IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

Impact Factor: 7.301

Volume 3, Issue 6, April 2023

Using MLT to Anticipate for Thyroid Sickness

P. S. L. Sravani[1], K. Ramya[2], K. Rajani[3], K. Vani Varshini[4], B. Sai Manoj Kumar[5]

Assistant Professor, Department of Computer Science and Engineering[1]

Students, Department of Computer Science and Engineering[2,3,4,5]

Raghu Institute of Technology, Visakhapatnam, AP, India

Abstract: *The idea that thyroid disorders are the primary factor in medical diagnosis and prediction is a complicated premise in the medical field. One of our body's most important organs is the thyroid. Thyroid hormones are released and regulate metabolism. The body's capacity to regulate its metabolism is impacted by both thyroid hormone overproduction and underproduction. It is essential to use machine learning in the prediction of illnesses and the investigation of classification models for thyroid disease based on hospital dataset data. A good knowledge base in the form of a hybrid model is necessary to deal with dynamic learning activities like medical diagnosis and prediction. Thyroid might be identified and repressed utilizing straightforward AI draws near. Using an SVM model to predict the likelihood of a thyroid patient is common practice. Whenever a patient is in danger of creating thyroid illness, our framework should propose home cures, alerts and medication.*

Keywords: Machine learning, Classification algorithms, Decision trees, KNN, K-means, ANN

I. INTRODUCTION

Thyroid disorders today disrupt the normal function of the thyroid gland, resulting in abnormal hormone production and hyperthyroidism[1][3]. In the developed world, hypothyroidism is thought to affect about 4-5% of people.If left untreated, hypothyroidism can lead to high cholesterol, high blood pressure, cardiovascular problems, reduced fertility, and depression. Expert advisory systems (EAS) are made possible by computer science professionals and medical science technology [4][7] in order to accurately diagnose a variety of diseases. The clinical experts are made to utilize these frameworks because of a few created blunders during general conclusion process [5]. Illness determination activities utilizing EAS are performed in light of sets of illness side effects. These frameworks depend on AI strategy which assists the doctor with limiting the expenses and time in successful analyses. This work is a half breed engineering configuration outfitted effectively utilizing machine learning procedures. This work aims to develop an efficient and spotless method for identifying thyroid disease in humans[9][10]. According to my literature review, the data considered for thyroid disease diagnoses (TDD) is inconsistent, redundant, and contains missing attribute values. However, there are several mechanisms implemented on thyroid data sets that produced surprising results. An expert advisory system based on hybrid architecture is being proposed for the purpose of determining the thyroid gland's positive disease growth. A string matching framework (SMS) was at the beginning created, which can foresee the genuine TDD in view of the information accessible [12] [13].

II. SYSTEM ANALYSIS

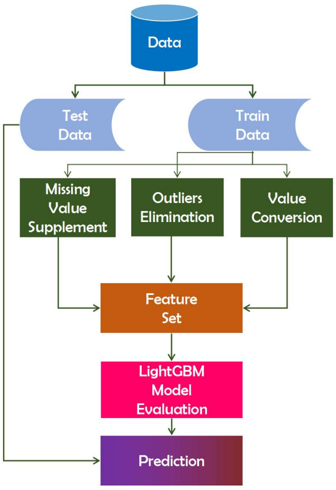
2.1 Existing Analysis

In light of this circumstance, it makes sense to employ machine learning-based techniques that make use of prior experience and knowledge to automatically classify malware by performing static code analysis on unknown binary code. As per the direction, this paper utilizes the connected advances of AI based techniques and investigates the utilization of this strategy in the order of malware

2.2 Proposed System

Working first with cascade one-sided perceptron’s and then with cascade kernelized one-sided perceptron’s, we present the concepts that underpin our framework in this paper. The concepts behind this framework were put through a scaling process that enables us to work with very large datasets of malware and clean files after being tested successfully on medium-sized datasets of malware and clean files.

DOI: 10.48175/IJARSCT-9428



Application of machine learning algorithms to predict the thyroid disease risk

2.3 Algorithm

Support Vector Machine, also known as SVM, is one of the most widely used supervised learning algorithms. It can be used to solve regression and classification problems. Nonetheless, basically, it is utilized for Grouping issues in AI.The SVM algorithm's objective is to find the most effective line or decision boundary for classifying n-dimensional space, allowing us to quickly place a new data point in the appropriate category in the future. A hyperplane is the name given to this best decision boundary.SVM picks the outrageous focuses/vectors that assistance in making the hyperplane. The algorithm is referred to as a Support Vector Machine because these extreme cases are referred to as support vectors. Take a look at the diagram below, which shows two distinct categories that are separated by a decision boundary or hyperplane

In recent years, a lot of work has been done to diagnose specific thyroid diseases. Many creatorshave utilized different sorts of information mining method. By using a variety of datasets and algorithms related to the work that needs to be done in the future in order to achieve more effective and superior results, the authors demonstrated that they could find diseases that are comparable to the thyroid with sufficient certainty and an adequate method. The purpose of this paper is to interpret a variety of data mining mechanisms and statistical attributes that have become increasingly popular in recent years for interpreting thyroid diseases with certainty by a variety of authors to achieve a variety of outcomes and approaches. Random forest, decision tree, naive Bayes, SVM, and ANN are just a few of the machine learning algorithms that are frequently utilized in prognostic and common diseases. Although the development of a machine learning-placed disease prediction mechanism and a medical determination is a nontrivial task, there are essential issues, such as the acquisition of data, compilation, and grouping, that are used to train the machine learning structures. These issues include diseases related to heart disease, diabetes, Parkinson's disease, hypertension, the Ebola virus (EV) , diagnoses and forecasting, R-NA sequenced data analysis, and allocation of biomedical imaging Estimation of large biomedical data sets over a long period of time is desired but practically nonexistent in actual activity issues .A methodical strategy for using a neural network's back propagation algorithm to make an earlier diagnosis of thyroid disease. Back propagation of an error that is being used for prior disease predictions is delicate and established by ANN. The impact of ANN is that it is trained using empirical data and tested using data that was not used during the training process, proving its validity. The advanced neural network, which uses as a substitute for prior disease predictions, concludes in good agreement with the initial data.

