



# International Journal of HRM and Organizational Behavior



[www.ijhrmob.com](http://www.ijhrmob.com)

[editor@ijhrmob.com](mailto:editor@ijhrmob.com)

# RANSOMWARE DETECTION USING PROCESSOR AND DISK USE DATA

K YATHEENDRA<sup>1</sup>, TARIMINI SUREKHA<sup>2</sup>, Y MADHURI<sup>3</sup>

<sup>1</sup>Associate Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,  
Email: k.yatheendra84@gmail.com

<sup>2</sup>P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email:  
tsurekha10@gmail.com

<sup>3</sup>Assistant Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,  
Email:madhuri.yandra94@gmail.com

**Abstract:** Process observing and information investigation are insufficient for ransomware identification, consequently the venture tends to it. A dependable and successful ransomware recognition framework for virtual machines is the objective. Explicit central processor and plate I/O occasions for the entire VM from the host framework are gathered. The venture utilizes machine learning (ML), quite an random forest (RF) classifier, to develop a recognition model. Ransomware defilement and checking above are diminished with this method. The recommended arrangement defeats a ransomware location trouble by being strong to client responsibilities. Staying away from persistent checking of each and every objective machine process keeps the model adaptable to different client conditions. We assess the venture utilizing 22 ransomware tests and client jobs. This Venture gives a precise recognition technique to battle ransomware. The task diminishes checking costs, speeds discovery, and adjusts to changing ransomware varieties by utilizing picked computer processor and circle I/O occasions and ML. This work added Convolutional Neural Network 2D (CNN2D) and an outfit model with a democratic classifier to improve ransomware detection. The voting classifier, which consolidated various ML classifiers, made last forecasts with 99% accuracy, demonstrating that joining models further develops identification.

**Index terms** - Deep learning, disk statistics, hardware performance counters, machine learning, ransomware, virtual machines.

## 1. INTRODUCTION

Ransomware locks or scrambles documents on an objective PC, making it pointless. Cybercriminals coercion casualties utilizing ransomware. Country state entertainers might harm their rivals' key foundation by means of ransomware. These assaults for the most part include information exfiltration to drive casualties to pay a payment or sell their information on the dull web. Around 70% of firms were ransomware casualties in 2022 [1]. By 2031, ransomware will focus on a business, client, or device like clockwork, up from 11 seconds in 2021. Harm cost \$20 billion of every 2021 and may approach \$265 billion. By 2031 [2]. Ransomware discovery has been concentrated on by a few specialists.[57]

Signature-based location [3], [4] breaks down the objective framework for records that match antivirus programming hash values for known ransomware. Polymorphic and transformative malware can keep away from signature-based location [4], [5]. Accordingly, conduct or runtime ransomware location supplements signature-based discovery. The social examination is a unique examination of the

ransomware's movement following disease. Dissimilar to other infections, ransomware should follow a specific succession to encode however many records as doable quickly. As of late delivered ransomware like LockBit2.0, Darkside, and BlackMatter encode just the underlying few bytes of documents to quickly hinder more records [6]. Since ransomware should quickly encode client information, its runtime conduct might vary from that of a harmless application. The thought is that a ransomware-contaminated machine should act unusually. Ransomware should get to records from the hard circle and use the computer chip to encode information, which increments action. ML approaches can recognize this.

Runtime identification on the objective PC incorporates observing cycles, parts, and subsystems, gathering occasion information, and assessing it for peculiarities [7], [8]. Ransomware might make new cycles and exercises to conceal its runtime movement. A framework under attack has elevated movement, which might be distinguished with legitimate examination. Runtime identification on the objective PC is asset escalated and obtrusive since ransomware projects might be difficult to perceive and many cycles should be watched. Ransomware planned to close down running cycles prior to encoding documents can impair such checking.

HPCs are unique registers that count processor and framework occasions per cycle or framework. Current processors can count many processor and framework occasions, including guidance execution, reserve misses, and off-chip memory gets to. HPC information is frequently used for framework programming execution examination and tuning. Be that as it may, current review has inspected their malware recognition utilization [9], [10], [11], [12], [13], [14]. Alam et al.

[15] utilized HPC information from every framework cycle. Checking a few cycles could slow the framework, making it unworkable. Pundir et al. [7] acquired machine-level information. Nonetheless, their examinations are restricted to a solitary Windows virtual machine (VM) responsibility, and changing the quantity of projects may decisively influence identification exactness.

## 2. LITERATURE SURVEY

Internet is compromised by deceptions, Worms, and Spyware. Despite the fact that malware and its varieties vary significantly from content marks, they share more elevated level conduct characteristics that better show malware's goal. [4] Malware conduct extraction, the formal [3, 5, 9, 10, 12] Malware Behavior Feature (MBF) extraction technique, and the unsafe conduct highlight based malware identification calculation are analyzed in this review. We at long last made the MBF-based malware recognition framework, and analyses exhibit it can recognize new dangers.

Crypto-ransomware, among other infections, monetarily coerces individuals by scrambling their documents and keeping them prisoner. This causes yearly worldwide misfortunes of millions. A rising number of ransomware variants might sidestep hostile to infections and programming just malware discovery frameworks that influence static execution marks. [7] The RanStop equipment helped approach we depict in this study recognizes crypto-ransomware disease in ware processors early. RanStop notices miniature structural occasion sets and recognizes known and unseen crypto-ransomware variations involving equipment execution counters in current processors' exhibition observing units. This article prepares an

recurrent neural network-based ML design utilizing long short-term memory (LSTM) [15, 52, 54] to dissect equipment miniature engineering occasions while executing various ransomware variations and harmless applications. From associated HPCs' timeseries, we produce inborn factual elements to build RanStop's discovery exactness and limit commotion utilizing LSTM [52,54] and worldwide normal pooling. RanStop utilizes HPC information from 20 timestamps 100us separated to recognize ransomware inside 2ms of program execution. This location time is too soon for ransomware to cause serious harm. RanStop can distinguish ransomware with 97% accuracy in 50 random preliminaries when approved against harmless projects with social (sub-schedule driven) similarity to crypto-ransomware.[59]

Ransomware has of late (re)emerged as a well known infection that objectifies people and companies for cash. Current ransomware discovery components neglect to give an early admonition continuously, bringing about irreversible encryption of countless records. Post-encryption procedures (e.g., key extraction, document rebuilding) are additionally restricted. [27], [28] High misleading up-sides come about because of existing recognition instruments neglecting to recognize ransomware encryption and client started document activities (e.g., harmless encryption or pressure). This paper [8] presents RWGuard, a ransomware identification component that can identify crypto-ransomware continuously on a client's machine by (1) utilizing distraction strategies, (2) cautiously checking running cycles and the record framework for vindictive exercises, and (3) learning clients' encryption conduct to discard harmless document changes. Our methodology is tried against 14 of the most widely recognized ransomware families [22], [23], [24], [25], [26]. Our examinations exhibit

that RWGuard recognizes ransomware progressively with 0% false negatives and as low as 0.1% false positives, with an above of around 1.9%.

The development of PCs in any area prompts malware. Numerous frameworks, including the most recent versatile ones, have infections, rootkits, spyware, adware, and other malware. Due of the numerous ways of evading anti-virus (AV) programming, malware dangers persevere and create. Aggressors use AV programming blemishes to get into frameworks these days. This exploration [9] researches in the event that presentation counters might be utilized to foster an equipment malware discovery. We utilize execution counter information to distinguish malware and find that our identification strategies are impervious to little malware program alterations. We can distinguish a few varieties inside a malware family on Android ARM and Intel Linux frameworks subsequent to assessing a restricted assortment of varieties. Our proposed equipment changes permit the malware identifier to run securely behind the framework programming, empowering more straightforward, less buggy AV arrangements than programming AV. Equipment AV approaches' power and security could further develop online infection recognition.

Ongoing examination recommends microarchitectural execution examples can recognize malware. Signature-based locators catch malware by looking at a program's execution design (signature) to known malware programs. In this work [10], we present irregularity based equipment malware identifiers that don't need marks for malware recognition [9], [10], [11], [12], [13], [14], and can distinguish more malware, including novel ones. In light of execution counter information, we apply unsupervised ML to foster profiles of typical program execution and

recognize significant varieties in program conduct brought about by malware double-dealing. We show that certifiable double-dealing of famous applications like IE and Adobe PDF Peruser on Windows/x86 can be dependably distinguished. We likewise examine the restrictions and trouble of sending this technique against a gifted enemy attempting to keep away from peculiarity based discovery. The proposed locator supplements signature-based indicators and further develops security.

### 3. METHODOLOGY

#### i) Proposed Work:

The proposed virtual machine ransomware recognition strategy is creative. From the host PC, it catches extensive central processor and circle I/O occasions. Strong location models are made utilizing ML, explicitly an RF classifier [52]. This technique keeps away from the weight of observing each target machine process, restricting ransomware information pollution. It additionally adapts to client responsibility changes. With the [52] RF classifier outflanking different classifiers, the proposed framework distinguishes known and new malware rapidly and precisely. This article added Convolutional Neural Network 2D (CNN2D) and a gathering model with a voting classifier to increment ransomware discovery exactness. The voting classifier, which joined various ML classifiers, made last forecasts with 99% accuracy, demonstrating that consolidating models further develops identification.

#### ii) System Architecture:

This study presents fast ransomware discovery on a Windows 10 VM. We catch HPC and circle I/O information on the host PC. The virtual machine (VM)

knows nothing about checking and information gathering, which meaningfully affects its exhibition. [24] Machine learning(ML) calculations assess information and distinguish ransomware execution [52]. Our answer safeguards cloud VM clients well. We use have PC HPC and circle I/O measurements to accurately recognize malware. Our strategy keeps away from the expense of checking different objective machine activities and information tainting by ransomware intended to forestall reconnaissance.

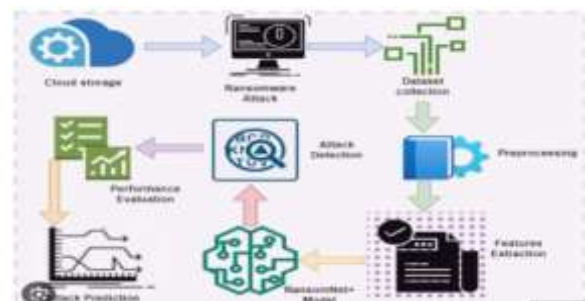


Fig 1 Proposed architecture

#### iii) Dataset collection:

The examination utilizes a HPC dataset that records processor and circle I/O occasions all through virtual machine running. This dataset is painstakingly chosen to mirror an assortment of framework activities, making it ideal for ransomware location model preparation and testing. The HPC dataset reproduces genuine world ransomware ways of behaving utilizing known ransomware tests for model alignment and obscure examples for power testing [16], [17], [18], [19], [20].

