

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Advanced Computing	
Series Title		
Chapter Title	Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination	
Copyright Year	2021	
Copyright HolderName	Springer Nature Singapore Pte Ltd.	
Corresponding Author	Family Name	Inampudi
	Particle	
	Given Name	Srividya
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Fr. Conceicao Rodrigues Institute of Technology
	Address	Vashi, Navi Mumbai, India
	Email	srivi2k@gmail.com
Author	Family Name	Johnson
	Particle	
	Given Name	Greshma
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Saintgits College of Engineering
	Address	Kottayam, India
	Email	
Author	Family Name	Jhaveri
	Particle	
	Given Name	Jay
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Vivekanand Education Society's Institute of Technology
	Address	Mumbai, India
	Email	
Author	Family Name	Niranjan
	Particle	
	Given Name	S.
	Prefix	
	Suffix	

	Role	
	Division	
	Organization	Sri Ramakrishna Engineering College
	Address	Coimbatore, India
	Email	
Author	Family Name	Chaurasia
	Particle	
	Given Name	Kuldeep
	Prefix	
	Suffix	
	Role	
	Division	School of Engineering and Applied Sciences
	Organization	Bennett University
	Address	Delhi, India
	Email	
	Family Name	Dixit
	Particle	
	Given Name	Mayank
Author	Prefix	
	Suffix	
	Role	
	Division	School of Engineering and Applied Sciences
	Organization	Bennett University
	Address	Delhi, India
	Email	
	Family Name	
	Particle	
	Given Name	
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	
	Address	
	Email	
Abstract	<p>The H1N1 Flu that came into existence in 2009 had a great impact on the lives of people around the world. It was a life-threatening season to hundreds of people mainly below 65 years old which eventually made the World Health Organization (WHO) to declare it as the greatest pandemic in more than 40 years. To find out the vaccination status National 2009 H1N1 Flu Survey (NHFS) was conducted in U.S. In this paper, the data from the above survey was used to develop a model that predicts how likely people got H1N1 and seasonal flu vaccine. For this purpose, various Machine Learning (ML) and Artificial Neural Network (ANN) models are used to determine the probability of person receiving H1N1 and Seasonal Flu vaccine.</p>	
Keywords (separated by '-')	Machine learning - H1N1 - Seasonal flu - SVM	



Machine Learning Based Prediction of H1N1 and Seasonal Flu Vaccination

Srividya Inampudi¹✉, Greshma Johnson², Jay Jhaveri³, S. