Prediction of Swine Flu using a Hybrid Voting Algorithm

N. Thirupathi Rao[a], Debnath Bhattacharyya[b], Eali Stephen Neal Joshua[c] and Ch.V. Satyanarayana[d]

[a] Department of CSE, Vignan'sInstitute of Information Technology, Visakhapatnam, AP, India.
[b]Department of CSE, K L Deemed to be University, KLEF, Guntur, AP, India.
[c]Department of CSE, Vignan'sInstitute of Information Technology, Visakhapatnam, AP, India.
[d]Department of CSE, Raghu Engineering College, Visakhapatnam, AP, India.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Swine influenza disease is one of the causes that results in mortality in the modern society. The only way to eradicate this disease and save a person's life is to detect it at an early stage and cure it. Swine flu spreads easily by droplet transmission through cough, sneeze or even by touch. The main goal of this article is to find out which type of swine flu virus the patient suffering from by using Machine Learning algorithms. This article summarizes a prototype using data mining techniques, namely Random Forest, Support Vector Machine and a hybrid Voting Classifier Algorithm. And to determine which machine learning algorithm helps in finding accurate results in many fields. Using the patients' symptoms as test data it was evaluated if the disease was prevailing or not. Hence, in this current article, it was explored to see if swine flu can be detected in its early stages so that it can be cured immediately after detection in an early stage. Along with the Swine Flu dataset of patients with symptoms classifiers such as Naive Bayes, k nearest neighbor, SVM, Random Forest and generated ROC curves were used to provide accurate predictions

Keywords: Predictive modeling, Classification, Regression, Symptom's dataset, Genetic Algorithms, Machine learning.

1. Introduction

Firstly, Swine Flu is caused by a virus called influenza, which spreads from pigs to humans. It is a contagious disease which is deadly and widely spread. There is a vaccine for swine flu and can be cured only in the initial stages. Many people died due to delay in the prediction of disease. Almost 2,03,000 people died in the year 2009 due to late prediction. As prevention is better than cure, detecting the virus and preventing it in the initial stage would save several people's lives. The manual lab tests are expensive and costs nearly 8000rs.This project also helps the needy people who cannot afford the lab tests and reports[1,2].

One of the most common viruses caused in humans was the H1N1 influenza. This virus was caused by one of the common causes was with the pigs. The other common types of influenza are like the H1N2, H3N2 etc. Some common symptoms were observed in humans in recent times with these two viruses also. The H1N1 virus had become vary drastic and dangerous for the mankind in the year 2009 in India with a wide spread across the country. These viruses are spreading across the country as a seasonal spread by times[3].

Swine flu is observed rarely, but once its case is registered it spreads rapidly. The only way to eradicate this disease and save a person's life is to detect it at early stage and cure it. It is also difficult to detect it in labs because its symptoms are similar to other flu's. So, in the current article, an attempt was made to predict influenza virus in an early stage using classifiers such as Random Forest, Support Vector Machine with the combination of a hybrid Voting algorithm. Previous patients' dataset was used as training data along with classifiers as algorithm. Providing patients' symptoms as test data, it was evaluated if the disease was prevailing or not. Swine Flu is a commonly occurred disease in both male and female. However, any time a virus affects humans from a different source, it is a concern. Scientists cannot always tell how a new virus will affect people, or how it will change over time.

The motivation behind this problem is to improve efficiency of predicting Swine flu as well as to generate results quickly based on the symptoms. Using the latest dataset of Swine flu patients, we compared various data validation techniques to predict the best accuracy rate. While pre-processing the dataset, we calculated the performance measures like accuracy, recall and precision to generate Receiver Operating Characteristic (ROC) curves.

1.1. Pandemic(H1N1)2009 Cases[SwineFlu]

a. Case Suspected

A person having this flu can be considered as a case and the case can be treated as a suspect of this flu as the patient had a high fever, breath shortness and dry cough.

b. Case Confirmed

This case is considered when a patient with the symptoms of flu was identified and his situation is identified by a standard laboratory arranged by the local government.

2. Literature Review

Several authors had done various works on the area related to the current article.Some of the articles and authors who had done some related works are discussed in thecurrentsectiontoidentifythesimilarsetofworks.

JyotiSoni, UjmaAnsariandDipeshSharma[2] discussed about various techniques for identifying the heart diseases. They had compared various machine learning techniques for the identification of heart diseases. They had used the genetic algorithm for the prediction of heart diseases.

Mai Shouman, Tim Turner, and Rob Stocker et. al.[3] discussed in detail about the prediction of heart-based diseases by using the machine learning algorithms. They used the KNN algorithm and verified its efficiency for the prediction of heart diseases. They also presented the performance of KNN algorithm and tabulated its results for better understanding of the model.

Mangesh J. Shinde, S. S. Pawar[4] developed and discussed in detail about the identification of Swineflu with the help of decision trees and naïve bayes models. They had modeled the algorithm and results are generated such that to predict the Swine flu.

Ms. Ankita R. Borkar, Dr. Prashant R. Deshmukh[5] have worked on prediction of swine flu using K-NN classifier using the Euclidean distance by considering various parameters such as effected area in the country and various swine flu symptoms and found the highest rate on flu predictions in a particular area.

Harshavardhan Achrekar, Avinash Gandhe, Ross Lazarus, Ssu-Hsin Yu and Benyuan Liu[6] had worked on the identification and the spread of Influenza at the places of United States. They worked and discussed about the implementing of classification of text which was related to the Influenza with the help of tweets being generated related to this flu. The implementation source of data was collected form the Twitter.

Thakkar, Hasan and Desai[7] had discussed and studied about the death and recovered rates of this Flue. The authors had developed a model through which the identification of flu was made simple and easy. They had used both KNN algorithm and Naïve bayes classifier techniques.

From the above works, almost all the works are done with the various machine learning algorithms. Very few works are done with the combination of both machine learning and decision algorithms. Hence, in the current work an attempt has been made to analyze our new dataset with the combination of both these types of algorithms and tried to analyze the performance of these algorithms. The results and representations are discussed in the form of ROC Curve and other values.

2.1 Existing Systems and Its Drawbacks

In 2009, many people faced the fatal disease, Swine Flu which had taken away lives of innocent humans and animals. Researches predicted that almost 2lakh people faced death due to this dangerous virus. This count is not even expected and also not registered in the hospitals. All this happened due to delay in predicting the virus at labs[8, 9]. Many people could not afford for the tests such as blood tests, X-ray’s also lost their lives. These manual tests take me and delay the prediction. Since swine flu can only be treated at initial stage, using this project we predict the virus immediately.

Influenza is contagious virus, once it is found in a person there are huge chances that it effects and spread to other persons. It follows droplet transmission either by sneezing or cough or by direct touch of effected persons. If it is not detected at early stage and eradicated, it spread rapidly. Early-stage detection takes time in existing system such as lab tests which include chest x-rays, nasal swab test, blood tests, CBC scans[10, 11,12]. All such manual tests are of high cost which are not affordable by poor people. Even though existing systems are helpful, they take more time to detect the virus which may not save the person's life and also harm other individuals.

2.2 Proposed System

The proposed system, was designed in a way to enhance the results in the existing system by using updated approaches and technologies such as Database Management System, Machine learning. We use classifiers such as Support Vector Machine, Random Forest and a Hybrid Voting Classifier algorithm.

a) In the current work, the Swine flu dataset of patients was considered and uses a classifier to build the probabilities and then we give a sample test data and thus we evaluate whether the virus is prevailing in that person or not.
b) This is a simple process and it takes less time. Using this we detect the virus at early stage so that it can be cured easily and save the person's life.
c) We use different classifiers to detect the virus and generate ROC curves (Receiver Operating Characteristic) which has True positive and false positive values.
d) Thus, this proposed system is more efficient, accurate and helpful.
e) We have researched and make comparisons of different algorithms such as Support Vector Machine, Random Forest using performance measures and also compared various classifiers on accuracy rates. The conclusion reached of this research that Random Forest and Voting Classifier has better accuracy on our data sets.