ProcessID	ProcessName	ParentProcessID	ParentProcessName	ProcessType	ProcessState	ProcessPriority	ProcessSessionID	ProcessSessionName	ProcessSessionPath	ProcessSessionType	ProcessSessionPath	ProcessSessionType
0	System	0	System	System	Running	High	0	System	C:\Windows\System32\smss.exe	System	C:\Windows\System32\smss.exe	System
4	smss.exe	0	System	System	Running	High	0	System	C:\Windows\System32\smss.exe	System	C:\Windows\System32\smss.exe	System
8	csrss.exe	4	smss.exe	System	Running	High	0	System	C:\Windows\System32\csrss.exe	System	C:\Windows\System32\csrss.exe	System
12	conhost.exe	8	csrss.exe	System	Running	High	0	System	C:\Windows\System32\conhost.exe	System	C:\Windows\System32\conhost.exe	System
16	cmd.exe	12	conhost.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
20	cmd.exe	16	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
24	cmd.exe	20	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
28	cmd.exe	24	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
32	cmd.exe	28	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
36	cmd.exe	32	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
40	cmd.exe	36	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
44	cmd.exe	40	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
48	cmd.exe	44	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
52	cmd.exe	48	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
56	cmd.exe	52	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
60	cmd.exe	56	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
64	cmd.exe	60	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
68	cmd.exe	64	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
72	cmd.exe	68	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
76	cmd.exe	72	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
80	cmd.exe	76	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
84	cmd.exe	80	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
88	cmd.exe	84	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
92	cmd.exe	88	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
96	cmd.exe	92	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System
100	cmd.exe	96	cmd.exe	System	Running	High	0	System	C:\Windows\System32\cmd.exe	System	C:\Windows\System32\cmd.exe	System

Fig 2 Dataset

iv) Data Processing:

Data processing transforms raw information into business-helpful data. Information researchers accumulate, sort out, clean, check, break down, and orchestrate information into diagrams or papers. Data can be handled physically, precisely, or electronically. Data ought to be more significant and decision-production simpler. Organizations might upgrade activities and settle on basic decisions quicker. PC programming improvement and other mechanized information handling innovations add to this. Big data can be transformed into significant bits of knowledge for quality administration and independent direction.

v) Feature selection:

Feature selection chooses the most steady, non-repetitive, and pertinent elements for model turn of events. As data sets extend in amount and assortment, purposefully bringing down their size is significant. The fundamental reason for feature selection is to increment prescient model execution and limit processing cost.

One of the vital pieces of feature engineering is picking the main attributes for machine learning algorithms. To diminish input factors, feature selection methodologies take out copy or superfluous elements and limit the assortment to those generally critical to the ML model. Rather than permitting the ML model pick the main qualities, feature selection ahead of time enjoys a few benefits.[61]

vi) Algorithms:

**Long Short Term Memory (LSTM):** LSTM RNNs take care of the disappearing angle issue in normal

RNNs. A memory cell gives the model catch long haul conditions access consecutive information, making it reasonable for time series or successive example issues. [15] The examination might utilize LSTMs to demonstrate and fathom fleeting associations, which are basic in ransomware detection when framework occasions and ways of behaving matter. [45] Over the long run, LSTMs can record complex examples, improving ransomware detection.

```
LSTM
X_train = X_train.values
X_test = X_test.values

from keras.layers import LSTM
from keras.models import Sequential
from keras.callbacks import EarlyStopping, ModelCheckpoint

# Create the LSTM model
model = Sequential([LSTM(100, input_shape=(X_train.shape[1], X_train.shape[0]))])

# Compile the model
model.compile(optimizer='adam', loss='mse')

# Train the model
model.fit(X_train, y_train, validation_data=(X_test, y_test), callbacks=[EarlyStopping(monitor='val_loss', patience=10), ModelCheckpoint(filepath='lstm_model.h5', save_best_only=True)])

# Predict on test data
y_pred = model.predict(X_test)
```

Fig 3 LSTM

**Deep Neural Network (DNN):** Deep neural networks incorporate a few secret layers among information and result. These networks can learn muddled progressive information portrayals, making them helpful for highlight reflection and portrayal errands. The task might utilize DNNs since they can learn complex information attributes and connections. DNNs can remove elements and train undeniable level portrayals for ransomware detection, where nuanced and confounded examples might happen [8], [13], [14].

```
DNN
# Import necessary libraries
import numpy as np
import tensorflow as tf
from tensorflow.keras import layers, models

# Prepare training and testing data
X_train, y_train, X_test, y_test = load_data()

# Create the DNN model
model = models.Sequential([
    layers.Dense(100, activation='relu'),
    layers.Dense(50, activation='relu'),
    layers.Dense(10, activation='softmax')
])

# Compile the model
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

# Train the model
model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=10)

# Predict on test data
y_pred = model.predict(X_test)
```

Fig 4 DNN

**XGBoost:** Extreme Gradient Boosting (XGBoost) is a gradient boosting ML calculation. It consecutively fosters a troupe of powerless students (ordinarily decision trees) that right each other's blunders, making a hearty and exact model. XGBoost is probable utilized for its order and enormous dataset proficiency. XGBoost can estimate ransomware activities and assist with making an exact discovery model [8], [13], [14].

```

XGBoost

#now train XGBoost algorithm
xgb_cls = XGBClassifier(n_estimators=10, learning_rate=0.09, max_depth=2)
xgb_cls.fit(X_train, y_train)
predict = xgb_cls.predict(X_test)
calculateMetrics("XGBoost", predict, y_test)

rf_acc = accuracy_score(predict, y_test)
rf_prec = precision_score(predict, y_test, average='macro')
rf_rec = recall_score(predict, y_test, average='macro')
rf_f1 = f1_score(predict, y_test, average='macro')

storeResults('XGBoost', rf_acc, rf_prec, rf_rec, rf_f1)
    
```