In the creators examined and compare the four grouping models in particular Credulous Bayes, Choice Tree,Multi-facet Perceptron and Outspread Premise Capability Network. The conclusion demonstrates remarkable classification model accuracy. The other classification models are inferior to the Decision Tree model. In this work, 29 dataset attributes are conscripted and enforced using the Feature Selection technique Chi-Square. The datasets are filtered using unsupervised coated filters to convert the attributes' continuous values into nominal values, reducing the number of attributes from 29 to 10.AI (ML) is a division of fake knowledge and is penetrated in the elements oflogical exploration

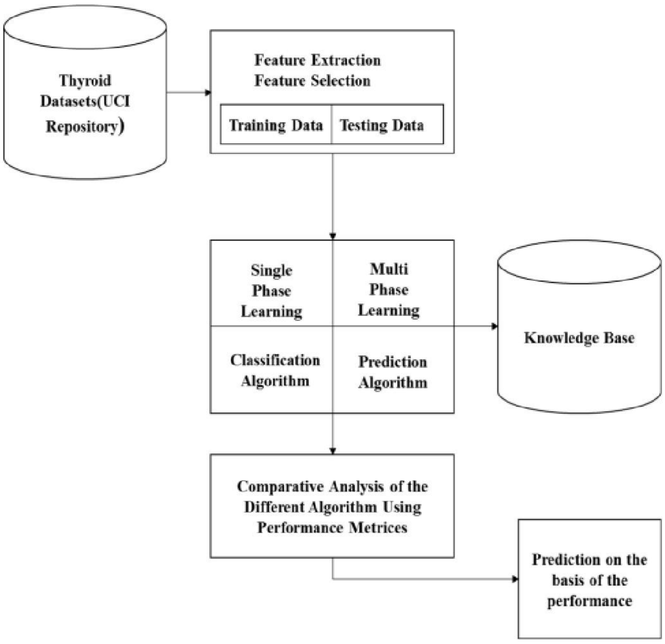
DOI: 10.48175/IJARSCT-9428

at developing advances. Algorithms are able to review from experience more easily thanks to machine learning. Classical epidemiology is an advanced blended recent data science approach to strap the capabilities of cultured data and has been induced by the input explosion that is connected with an expanding computational capability.The particular tools look into nearby clinically relevant connections between input and output criteria in order to take into account large data sets. To alter surgical agreements, factual analyses of surgical conclusions are highly deceptive. The description of the patient's companion, which aids from surgery in the arbitration, is one of the crucial aspects of surgical agreements. Computers can learn from previous data to make precise predictions based on current data using machine learning. The informative aspect produces highly authoritative prediction algorithms that are able to replicate the formerly exotic communication in vast, confusing data sets and adapt to effective data aura .

A favorable framework for the creation of machine learning models is provided by the composite characteristics and curative procedures used to treat thyroid disorders, which provide ample clustering of intricate and diverse data . This suggests that machine learning models could be used in a lot of situations and points to a growing trend toward rigorous treatments that are tailored to each patient. In the field of AI, a broad disparity could becreated in the midst of administered and solo learning.

Supervised learning algorithms crop a model that makes predictions on previously fictitious data from "labelled" training data. Unsupervised learning algorithms may take in a large number of unlabelled genomics data as input and analyze previously anonymous assemblages of data. For unsupervised mechanism of learning, only unlabelled data are possible, and the algorithms look for analogies and devices. These algorithms could be used to create labels in order to eventually train a supervised model. They could also be dominant in previously formerly arrangements in complex data that are not primarily measurable by humans [9]. To generate a desired output from a given set of input variables, a programmer manually creates a set of information known as "the programs" in conventional programming. InAI, the data sources are furnished along withthe hunger for result and PC calculations are asked todetermine the "rules from the arranged preparation information". Computerized learning is a good way to interpret a lot of data, design hidden communications in composite data sets, and adapt to dynamic atmosphere.

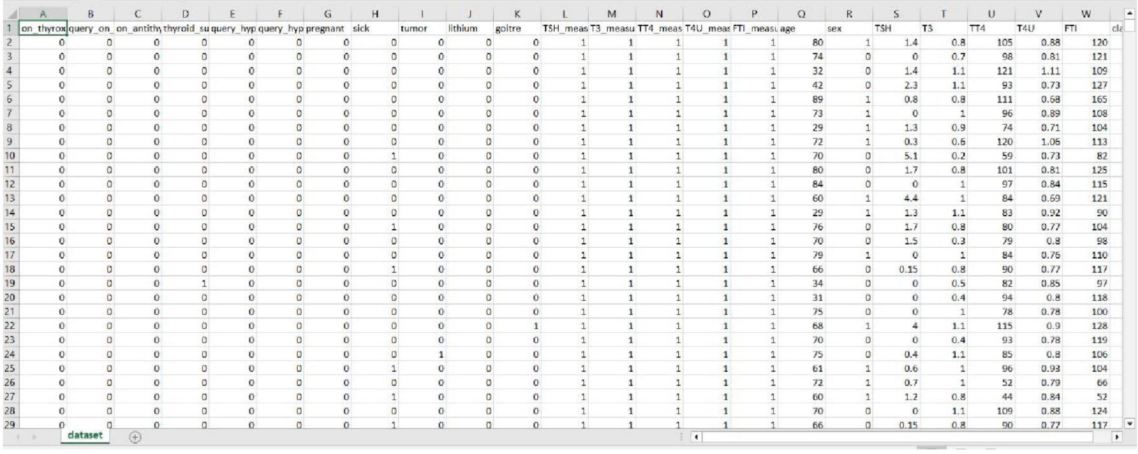
Algorithms in the learning mechanism strive to achieve an excellent aggregate of input variables (features), and weights are added to these features in the model to reduce the gap between expected and significant results. The enforced machine learning techniques are recycled to develop abstraction devices or frame a model and use the accomplished devices or models in making predictions in the future for anonymous cases. Machine learning is used to train the system over vast databases.