3. Algorithms

Several algorithms like Support Vector Machine, Random Forest and Voting algorithms for prediction of diseases were used in the current work. They were discussed in detail as,

1. Support Vector Machine
2. Random Forest
3. Hybrid Voting Classifier

3.1 Support Vector Machine

This Support Vector Machine Algorithm (SVM) belongs to the family of supervised machine learning algorithms. These supervised algorithms can be used for regression or the classification problems based on the requirement of the problem. The major task processed in the current algorithm was that the number of features we have will be analyzed with the value of each feature for being a value of particular coordinate[12]. In general, the support vectors are the points of data which lies very near to the hyperplane. The margin will be calculated as the distance between the nearest data point to the hyperplane. The final outcome or the output expected was the best hyperplane to be chosen with the highest margin value.

3.2 Random Forest

Like SVM algorithm, Random forest algorithm is also a supervised learning algorithm. In the current algorithm forest means a set of decision trees. This mode of trees development was done with the help of the technique called the bagging method. In general, the basic idea of this model was to the combine the learned models such that to get the maximum overall result. In other words, we can say that it generates multiple decision trees in number and combines all these decision trees such that to get the maximum output. As a result, we may predict or get the stable output or stable prediction of values.

3.3 Hybrid Voting ClassifierAlgorithm

The voting algorithm was proposed by the Boyer and Moore. This algorithm is mainly used for finding the majority in a set of elements arranged in a sequential format. The majority value can be calculated by using the techniques known as the constant space and the time in a linear fashion. The algorithm mainly identifies the majority element exist in the set of elements considered for the model to test. The element is identified if the element is being repeated many times with the same value. Once the setup is run and the value identified, it is verified again with the second time run of the program and the value was verified again and again for better analysis and performance of the model with the algorithm. The concept of voting is very simple and easy to understand when we are trying to solve the problems with the help of machine learning algorithms. Its different classifier with the other types of classifiers in the machine learning models.

4. Implementation and Results

For better understanding the performance of considered algorithms in predicting the Swine Flu presence from the data set collected through various local hospitals in Visakhapatnam Area. The entire implementation pf the total process had completed and compiled in the Python programming platform.

4.1 Methodology

Using the training data set along with different algorithms was used to predict the final output. At first, the general machine learning model is used for the prediction which includes the collection of data, preprocessing of data and then the data is being fed to the algorithm for executing. The results generated will be tabulated and the accuracy of each algorithm had noted. At last, the comparison of three algorithms had presented for better understanding of the performance of the model. The method followed for the implementation of dataset with respect to both SVM algorithm and the Random forest had shown in the following figures 3 and 4.

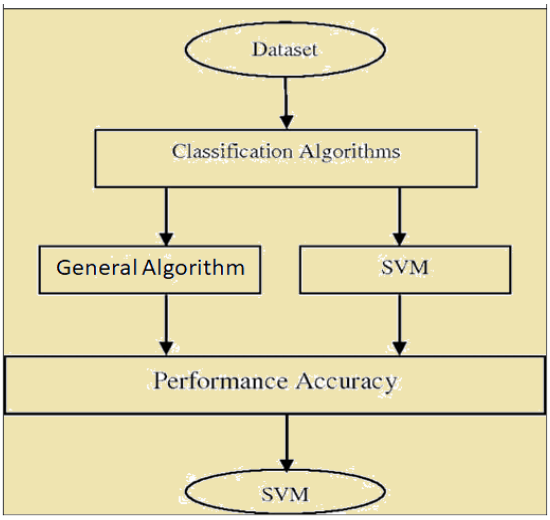


Figure 4 Swine Flu Prediction using Support Vector Machine (SVM)