Fig 5 Xgboost

**Random Forest:** Random Forest fabricates a few choice trees during preparing. It yields the class mode or relapse mean of individual trees. The undertaking involves Random Forest for muddled arrangement errands. An Random Forest outfit can improve ransomware recognition by incorporating numerous choice trees to make a strong and reliable model.

```

#training random forest algorithm
rf = RandomForestClassifier(n_estimators=10, criterion='gini', max_features='log2', min_weight_fraction_leaf=0.3)
rf.fit(X_train, y_train)
predict = rf.predict(X_test)
calculateMetrics("Random Forest", predict, y_test)

rf_acc = accuracy_score(predict, y_test)
rf_prec = precision_score(predict, y_test, average='macro')
rf_rec = recall_score(predict, y_test, average='macro')
rf_f1 = f1_score(predict, y_test, average='macro')

storeResults('Random Froest', rf_acc, rf_prec, rf_rec, rf_f1)
    
```

Fig 6 Random forest

**Decision Tree:** Every hub in a Decision Tree demonstrates a choice in light of information highlights. It recursively isolates the dataset into

subgroups to track down terminal hubs for official conclusions or expectations. Decision Trees show dynamic cycles essentially and interpretably. decision Trees can help make sense of ransomware recognition results by uncovering the successive stages prompting a decision [52].

```

Decision Tree

#now train decision tree classifier with hyper parameters
dt_cls = DecisionTreeClassifier(criterion = 'entropy', max_leaf_nodes=100, random_state=1)
dt_cls.fit(X_train, y_train)
predict = dt_cls.predict(X_test)
calculateMetrics("Decision Tree", predict, y_test)

dt_acc = accuracy_score(predict, y_test)
dt_prec = precision_score(predict, y_test, average='macro')
dt_rec = recall_score(predict, y_test, average='macro')
dt_f1 = f1_score(predict, y_test, average='macro')

storeResults('Decision Tree', dt_acc, dt_prec, dt_rec, dt_f1)
    
```

Fig 7 Decision tree

**K – Nearest Neighbor (KNN):**

A supervised ML technique for grouping and relapse is KNN. It arranges approaching data of interest by greater part class vote or estimates their qualities by averaging their k closest neighbors in highlight space. KNN is possible utilized on the grounds that it catches neighborhood information designs without any problem. In ransomware identification, when slight changes might exist, KNN might find equivalent dataset designs.[63]

```

KNN

#now training KNN algorithm
knn_cls = KNeighborsClassifier(n_neighbors=500)
knn_cls.fit(X_train, y_train)
predict = knn_cls.predict(X_test)
calculateMetrics("KNN", predict, y_test)

knn_acc = accuracy_score(predict, y_test)
knn_prec = precision_score(predict, y_test, average='macro')
knn_rec = recall_score(predict, y_test, average='macro')
knn_f1 = f1_score(predict, y_test, average='macro')

storeResults('KNN', knn_acc, knn_prec, knn_rec, knn_f1)
    
```

Fig 8 KNN

**Support Vector Machine (SVM):**

A managed ML calculation for characterization and relapse is SVM. It boosts edge between classes by finding a hyperplane that best characterizes information or predicts a nonstop outcome. SVM is utilized to find ideal choice limits and handle high-layered information. SVM's express choice limits can help arrange malware in confounded highlight spaces [52].

### SVM

```
#now train SVM algorithm on training features and then test on testing features to calculate accuracy and other metrics
svm_cls = svm.SVC(kernel='poly', gamma='scale', C=0.004)
svm_cls.fit(X_train, y_train)
predict = svm_cls.predict(X_test)
calculateMetrics('SVM', predict, y_test)

|: svm_acc = accuracy_score(predict, y_test)
svm_prec = precision_score(predict, y_test, average='macro')
svm_rec = recall_score(predict, y_test, average='macro')
svm_f1 = f1_score(predict, y_test, average='macro')
storeResults('SVM', svm_acc, svm_prec, svm_rec, svm_f1)
```

Fig 9 SVM

### 2D Convolutional Neural Network (CNN2D):

CNN2D processes network like information for picture investigation utilizing deep learning. This venture adjusts CNN2D to constant information by utilizing convolutional layers to learn various leveled highlights and examples consequently. CNN2D processes ceaseless information to remove muddled attributes consequently. In ransomware identification, designs in consecutive and ceaseless framework movement are critical. CNN2D can catch itemized data, further developing the recognition model.

```
CNN
from sklearn.preprocessing import StandardScaler
X_train_scaled = StandardScaler().fit_transform(X_train)
X_test_scaled = StandardScaler().fit_transform(X_test)

#Define function for CNN model object
def CNN(x_train, x_test):
    #Building CNN layer with 32 filters to extract feature features using 2D conv
    conv_model = Conv2D(32, (3, 3), input_shape=(1, 1, 28, 28), activation='relu')
    #adding maxpooling layer to extract 2D feature features from previous 32 layer
    pool_model = MaxPooling2D(pool_size=(2, 2))
    #Building another CNN layer with 64 filters
    conv_model.add(Conv2D(64, (3, 3), activation='relu'))
    #adding another maxpooling layer
    pool_model.add(MaxPooling2D(pool_size=(2, 2)))
    #Flattening the output of previous layer
    flat_model = Flatten()
    #adding another fully connected layer
    fc_model = Dense(1000)
    #adding another fully connected layer with softmax activation
    output_model = Dense(10, activation='softmax')
    #combining the two fully connected layers
    model = flat_model(fc_model(output_model))
    #compiling the model with adam optimizer and categorical_crossentropy loss function
    model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
    #fitting the model on training data
    model.fit(X_train_scaled, y_train, batch_size=32, epochs=10, validation_data=(X_test_scaled, y_test), verbose=1)
    #predicting on test data
    predict = model.predict(X_test_scaled)
    return predict
```