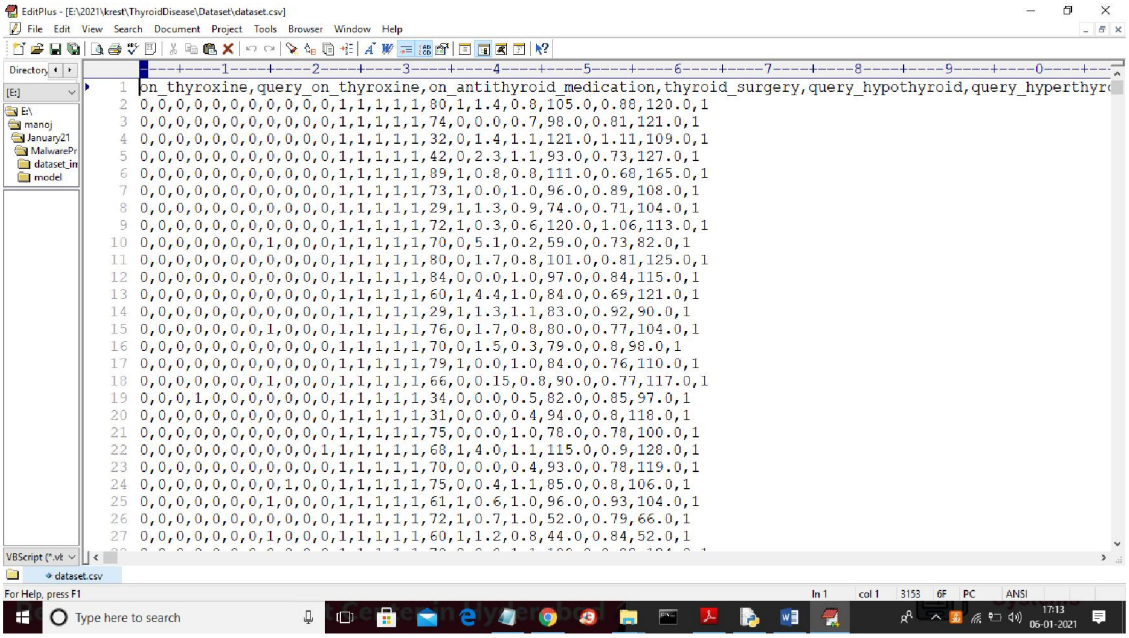
ARCHITECTURE OF THYRIOD PREDICTION SYSTEM

DOI: 10.48175/IJARSCT-9428

2.4 Sample Dataset



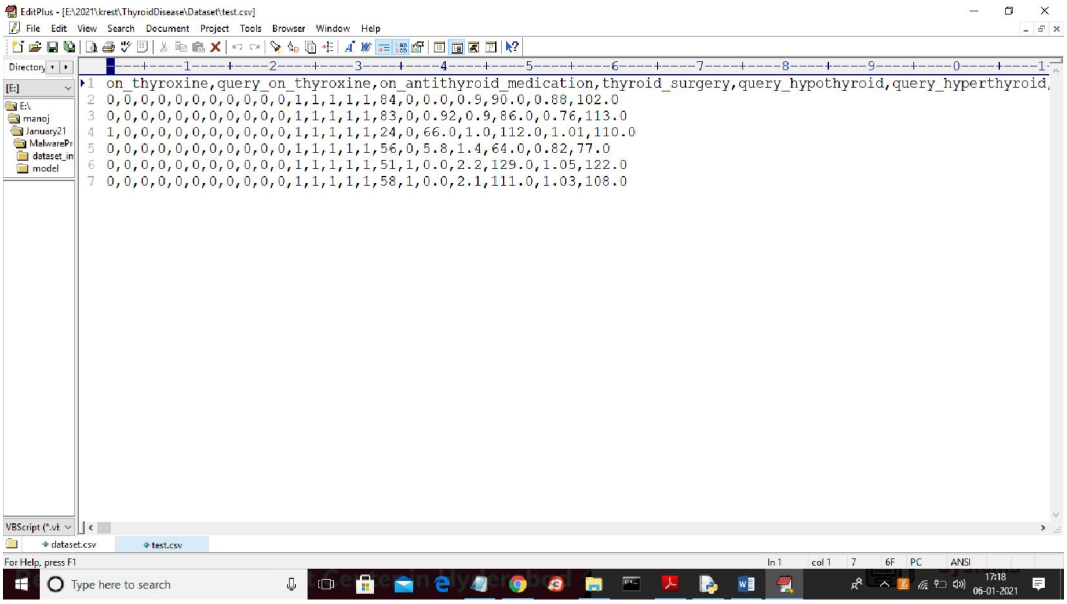
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	on_thyrox	query_on_	on_antithy	thyroid_su	query_hyp	query_hyp	pregnant	sick	tumor	lithium	goitre	TSH_meas	T3_measu	TT4_meas	T4U_meas	FTI_measu	age	sex	TSH	T3	TT4	T4U	FTI	cla
2	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	80	1	1.4	0.8	105	0.88	120	
3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	74	0	0	0.7	98	0.81	121	
4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	32	0	1.4	1.1	121	1.11	109	
5	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	42	0	2.3	1.1	93	0.73	127	
6	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	89	1	0.8	0.8	111	0.68	165	
7	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	73	1	0	1	96	0.89	108	
8	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	29	1	1.3	0.9	74	0.71	104	
9	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	72	1	0.3	0.6	120	1.06	113	
10	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	70	0	5.1	0.2	59	0.73	82	
11	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	80	0	1.7	0.8	101	0.81	125	
12	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	84	0	0	1	97	0.84	115	
13	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	60	1	4.4	1	84	0.69	121	
14	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	29	1	1.3	1.1	83	0.92	90	
15	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	76	0	1.7	0.8	80	0.77	104	
16	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	70	0	1.5	0.3	79	0.8	98	
17	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	79	1	0	1	84	0.76	110	
18	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	66	0	0.15	0.8	90	0.77	117	
19	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	34	0	0	0.5	82	0.85	97	
20	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	31	0	0	0.4	94	0.8	118	
21	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	75	0	0	1	78	0.78	100	
22	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	68	1	4	1.1	115	0.9	128	
23	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	70	0	0	0.4	93	0.78	119	
24	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	75	0	0.4	1.1	85	0.8	106	
25	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	61	1	0.6	1	96	0.93	104	
26	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	72	1	0.7	1	52	0.79	66	