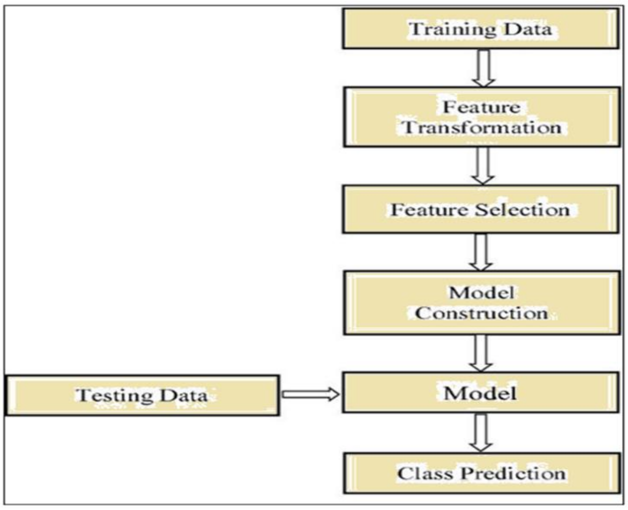


Figure 5 Swine Flu Prediction using Random Forest

The results for various algorithms like the Support Vector Machine algorithm, the Random Forest Algorithm and the mixed algorithm or the hybrid algorithm the Voting Algorithm which was a combination of both SVM and Random Forest algorithm had evaluated and represented the results in the section.

The input submission to each algorithm individually and the accuracy obtained from that particular algorithm had shown in each section separately and also the graphical representation of the model also represented for better understanding of the model and its performance.

4.2 Initial Diagnosis Method

Before going for the implementation of machine learning algorithms, basic algorithm had implemented for testing the model and results are tabulated.

The following table 1 shows the initial diagnosis symptoms and their values considered from the patients database.

Table 1 Initial Symptoms for initial diagnosis

Sl.No.	DryCough	Cold	BodyPains	Effected area	Immunity	Severity of Symptoms	Initial Diagnosis
0	YES	NO	NO	YES	HIGH	YES	NO
1	YES	NO	NO	YES	HIGH	NO	NO
2	NO	YES	NO	YES	LOW	YES	YES
3	NO	YES	YES	YES	HIGH	NO	NO
4	NO	NO	YES	NO	HIGH	NO	NO
...	...	...	...	...	...	...	...
746	NO	NO	NO	YES	LOW	YES	YES
747	NO	NO	YES	YES	LOW	YES	NO
748	YES	YES	NO	NO	HIGH	YES	YES
749	YES	NO	YES	NO	LOW	NO	YES
750	NO	YES	YES	NO	LOW	YES	YES

751rows×7columns

In the table 1, a total of 751 patient records had collected, tabulated and used for the processing of the model. The initial symptoms and features for initial diagnosis before preprocessing had displayed as,

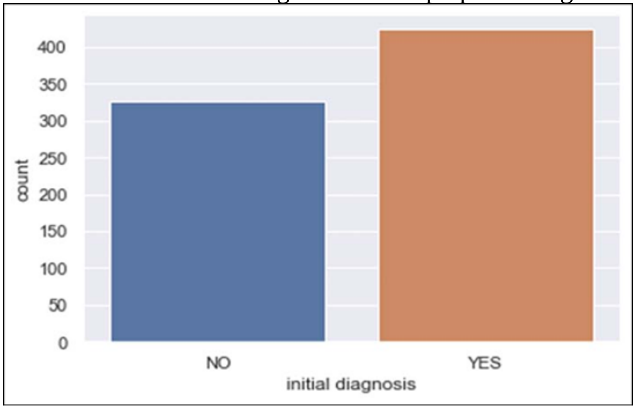


Figure 6 Representation of two classes (Yes and No) for Initial Diagnosis feature in training data or representation of count plot for Initial Diagnosis feature

Pre-processing of Data:

The pre-processing of the data should be done such that to remove or make corrected if any corrections or mistakes were present in the data during the collection of data.

After the preprocessing of the data had completed, the resultant pair plot for training data can be represented as follows at figure 7.