Fig 10 CNN2D

### Voting Classifier:

Voting Classifiers use greater part casting a ballot or averaging to consolidate model forecasts. It utilizes fluctuated calculations to work on model execution. The Voting Classifier utilizes calculation expectations. This ensemble procedure increments ransomware discovery framework vigor and accuracy by consolidating a few model perspectives.

```
Voting Classifier
from sklearn.ensemble import RandomForestClassifier, VotingClassifier, AdaBoostClassifier
clf1 = AdaBoostClassifier(n_estimators=10, random_state=0)
clf2 = RandomForestClassifier(n_estimators=5, random_state=1)
vclf = VotingClassifier(estimators=[('ad', clf1), ('rf', clf2)], voting='soft')
vclf.fit(X_train, y_train)
predict = vclf.predict(X_test)
calculateMetrics('Voting Classifier', predict, y_test)

rf_acc = accuracy_score(predict, y_test)
rf_prec = precision_score(predict, y_test, average='macro')
rf_rec = recall_score(predict, y_test, average='macro')
rf_f1 = f1_score(predict, y_test, average='macro')
storeResults('Voting Classifier', rf_acc, rf_prec, rf_rec, rf_f1)
```

Fig 11 Voting classifier

## 4. EXPERIMENTAL RESULTS

**Precision:** Precision quantifies the percentage of certain events or tests that are well characterized. To attain accuracy, use the formula:

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} = \frac{TP}{(TP + FP)}$$



$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

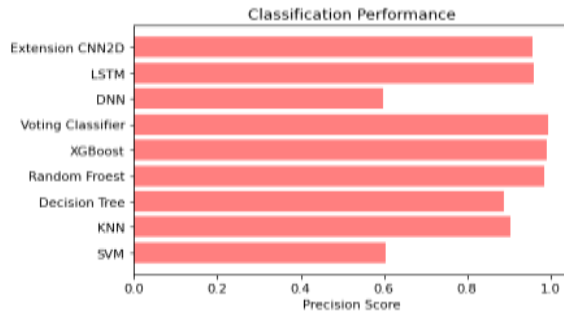


Fig 12 Precision comparison graph

**Recall:** ML recall measures a model's ability to catch all class occurrences. The model's ability to recognize a certain type of event is measured by the percentage of precisely anticipated positive prospects that turn into real earnings.[65]

$$\text{Recall} = \frac{TP}{TP + FN}$$

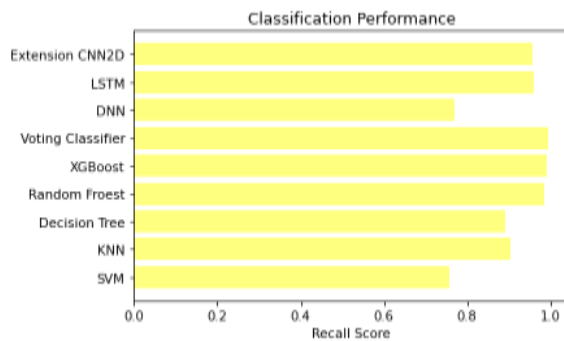


Fig 13 Recall comparison graph

**Accuracy:** The model's accuracy is the percentage of true predictions at a grouping position.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

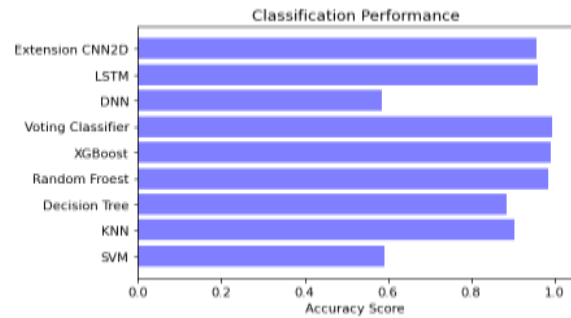


Fig 14 Accuracy graph

**F1 Score:** The F1 score captures both false positives and false negatives, making it a harmonized precision and validation technique for unbalanced data sets.

$$\text{F1 Score} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} * 100$$

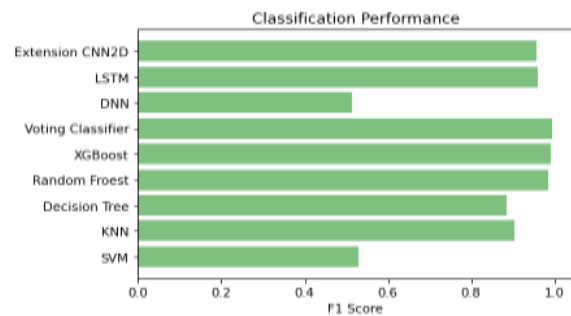


Fig 15 F1Score

ML Model	Accuracy	Precision	Recall	F1 score
SVM	0.60	0.60	0.75	0.67
KNN	0.90	0.90	0.90	0.90
Decision Tree	0.90	0.90	0.90	0.90
Random Forest	0.98	0.98	0.98	0.98
XGBoost	0.98	0.98	0.98	0.98
Extension Voting Classifier	0.98	0.98	0.98	0.98
DNN	0.60	0.60	0.75	0.67
LSTM	0.95	0.95	0.95	0.95
Extension CNN2D	0.95	0.95	0.95	0.95

