the method was applied in practical scenarios and its impact on improving security.

In conclusion, the results and discussion of Kernel-Based SVMs for cybersecurity should address accuracy, complexity, scalability, the impact of kernel selection, hyperparameter optimization, comparisons to other methods, interpretability, resilience, and real-world applications. These discussions will help researchers and practitioners understand the performance and utility of Kernel-Based SVMs in addressing cyber threats.



# CHAPTER 4

## CONCLUSION



## CHAPTER 4

### 4.1 CONCLUSION

In conclusion, the literature survey revealed several limitations in existing systems for big data cyber-security based on Support Vector Machines (SVM). These limitations include limited scalability, manual hyper parameter selection, inability to handle conflicting objectives, inflexibility in feature selection, limited adaptability to evolving cyber threats, and lack of explainability. These shortcomings highlight the need for an improved approach that can address these challenges effectively. A Bi-Objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security 14 The proposed system, a Bi-objective Hyper-heuristic approach for SVM in big data cybersecurity, aims to overcome these limitations. By leveraging hyper-heuristics, the proposed system will automatically select optimal hyper parameters and feature selection techniques, improving the accuracy and efficiency of SVM-based cyber-security systems. Additionally, the bi-objective optimization framework will balance the detection rate of cyber threats with the minimization of false positives, providing security analysts with a decision support tool. The development of the proposed system represents an important step towards enhancing the cyber-security posture of organizations in the era of big data. By addressing the limitations of existing systems, the proposed approach has the potential to improve the scalability, adaptability, and decision-making capabilities of cyber-security systems, leading to more effective threat detection and prevention.

# REFERENCES

## REFERENCES

- N. R. Sabar, X. Yi and A. Song, "A Bi-objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security," in IEEE Access, vol. 6, pp. 10421-10431, 2018, doi: 10.1109/ACCESS.2018.2801792.
- Herbert Schildt, Edition (2003) 'The Complete Reference JAVA 2' Tata McGraw Hill Publications
- M. Ahmadi, D. Ulyanov, S. Semenov, M. Trofimov and G. Giacinto, "Novel feature extraction selection and fusion for effective malware family classification", Proc. 6th ACM Conf. Data Appl. Secur. Privacy, pp. 183-194, 2016.
- Michael Foley and Mark McCulley, Edition (2002) 'JFC Unleashed' Prentice-Hall India
- M. P. Basgalupp, R. C. Barros and V. Podgorelec, "Evolving decision-tree induction algorithms with a multi-objective hyper-heuristic", Proc. 30th Annu. ACM Symp. Appl. Comput., pp. 110-117, 2015.