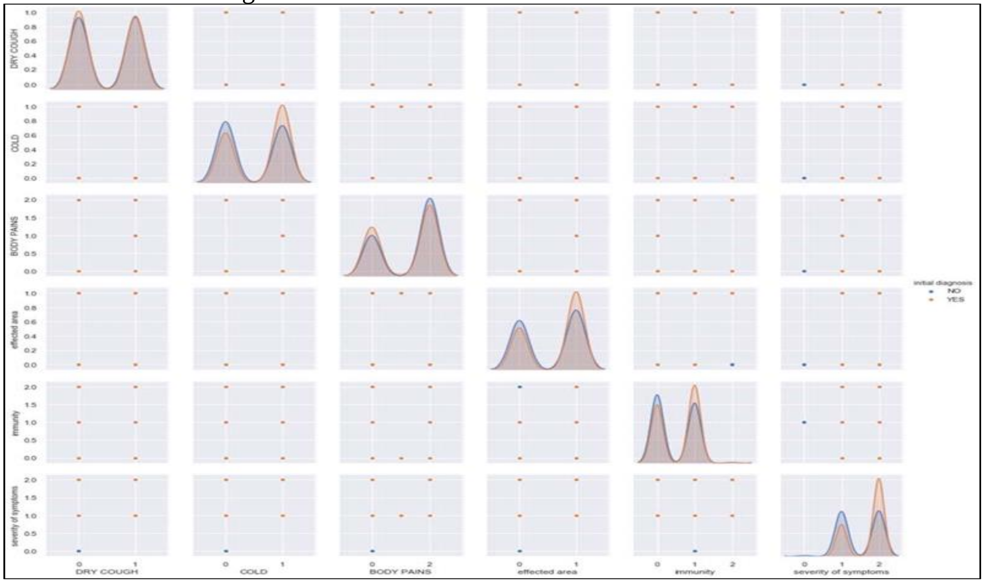


Figure 7 Representation of pair plot for training data features

The output of the initial diagnosis method with respect to various performance measures calculates and presented in the below table 2.

Table 2. Initial diagnosis results

Parameter	DryCough	Cold	BodyPains	EffectedArea	Immunity	SeverityofSymptoms
Count	751.000000	751.00000	751.000000	751.000000	751.00000	751.000000
Mean	0.491345	0.553928	1.234354	0.605859	0.536618	1.613848
Std	0.500258	0.497415	0.969367	0.488991	0.512176	0.495333
Min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000
50%	0.000000	1.000000	2.000000	1.000000	1.000000	2.000000
75%	1.000000	1.000000	2.000000	1.000000	1.000000	2.000000
max	1.000000	1.000000	2.000000	1.000000	2.000000	2.000000

The final output of the confirmation of whether the patient effected with Swine Flu or not could be represented with the following table 3 as,

Table 3 Output of the Initial diagnosis method for confirming the Swine Flu

S.No.	**InitialDiagnosis**
0	0
1	0
2	1
3	0
4	0
...	...
746	1
747	0
748	1
749	1
750	1

4.3 Random Forest Algorithm

One of the algorithms chosen for identifying the prediction of Swine flu with the dataset collected. The implementation of the algorithm had completed and the results with respect to the accuracy of the algorithm had achieved as,

The array output generated from the algorithm is as follows,

```
array([0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0,
1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0,
0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1,
0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1,
1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0,
0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1,
0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0,
0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0,
1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0,
0, 1, 1, 0, 1, 1, 1, 1, 0])
```

The accuracy of the current random forest algorithm achieved was,

Accuracy: 0.7888446215139442

The ROC curve generated for the performance of Random Forest algorithm was shown at figure

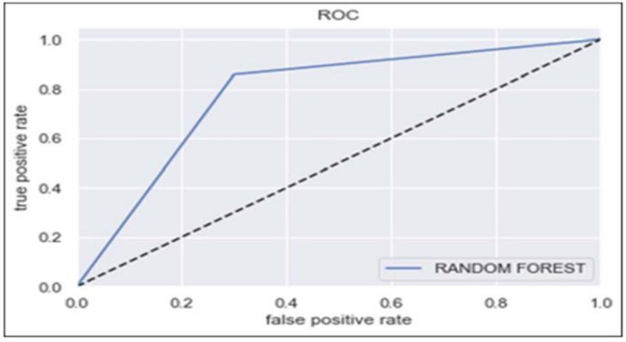


Figure 8 Receiver operating characteristics curve for test data using Random Forest algorithm

4.4 Support Vector Machine Algorithm

The other second algorithm used for the prediction of Swine Flu with the collected data set for the analysis of the data was the Support Vector Machine Algorithm.