Fig 16 Performance Evaluation



Fig 17 Home page



Fig 19 Login page



Fig 18 Signin page

rd\_total\_times

wr\_total\_times

flush\_total\_times

Predict

Fig 20 User input

Prediction Result: **Benign!**

Fig 21 Predict result for given input

## 5. CONCLUSION

The venture utilizes virtualization innovations, equipment execution counters, and IO occasions information to improve ransomware identification exactness while diminishing framework execution [9], [10], [11], [12], [13], [14]. The exploration broadly tests SVM, KNN, DT, RF, XGBOOST, DNN, and LSTM, finding that RF and XGBOOST [52] dependably anticipate ransomware action. The review thinks about DL models like DNN and LSTM against exemplary ML calculations to increment forecast technique assortment. A freely open dataset from different projects advances joint effort and permits scholastics to evaluate their ransomware discovery strategies. The task flawlessly coordinates Flask for web structure and SQLite for client enlistment and verification, giving an easy to use interface for information section, preprocessing, and model expectations, working on functional materialness.

## 6. FUTURE SCOPE

Since the current review zeroed in on a blend of known and obscure ransomware, further exploration could assess the recommended strategy's capacity to distinguish novel and arising varieties. Expanding on the review's adaptability to various responsibilities, analyzing how client jobs influence ransomware location [7] could yield critical data. With almost 100% ransomware recognition accuracy, the undertaking's democratic classifier has performed well. Front-end highlight esteem testing demonstrates its unwavering quality in distinguishing and alleviating ransomware assaults. Genuine testing of the proposed procedure would uncover its adequacy in recognizing ransomware assaults in live creation frameworks. The undertaking could improve

ransomware discovery [24, 25, 34] by adding information sources or qualities to the ML model. Cooperation with network protection trained professionals and associations takes into account approval and move toward refinement in view of real-world experience.

## REFERENCES

- [1] SR Department. (2022). Ransomware victimization rate 2022. Accessed: Apr. 6, 2022. [Online]. Available: <https://www.statista.com/statistics/204457/businesses-ransomware-attack-rate/>
- [2] D. Braue. (2022). Ransomware Damage Costs. Accessed: Sep. 16, 2022. [Online]. Available: <https://cybersecurityventures.com/globalransomware-damage-costs-predicted-to-reach-250-billion-usd-by-2031/>
- [3] Logix Consulting. (2020). What is Signature Based Malware Detection. Accessed: Apr. 3, 2023. [Online]. Available: <https://www.logixconsulting.com/2020/12/15/what-is-signature-based-malware-detection/>
- [4] W. Liu, P. Ren, K. Liu, and H.-X. Duan, "Behavior-based malware analysis and detection," in Proc. 1st Int. Workshop Complex. Data Mining, Sep. 2011, pp. 39–42.
- [5] (2021). Polymorphic Malware. Accessed: Apr. 3, 2023. [Online]. Available: <https://www.thesslstore.com/blog/polymorphic-malware-andmetamorphic-malware-what-you-need-to-know/>
- [6] M. Loman. (2021). Lockfile Ransomware's Box of Tricks: Intermittent Encryption and Evasion.

Accessed: Nov. 16, 2021. [Online]. Available: <https://news.sophos.com/en-us/2021/08/27/lockfile-ransomwares-box-oftricks-intermittent-encryption-and-evasion/>

[7] N. Pundir, M. Tehranipoor, and F. Rahman, "RanStop: A hardwareassisted runtime crypto-ransomware detection technique," 2020, arXiv:2011.12248.

[8] S. Mehnaz, A. Mudgerikar, and E. Bertino, "RWGuard: A real-time detection system against cryptographic ransomware," in Proc. Int. Symp. Res. Attacks, Intrusions, Defenses. Cham, Switzerland: Springer, 2018, pp. 114–136.

[9] J. Demme, M. Maycock, J. Schmitz, A. Tang, A. Waksman, S. Sethumadhavan, and S. Stolfo, "On the feasibility of online malware detection with performance counters," ACM SIGARCH Comput. Archit. News, vol. 41, no. 3, pp. 559–570, Jun. 2013.

[10] A. Tang, S. Sethumadhavan, and S. J. Stolfo, "Unsupervised anomalybased malware detection using hardware features," in Proc. Int. Workshop Recent Adv. Intrusion Detection. Cham, Switzerland: Springer, 2014, pp. 109–129.

[11] S. Das, J. Werner, M. Antonakakis, M. Polychronakis, and F. Monroe, "SoK: The challenges, pitfalls, and perils of using hardware performance counters for security," in Proc. IEEE Symp. Secur. Privacy (SP), May 2019, pp. 20–38.

[12] S. P. Kadiyala, P. Jadhav, S.-K. Lam, and T. Srikanthan, "Hardware performance counter-based fine-grained malware detection," ACM Trans. Embedded Comput. Syst., vol. 19, no. 5, pp. 1–17, Sep. 2020.

[13] B. Zhou, A. Gupta, R. Jahanshahi, M. Egele, and A. Joshi, "Hardware performance counters can detect malware: Myth or fact?" in Proc. Asia Conf. Comput. Commun. Secur., May 2018, pp. 457–468.

[14] S. Aurangzeb, R. N. B. Rais, M. Aleem, M. A. Islam, and M. A. Iqbal, "On the classification of microsoft-windows ransomware using hardware profile," PeerJ Comput. Sci., vol. 7, p. e361, Feb. 2021.

[15] M. Alam, S. Bhattacharya, S. Dutta, S. Sinha, D. Mukhopadhyay, and A. Chattopadhyay, "RATAFIA: Ransomware analysis using time and frequency informed autoencoders," in Proc. IEEE Int. Symp. Hardw. Oriented Secur. Trust (HOST), May 2019, pp. 218–227.