27	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	60	1	1.2	0.8	44	0.84	52	
28	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	70	0	0	1.1	109	0.88	124	
29	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	66	0	0.15	0.8	90	0.77	117	



```
on_thyroxine,query_on_thyroxine,on_antithyroid_medication,thyroid_surgery,query_hypothyroid,query_hyperthyr
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,80,1,1.4,0.8,105.0,0.88,120.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,74,0,0.0,0.7,98.0,0.81,121.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,32,0,1.4,1.1,121.0,1.11,109.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,42,0,2.3,1.1,93.0,0.73,127.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,89,1,0.8,0.8,111.0,0.68,165.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,73,1,0.0,1.0,96.0,0.89,108.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,29,1,1.3,0.9,74.0,0.71,104.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,72,1,0.3,0.6,120.0,1.06,113.0,1
0,0,0,0,0,0,0,1,0,0,0,1,1,1,1,1,70,0,5.1,0.2,59.0,0.73,82.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,80,0,1.7,0.8,101.0,0.81,125.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,84,0,0.0,1.0,97.0,0.84,115.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,60,1,4.4,1.0,84.0,0.69,121.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,29,1,1.3,1.1,83.0,0.92,90.0,1
0,0,0,0,0,0,0,1,0,0,0,1,1,1,1,1,76,0,1.7,0.8,80.0,0.77,104.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,70,0,1.5,0.3,79.0,0.8,98.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,79,1,0.0,1.0,84.0,0.76,110.0,1
0,0,0,0,0,0,0,1,0,0,0,1,1,1,1,1,66,0,0.15,0.8,90.0,0.77,117.0,1
0,0,0,1,0,0,0,0,0,0,0,1,1,1,1,1,34,0,0.0,0.5,82.0,0.85,97.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,31,0,0.0,0.4,94.0,0.8,118.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,75,0,0.0,1.0,78.0,0.78,100.0,1
0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,68,1,4.0,1.1,115.0,0.9,128.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,70,0,0.0,0.4,93.0,0.78,119.0,1
0,0,0,0,0,0,0,0,1,0,0,1,1,1,1,1,75,0,0.4,1.1,85.0,0.8,106.0,1
0,0,0,0,0,0,0,1,0,0,0,1,1,1,1,1,61,1,0.6,1.0,96.0,0.93,104.0,1
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,72,1,0.7,1.0,52.0,0.79,66.0,1
0,0,0,0,0,0,0,1,0,0,0,1,1,1,1,1,60,1,1.2,0.8,44.0,0.84,52.0,1
```