The array output of the current algorithm is as follows,

```
array ([0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0,
0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0,
0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1,
0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0,
1, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0,
0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1,
0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0,
0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1,
1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0,
1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1,
1, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0,
0, 0, 1, 0, 1, 1, 1, 1, 0])
```

Accuracy: **0.701195219123506**

The ROC curve generated for the performance of Random Forest algorithm was shown at figure 6.

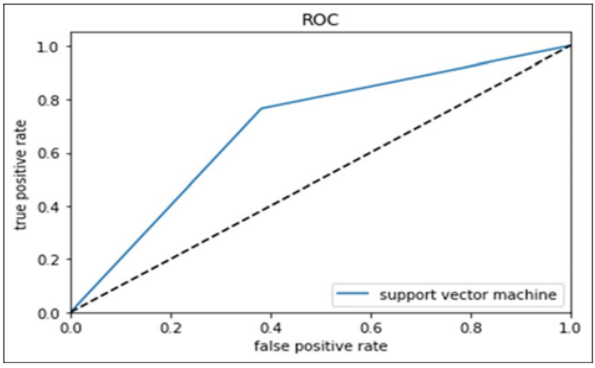


Figure 9 Receiver operator characteristic curve for test data using Support Vector Machine

4.5 Voting Algorithm

The other third algorithm which was a hybrid algorithm where both the Random forest algorithm and the support vector machine algorithms had combined for the test case to verify or to identify the best algorithm among the considered algorithms to verify their performance.

The output array for the current algorithm was as follows,

```
Array([0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0,
0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0,
0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1,
0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0,
1, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0,
0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1,
0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0,
0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1,
1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0,
1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1,
0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0,
0, 0, 1, 0, 1, 1, 1, 1, 0])
```

The accuracy of the current combined hybrid algorithm was,

Accuracy: **0.7330677290836654**

The ROC curve generated for the performance of Hybrid voting algorithm was shown at figure 10.

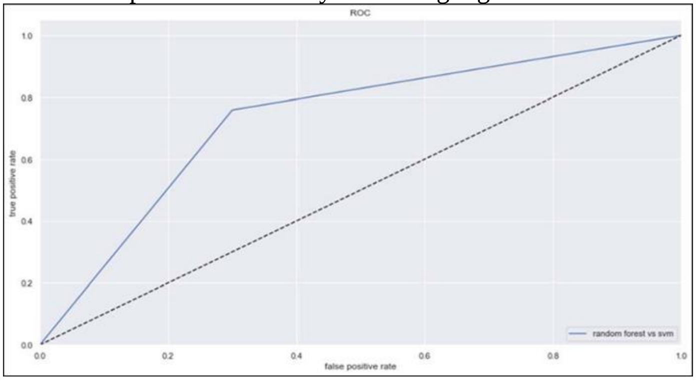
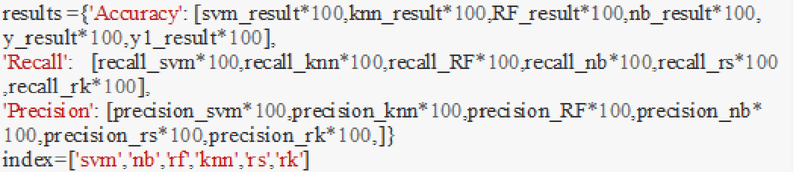


Figure 10 Receiver operator characteristic curve for test data for voting algorithm 1 (combining Random Forest and SVM)