[16] K. Thummapudi, R. Boppana, and P. Lama, "HPC 41 events 5 rounds," Harvard Dataverse, 2022, doi: 10.7910/DVN/MA5UPP.

[17] K. Thummapudi, R. Boppana, and P. Lama, "IO 41 events 5 rounds," Harvard Dataverse, 2022, doi: 10.7910/DVN/GHJFUT.

[18] K. Thummapudi, R. Boppana, and P. Lama, "HPC 5 events 7 rounds," Harvard Dataverse, 2022, doi: 10.7910/DVN/YAYW0J.

[19] K. Thummapudi, R. Boppana, and P. Lama, "Io 5 events 7 rounds," Harvard Dataverse, 2022, doi: 10.7910/DVN/R9FYPL.

[20] K. Thummapudi, R. Boppana, and P. Lama, "Scripts to reproduce results," Harvard Dataverse, 2023, doi: 10.7910/DVN/HSX6CS.

[21] M. Rhode, P. Burnap, and A. Wedgbury, "Real-time malware process detection and automated process

killing,” *Secur. Commun. Netw.*, vol. 2021, pp. 1–23, Dec. 2021.

[22] A. Kharraz and E. Kirida, “Redemption: Real-time protection against ransomware at end-hosts,” in *Proc. Int. Symp. Res. Attacks, Intrusions, Defenses*. Cham, Switzerland: Springer, 2017, pp. 98–119.

[23] F. Mbol, J.-M. Robert, and A. Sadighian, “An efficient approach to detect torrentlocker ransomware in computer systems,” in *Proc. Int. Conf. Cryptol. Netw. Secur.* Springer, 2016, pp. 532–541.

[24] K. Lee, S. Lee, and K. Yim, “Machine learning based file entropy analysis for ransomware detection in backup systems,” *IEEE Access*, vol. 7, pp. 110205–110215, 2019.

[25] C. J. Chew and V. Kumar, “Behaviour based ransomware detection,” in *Proc. Int. Conf. Comput. Their Appl.*, in *EPiC Series in Computing*, vol. 58. 2019, pp. 127–136.

[26] S. Homayoun, A. Dehghantanha, M. Ahmadzadeh, S. Hashemi, and R. Khayami, “Know abnormal, find evil: Frequent pattern mining for ransomware threat hunting and intelligence,” *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 2, pp. 341–351, Apr. 2020.

[27] A. Kharaz, S. Arshad, C. Mulliner, W. Robertson, and E. Kirida, “UNVEIL: A large-scale, automated approach to detecting ransomware (keynote),” in *Proc. IEEE 24th Int. Conf. Softw. Anal., Evol. Reengineering (SANER)*, Feb. 2017, pp. 757–772.

[28] A. Kharraz, W. Robertson, D. Balzarotti, L. Bilge, and E. Kirida, “Cutting the gordian knot: A look under the hood of ransomware attacks,” in *Proc. Int. Conf. Detection Intrusions Malware, Vulnerability*

*Assessment*. Cham, Switzerland: Springer, 2015, pp. 3–24.

[29] A. Continella, A. Guagnelli, G. Zingaro, G. De Pasquale, A. Barengi, S. Zanero, and F. Maggi, “ShieldFS: A self-healing, ransomware-aware filesystem,” in *Proc. 32nd Annu. Conf. Comput. Secur. Appl.*, Dec. 2016, pp. 336–347.

[30] M. Shukla, S. Mondal, and S. Lodha, “POSTER: Locally virtualized environment for mitigating ransomware threat,” in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2016, pp. 1784–1786.

[31] N. Scaife, H. Carter, P. Traynor, and K. R. B. Butler, “CryptoLock (and drop it): Stopping ransomware attacks on user data,” in *Proc. IEEE 36th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2016, pp. 303–312.

[32] D. Sgandurra, L. Muñoz-González, R. Mohsen, and E. C. Lupu, “Automated dynamic analysis of ransomware: Benefits, limitations and use for detection,” 2016, arXiv:1609.03020.

[33] P. Zavarsky and D. Lindskog, “Experimental analysis of ransomware on windows and Android platforms: Evolution and characterization,” *Proc. Comput. Sci.*, vol. 94, pp. 465–472, Jan. 2016.

[34] T. McIntosh, J. Jang-Jaccard, P. Watters, and T. Susnjak, “The inadequacy of entropy-based ransomware detection,” in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2019, pp. 181–189.

[35] Z. A. Genc, G. Lenzini, and D. Sgandurra, “On deception-based protection against cryptographic ransomware,” in *Proc. Int. Conf. Detection Intrusions*

Malware, Vulnerability Assessment Cham, Switzerland: Springer, 2019, pp. 219–239.

[36] S. Song, B. Kim, and S. Lee, “The effective ransomware prevention technique using process monitoring on Android platform,” *Mobile Inf. Syst.*, vol. 2016, pp. 1–9, Mar. 2016.

[37] K. Cabaj, M. Gregorczyk, and W. Mazurczyk, “Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics,” *Comput. Electr. Eng.*, vol. 66, pp. 353–368, Feb. 2018.

[38] R. Moussaileb, B. Bouget, A. Palisse, H. Le Boudier, N. Cuppens, and J.-L. Lanet, “Ransomware’s early mitigation mechanisms,” in *Proc. 13th Int. Conf. Availability, Rel. Secur.*, 2018, pp. 1–10.

[39] Z. A. Genc, G. Lenzini, and P. Y. Ryan, “No random, no ransom: A key to stop cryptographic ransomware,” in *Proc. Int. Conf. Detection Intrusions Malware, Vulnerability Assessment*. Cham, Switzerland: Springer, 2018, pp. 234–255.