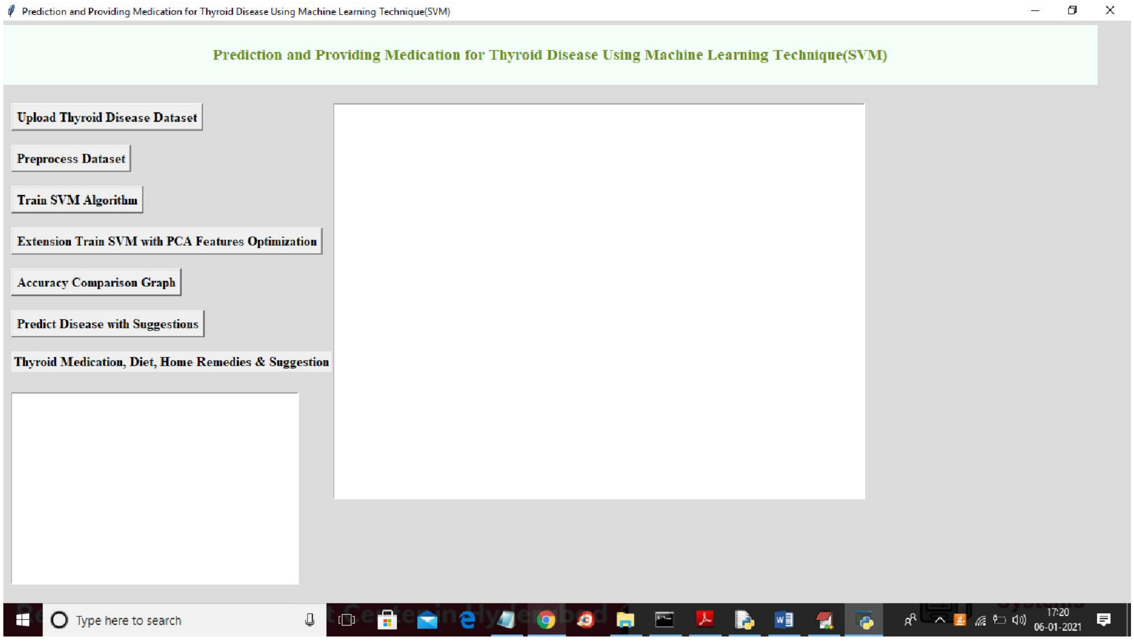
In above dataset first line contains section names and different columns contains values as 0 or 1 and on the off chance that patient is under thyroid drug or medical procedure, its segment worth will be 1 else 0 and in last section contains class mark as 0 or 1 where 0 methods patient record is typical and 1 method patient record contains thyroid sickness. There are more than 3000 rows and 24 columns in this dataset. We are using the PCA (principal component analysis) feature selection algorithm as an extension concept to optimize features or to reduce columns or features that are not important for prediction because all 24 columns are not available for prediction. In order to train the SVM algorithm, PCA will eliminate all unnecessary columns from the dataset and only use important attributes. As a result of optimizing features, SVM prediction accuracy can be improved.

DOI: 10.48175/IJARSCT-9428



EditPlus - [E:\2021\krest\ThyroidDisease\Dataset\test.csv]

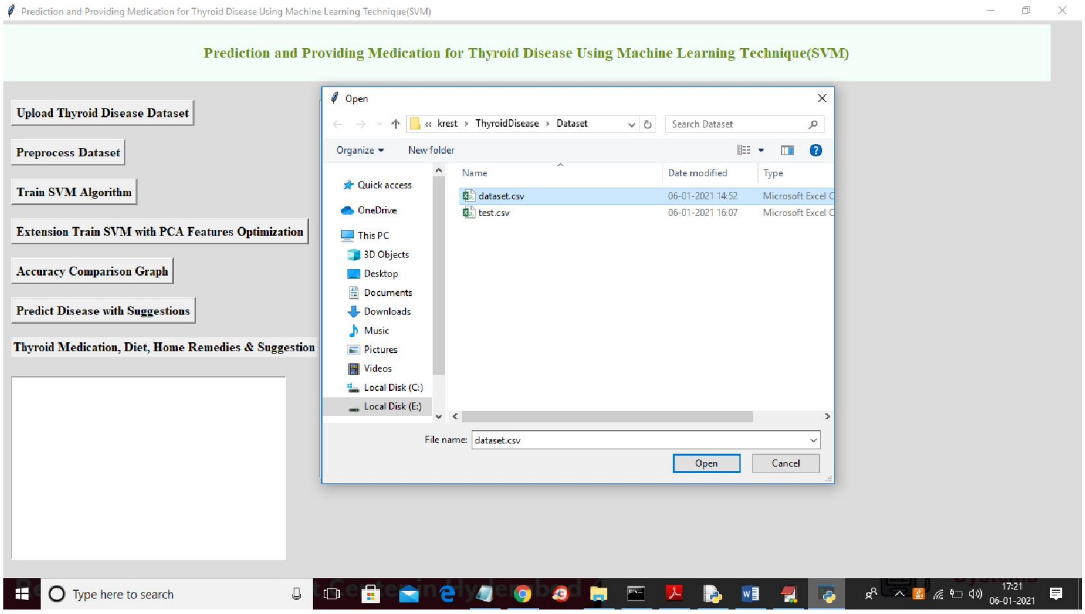
```
on_thyroxine,query_on_thyroxine,on_antithyroid_medication,thyroid_surgery,query_hypothyroid,query_hyperthyroid,
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,84,0,0.0,0.9,90.0,0.88,102.0
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,83,0,0.92,0.9,86.0,0.76,113.0
1,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,24,0,66.0,1.0,112.0,1.01,110.0
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,56,0,5.8,1.4,64.0,0.82,77.0
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,51,1,0.0,2.2,129.0,1.05,122.0
0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,58,1,0.0,2.1,111.0,1.03,108.0
```