5. Comparison of the Three Parameters Precision-Recall-Accuracy:

For better understanding of the performance of the algorithms with the current problem of identifying the Swine flu, three performance measures were calculated and a comparison had presented for better understanding of the model working with respect to all the algorithms.



```
results ={'Accuracy': [svm_result*100,knn_result*100,RF_result*100,nb_result*100,
y_result*100,y1_result*100],
'Recall':   [recall_svm*100,recall_knn*100,recall_RF*100,recall_nb*100,recall_rs*100
,recall_rk*100],
'Precision': [precision_svm*100,precision_knn*100,precision_RF*100,precision_nb*
100,precision_rs*100,precision_rk*100,]}
index=['svm','nb','rf','knn','rs','rk']
```

results =pd.DataFrame(results,index=index)

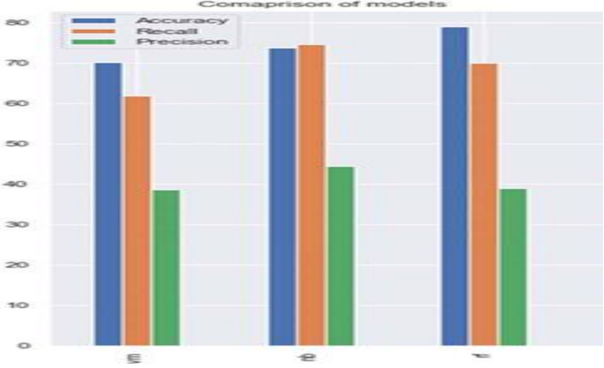


Figure 11 Comparison of accuracy, recall, precision of each algorithm used in Swine Flu prediction

The accuracy score achieved using Support Vector Machine is: **0.701195219123506%.**

The accuracy score achieved using Random Forest is: **0.7888446215139442 %.**

The accuracy score achieved using Voting Algorithm (Random Forest vs SVM) is: **0.7330677290836654 %.**

As a whole the, the total accuracy of all the algorithms had presented.

Table 4 Final Comparison of the accuracy of several algorithms usedinthe current work

Algorithms	Featuresusedinthedataset	Accuracy
SVM	DRY COUGH, COLD, BODY PAINS,EFFECTED AREAIMMUNITY, SEVERITYOFSYMPTOMS	0.70
RandomForest		0.78
VotingClassifier1		0.73

In the current work, several machine learning algorithms were used to predict whether a person was affected with swine flu or not based on the symptoms of the person. A dataset known as training data with previous patients' data was used with machine learning techniques along with the test data that has current data. As feature matrix, different symptoms were taken into consideration such as cold, dry cough, body pains, effected area, immunity etc. Tests were completed and compared between two machine learning algorithms such as SVM, Random Forest among which Random Forest Algorithm resulted in the highest accuracy above all the algorithms with 78 percent accuracy.

6. Conclusion

In the current work, Machine Learning Algorithms such as Support Vector Machine (SVM), Random Forest are used to classify the training data and also used Voting Classifier Algorithm to combine two algorithms to generate better performance or accuracy. We used patients' dataset as training data to build the training models for the selected machine learning classifiers. Several performance metrics are computed (accuracy rate, precision, false negative, false positive, true negative and true positive). We designed a model that can predict with high accuracy the symptoms of a person suffering from H1N1 influenza. In this we have taken Swine flu dataset of patients with a smaller number of features in training data and use a classifier to build the model and then we give a sample test data and thus we evaluate whether the virus is prevailing in that person or not. The Dataset has features such as symptoms, effected area, severity of symptoms, immunity power and thus diagnosis is done.