[40] M. M. Ahmadian, H. R. Shahriari, and S. M. Ghaffarian, “Connectionmonitor & connection-breaker: A novel approach for prevention and detection of high survivable ransoms,” in *Proc. 12th Int. Iranian Soc. Cryptol. Conf. Inf. Secur. Cryptol. (ISCISC)*, Sep. 2015, pp. 79–84.

[41] E. Kolodenker, W. Koch, G. Stringhini, and M. Egele, “PayBreak: Defense against cryptographic ransomware,” in *Proc. ACM Asia Conf. Comput. Commun. Secur.*, Apr. 2017, pp. 599–611.

[42] M. S. Kiraz, Z. A. Genc, and E. Ozturk, “Detecting large integer arithmetic for defense against

crypto ransomware,” *Cryptology ePrint Arch., Rep.*, vol. 558, p. 2017, Jan. 2017.

[43] (2022). What Systems Have You Seen Infected by Ransomware? Accessed: Apr. 3, 2023. [Online]. Available: <https://www.statista.com/statistics/701020/major-operating-systems-targeted-by-ransomware/>

[44] (2022). Linux Profiling With Performance Counters. Accessed: Apr. 3, 2023. [Online]. Available: <https://perf.wiki.kernel.org/index.php/MainPage>

[45] (2022). Likwid Performance Tools. Accessed: Apr. 3, 2023. [Online]. Available: <https://hpc.fau.de/research/tools/likwid/>

[46] K. Keahey, J. Anderson, Z. Zhen, P. Riteau, P. Ruth, D. Stanzione, M. Cevik, J. Colleran, H. S. Gunawi, C. Hammock, J. Mambretti, A. Barnes, F. Halbach, A. Rocha, and J. Stubbs, “Lessons learned from the chameleon testbed,” in *Proc. USENIX Annu. Tech. Conf.*, Jul. 2020, pp. 219–233.

[47] Alexa Top Websites. Accessed: Sep. 13, 2021. [Online]. Available: <https://www.alexa.com/topsites>

[48] Ninite. Accessed: Apr. 3, 2023. [Online]. Available: <https://ninite.com>

[49] API. (2023). Libvirt: Virsh Tool Manual. Accessed: Apr. 3, 2023. [Online]. Available: <https://libvirt.org/manpages/virsh.html>

[50] (2021). Virusshare. Accessed: Nov. 16, 2021. [Online]. Available: <https://virusshare.com>

[51] Intel. (2023). System Programming Guide. Accessed: Apr. 3, 2023. [Online]. Available:

<https://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-softwaredeveloper-vol-3b-part-2-manual.pdf>

[52] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

[53] J. Benesty, J. Chen, Y. Huang, and I. Cohen, “Pearson correlation coefficient,” in *Noise Reduction in Speech Processing*. Berlin, Germany: Springer, 2009, pp. 1–4.

[54] H. Jin, Q. Song, and X. Hu, “Auto-Keras: An efficient neural architecture search system,” in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 1946–1956.

[55] Wikipedia. (2021). Sensitivity and Specificity. Accessed: Apr. 3, 2023. [Online]. Available: <https://en.wikipedia.org/wiki/Sensitivityandspecificity>

[56] Hat Enterprise. (2023). Block I/O Tuning. [Online]. Available: <https://access.redhat.com/documentation/en-us/redhatenterpriselinux/7/html/virtualizationtuningandoptimizationguide/sect-virtualizationtuningoptimizationguide-blockio-techniques>

[57] G. Viswanath, “Hybrid encryption framework for securing big data storage in multi-cloud environment”, *Evolutionary intelligence*, vol.14, 2021, pp.691-698.

[58] Viswanath Gudditi, “Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage”, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol.12, 2021, pp.545-552.

[59] Viswanath Gudditi, “A Smart Recommendation System for Medicine using Intelligent NLP Techniques”, 2022 *International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 2022, pp.1081-1084.

[60] G.Viswanath, “Enhancing power unbiased cooperative media access control protocol in manets”, *International Journal of Engineering Inventions*, 2014, vol.4, pp.8-12.

[61] Viswanath G, “A Hybrid Particle Swarm Optimization and C4.5 for Network Intrusion Detection and Prevention System”, 2024, *International Journal of Computing*, DOI: <https://doi.org/10.47839/ijc.23.1.3442>, vol.23, 2024, pp.109-115.

[62] G.Viswanath, “A Real Time online Food Ording application based DJANGO Restfull Framework”, *Juni Khyat*, vol.13, 2023, pp.154-162.

[63] Gudditi Viswanath, “Distributed Utility-Based Energy Efficient Cooperative Medium Access Control in MANETS”, 2014, *International Journal of Engineering Inventions*, vol.4, pp.08-12.

[64] G.Viswanath,“ A Real-Time Video Based Vehicle Classification, Detection And Counting System”, 2023, *Industrial Engineering Journal*, vol.52, pp.474-480.

[65] G.Viswanath, “A Real- Time Case Scenario Based On Url Phishing Detection Through Login Urls

”, 2023, Material Science Technology, vol.22, pp.103-108.

[66] Manmohan Singh, Susheel Kumar Tiwari, G. Swapna, Kirti Verma, Vikas Prasad, Vinod Patidar, Dharmendra Sharma and Hemant Mewada, “A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification” published in Journal of Computer Science, Available at: <https://pdfs.semanticscholar.org/69ac/f07f2e756b79181e4f1e75f9e0f275a56b8e.pdf>