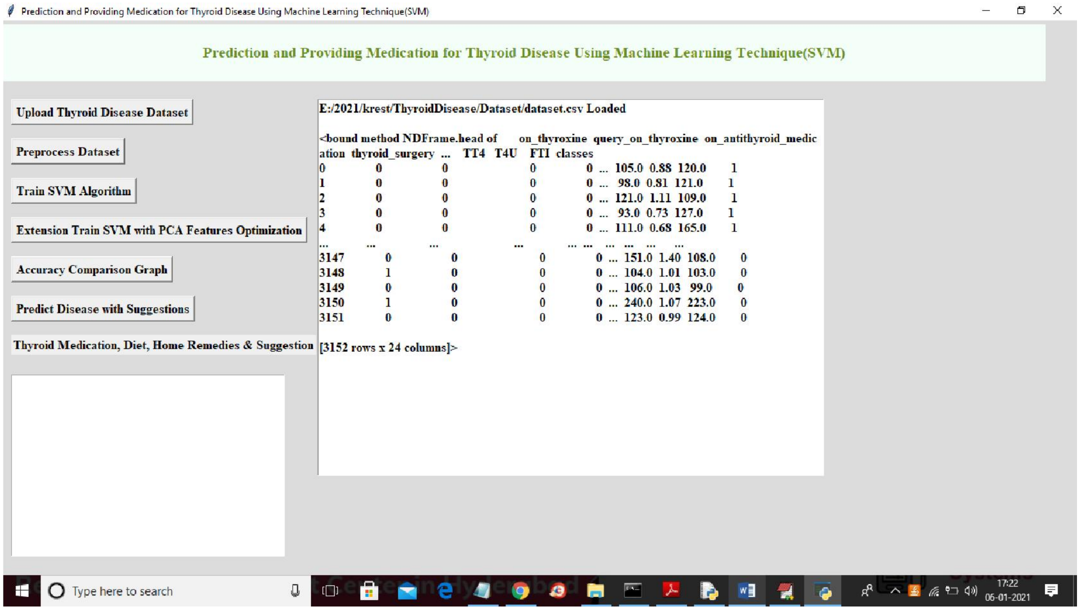


Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)

Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Type here to search
17:20
06-01-2021

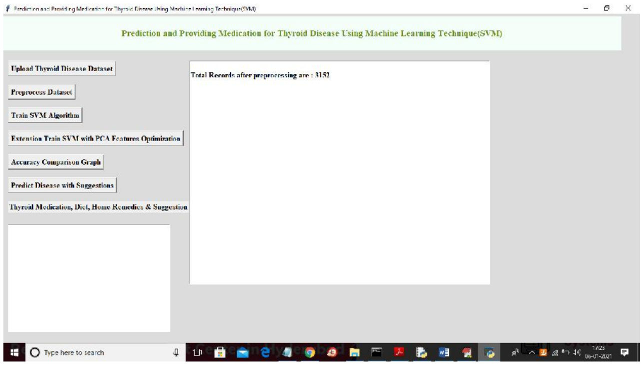
DOI: 10.48175/IJARSCT-9428


Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Open
« krest › ThyroidDisease › Dataset
Search Dataset
Organize ▾
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
Name
Date modified
Type
dataset.csv
06-01-2021 14:52
Microsoft Excel C
test.csv
06-01-2021 16:07
Microsoft Excel C
File name: dataset.csv
Open
Cancel
Type here to search
17:21
06-01-2021

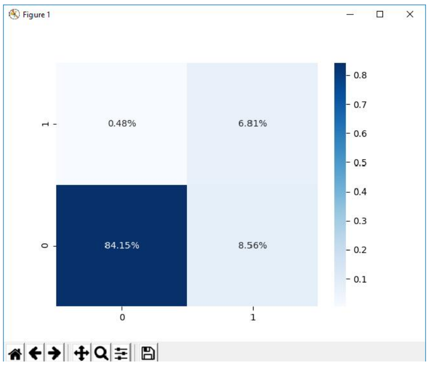

Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Type here to search
17:22
06-01-2021

```
E:/2021/krest/ThyroidDisease/Dataset/dataset.csv Loaded

<bound method NDFrame.head of      on_thyroxine  query_on_thyroxine  on_antithyroid_medic
ation  thyroid_surgery  ...   TT4   T4U   FTI  classes
0          0            0                0          0 ...  105.0  0.88  120.0      1
1          0            0                0          0 ...   98.0  0.81  121.0      1
2          0            0                0          0 ...  121.0  1.11  109.0      1
3          0            0                0          0 ...   93.0  0.73  127.0      1
4          0            0                0          0 ...  111.0  0.68  165.0      1
...       ...          ...              ...        ... ...    ...   ...    ...    ...
3147       0            0                0          0 ...  151.0  1.40  108.0      0
3148       1            0                0          0 ...  104.0  1.01  103.0      0
3149       0            0                0          0 ...  106.0  1.03   99.0      0
3150       1            0                0          0 ...  240.0  1.07  223.0      0
3151       0            0                0          0 ...  123.0  0.99  124.0      0

[3152 rows x 24 columns]>
```

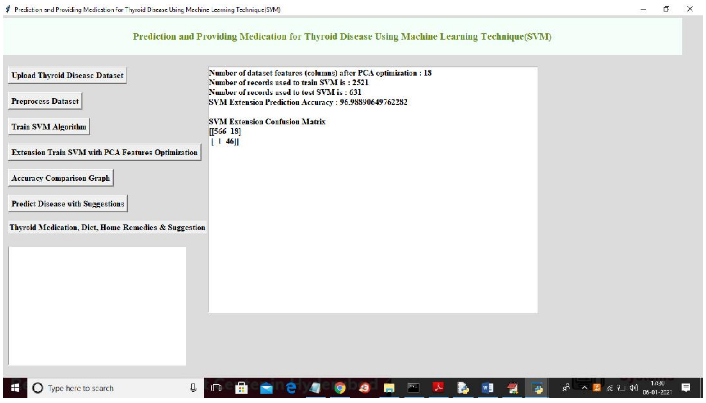
In above screen dataset loaded and displaying few records from dataset and then click on ‘Preprocess Dataset’ button to remove missing and NAN values from dataset and to separate X and Y values where X contains all dataset values and Y contains class label value.


Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Type here to search

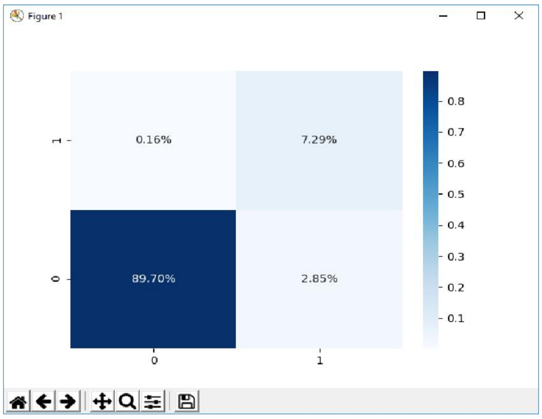
```
Total Records after preprocessing are : 3152
```



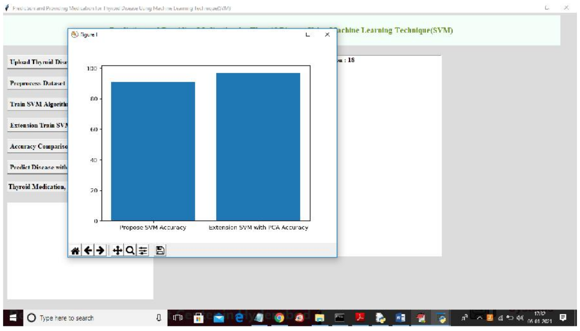
In above graph 84.15% and 6.81% is the true prediction and now click on 'Extension Train SVM with PCA Features Optimization' button to train SVM with PCA features optimization and to get below prediction accuracy

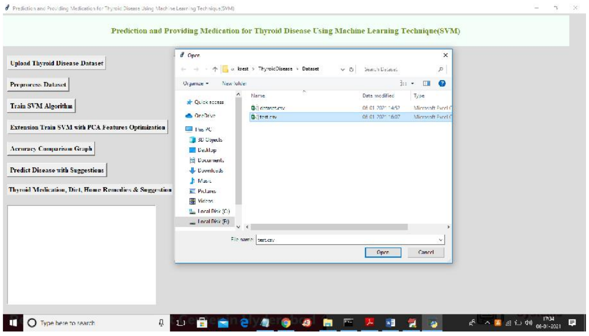

Prediction and Providing Medication for Thyroid Disease Using Mechine Learning Technique(SVM)
Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Type here to search

```
Number of dataset features (columns) after PCA optimization : 18
Number of records used to train SVM is : 2521
Number of records used to test SVM is : 631
SVM Extension Prediction Accuracy : 96.98890649762282

SVM Extension Confusion Matrix
[[566  18]
 [  1  46]]
```

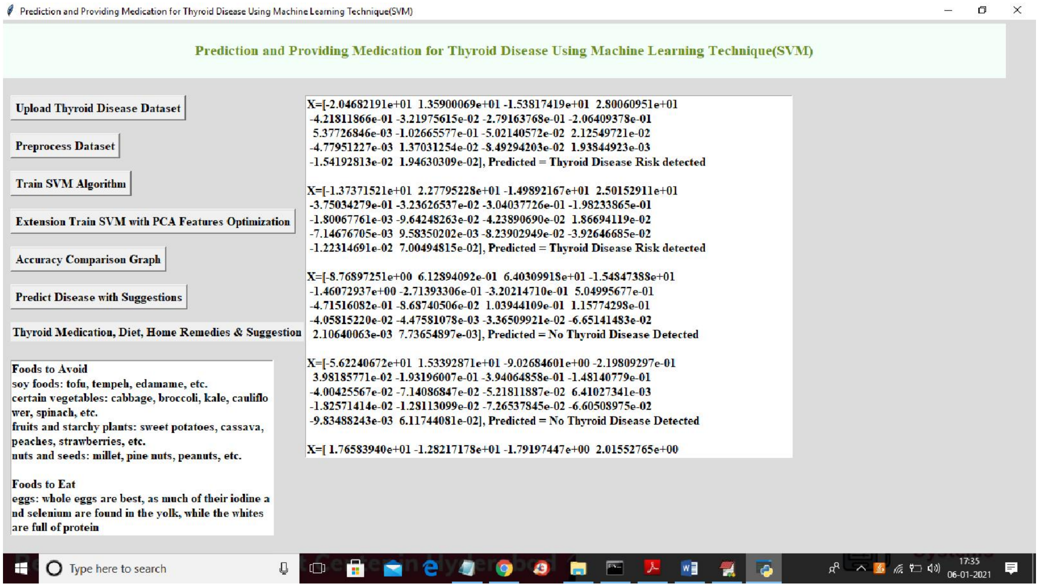


In above graph 89.70 and 7.29% is the correct prediction and other values are the false prediction. Now click on 'Accuracy Comparison Graph' button to get below accuracy comparison graph




Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Open
krest > ThyroidDisease > Dataset
Organize
New folder
Name
Date modified
Type
dataset.csv
test.csv
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
File name: test.csv
Open
Cancel
Type here to search

DOI: 10.48175/IJARSCT-9428


Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Prediction and Providing Medication for Thyroid Disease Using Machine Learning Technique(SVM)
Upload Thyroid Disease Dataset
Preprocess Dataset
Train SVM Algorithm
Extension Train SVM with PCA Features Optimization
Accuracy Comparison Graph
Predict Disease with Suggestions
Thyroid Medication, Diet, Home Remedies & Suggestion
Type here to search
17:35
06-01-2021

Foods to Avoid
soy foods: tofu, tempeh, edamame, etc.
certain vegetables: cabbage, broccoli, kale, cauliflo
wer, spinach, etc.
fruits and starchy plants: sweet potatoes, cassava,
peaches, strawberries, etc.
nuts and seeds: millet, pine nuts, peanuts, etc.