References

1. C. Griffin and R. Brooks. A note on the spread of worms in scale-free networks. IEEE TranonSys, Man, and Cyber, Part B: Cybernetics, 2006;36(1):198–202.
2. D. S. Jyoti Soni, Ujma Ansari. Predictive data mining for medical diagnosis: An overview of heart disease prediction. Int.J. Com.Appl. 2011;43-48.
3. Mai Shouman, Tim Turner, and Rob Stocker. Applying K-Nearest Neighbour in Diagnosing Heart disease Patients. IntJInfand EduTech 2012; 2(3):220-223.
4. Kumari, R., & Pandey, R. (2017). An Efficient Resource Scheduling Algorithm Using Knapsack. In *International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), ISSN (P): 2249-6831; ISSN (E): 2249-7943* (Vol. 7, No. 2, pp. 25-28). TJPRC Pvt. Ltd.
5. Mangesh J. Shinde, S. S. Pawar. Comparative Study of Decision Tree Algorithm and Naive Bayes Classifier for Swine Flu Prediction.IntJRes in Engg and Tech. 2015;4(6):45-50.
6. Khami, M. J. (2017). Unlimited Size of English Plain Text-in-Text Hiding Algorithm. *International Journal of Computer Science and Engineering (IJCSE)*, *6*.
7. Ms. Ankita R. Borkar, Dr. Prashant R. Deshmukh. Naïve Bayes Classifier for Prediction of Swine Flu Disease.Int Jof Adv Resin ComSciSoftEngg.2015;5(4):120-123.
8. Achrekar, Harshavardhan&Gandhe, Avinash& Lazarus, Ross & Yu, Ssu-Hsin& Liu, Benyuan. Twitter improves seasonal influenza prediction. HEALTHINF 2011
9. Patel, G., Patel, D. B., & Paghdal, K. M. (2016). Analysis of P&O MPPT algorithm for PV system. *International Journal of Electrical and Electronics Engineering (IJEEE)*, *5*(6), 1-10.
10. Thakkar, Hasan and Desai. Health care decision support system for swine flu prediction using Naïve Bayes classifier. Proceedings of the International Conference on Advances in Recent Technologies in Communication and Computing; 2010; Kottayam, India; 2010:101-105.
11. H. Qin, A. Shapiro, and L. Yang. Emerging infectious disease: A computational multi-agent model. ASE/IEEE International Conference onBioMedical Computing (BioMedCom).Dec 2012; USA; 2012:28–33.
12. Silambarasan, S., & Prabhakaran, G. (2018). Optimization of Process Parameters of Wire EDM Using Genetic Algorithm. *Carbon*, *100*, 0-036.
13. J. Xie and S. Jiang. A simple and fast algorithm for global k-means clustering. Second International Workshop onEducation Technology and Computer Science (ETCS)2010; 2: 36–40.
14. Sushma D., Thirupathi Rao N., Bhattacharyya D.A Comparative Study on Automated Detection of Malaria by Using Blood Smear Images. International Conference on Machine Intelligence and Soft Computing; September 2020; Advances in Intelligent Systems and Computing,Springer, Singapore, 1280:1-18.
15. Soliman, M. A., Abdel-All, N. H., Hussein, R. A., & El-Nini, W. M. (2016). Generalized Algorithm to the Extraction of Height Ridges in Riemannian Geometry. *Int. J. Manage. Infor. Technol. Eng*, *4*(2), 23-36.

16. B. Vamsi, B. Prakashand Debnath Bhattacharyya. Comparative Analysis of Prevalent Disease by Preprocessing Techniques Using Big Data and Machine Learning: An Extensive ReviewInternational Conference on Machine Intelligence and Soft Computing; September 2020; Advances in Intelligent Systems and Computing,Springer, Singapore, 1280:27-38.
17. B. Vamsi and Debnath Bhattacharyya. Prediction of brain stroke severity using machine learning.Revd'Int Art., 2020;34(6):753-761.
18. Yadav, S., & Sharma, G. (2015). Improvisation Of Data Mining Techniques In Cancer Site Among Various Patients Using Market Basket Analysis Algorithm. *BEST: International Journal of Management, Information Technology and Engineering (ISSN: 2348-0513) Volume*, *3*.
19. K. Thaiyalnayaki. Classification of Diabetes Using Deep Learning and SVM Techniques. Int J Cur ResRev. 2021;13(01):146-149.