Foods to Eat
eggs: whole eggs are best, as much of their iodine a
nd selenium are found in the yolk, while the whites
are full of protein

```
X=[-2.04682191e+01  1.35900069e+01 -1.53817419e+01  2.80060951e+01
 -4.21811866e-01 -3.21975615e-02 -2.79163768e-01 -2.06409378e-01
  5.37726846e-03 -1.02665577e-01 -5.02140572e-02  2.12549721e-02
 -4.77951227e-03  1.37031254e-02 -8.49294203e-02  1.93844923e-03
 -1.54192813e-02  1.94630309e-02], Predicted = Thyroid Disease Risk detected

X=[-1.37371521e+01  2.27795228e+01 -1.49892167e+01  2.50152911e+01
 -3.75034279e-01 -3.23626537e-02 -3.04037726e-01 -1.98233865e-01
 -1.80067761e-03 -9.64248263e-02 -4.23890690e-02  1.86694119e-02
 -7.14676705e-03  9.58350202e-03 -8.23902949e-02 -3.92646685e-02
 -1.22314691e-02  7.00494815e-02], Predicted = Thyroid Disease Risk detected

X=[-8.76897251e+00  6.12894092e-01  6.40309918e+01 -1.54847388e+01
 -1.46072937e+00 -2.71393306e-01 -3.20214710e-01  5.04995677e-01
 -4.71516082e-01 -8.68740506e-02  1.03944109e-01  1.15774298e-01
 -4.05815220e-02 -4.47581078e-03 -3.36509921e-02 -6.65141483e-02
  2.10640063e-03  7.73654897e-03], Predicted = No Thyroid Disease Detected

X=[-5.62240672e+01  1.53392871e+01 -9.02684601e+00 -2.19809297e-01
  3.98185771e-02 -1.93196007e-01 -3.94064858e-01 -1.48140779e-01
 -4.00425567e-02 -7.14086847e-02 -5.21811887e-02  6.41027341e-03
 -1.82571414e-02 -1.28113099e-02 -7.26537845e-02 -6.60508975e-02
 -9.83488243e-03  6.11744081e-02], Predicted = No Thyroid Disease Detected

X=[ 1.76583940e+01 -1.28217178e+01 -1.79197447e+00  2.01552765e+00
```

In above screen in brackets we can see each record test value and after bracket we can see value as thyroid risk detected or not and if detected then it left box we are showing diet and medication plan as suggestion

III. CONCLUSION

In above dataset first line contains section names and different columns contains values as 0 or 1 and on the off chance that patient is under thyroid drug or medical procedure, its segment worth will be 1 else 0 and in last section contains class mark as 0 or 1 where 0 methods patient record is typical and 1 method patient record contains thyroid sickness. There are more than 3000 rows and 24 columns in this dataset. We are using the PCA (principal component analysis) feature selection algorithm as an extension concept to optimize features or to reduce columns or features that are not important for prediction because all 24 columns are not available for prediction. In order to train the SVM algorithm, PCA will eliminate all unnecessary columns from the dataset and only use important attributes. As a result of optimizing features, SVM prediction accuracy can be improved.

REFERENCES

[1]. Santos, Y. K. Penya, J. Devesa, and P. G. Garcia, “N-grams-based file signatures for malware detection,” 2009.

[2]. K. Rieck, T. Holz, C. Willems, P. D¨ussel, and P. Laskov, “Learning and classification of malware behavior,” in DIMVA ’08: Proceedings of the 5th international conference on Detection of Intrusions and Malware, and Vulnerability Assessment. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 108–125.

[3]. E. Konstantinou, “Metamorphic virus: Analysis and detection,” 2008, Technical Report RHUL-MA-2008-2, Search Security Award M.Sc. thesis, 93 pages.

[4]. P. K. Chan and R. Lippmann, “Machine learning for computer security,” Journal of Machine Learning Research, vol. 6, pp. 2669–2672, 2006.

[5]. J. Z. Kolter and M. A. Maloof, “Learning to detect and classify malicious executables in the wild,” Journal of Machine Learning Research, vol. 7, pp. 2721–2744, December 2006, special Issue on Machine Learning in Computer Security.

[6]. Y. Ye, D. Wang, T. Li, and D. Ye, “Imds: intelligent malware detection system,” in KDD, P. Berkhin, R. Caruana, and X. Wu, Eds. ACM, 2007, pp. 1043–1047.

[7]. M. Chandrasekaran, V. Vidyaraman, and S. J. Upadhyaya, “Spycon: Emulating user activities to detect evasive spyware,” in IPCCC. IEEE Computer Society, 2007, pp. 502–509.

DOI: 10.48175/IJARSCT-9428

[8]. M. R. Chouchane, A. Walenstein, and A. Lakhotia, “Using Markov Chains to filter machine-morphed variants of malicious programs,” in Malicious and Unwanted Software, 2008. MALWARE 2008. 3rd International Conference on, 2008, pp. 77–84.

[9]. M. Stamp, S. Attaluri, and S. McGhee, “Profile hidden markov models and metamorphic virus detection,” Journal in Computer Virology, 2008.

[10]. R. Santamarta, “Generic detection and classification of polymorphic malware using neural pattern recognition,” 2006.

[11]. I Yoo, “Visualizing Windows executable viruses using self-organizing maps,” in VizSEC/DMSEC ’04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security. New York, NY, USA: ACM, 2004, pp. 82–89.

[12]. F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain,” pp. 89–114, 1988.

[13]. T. Mitchell, Machine Learning. McGraw-Hill Education (ISE Editions), October 1997.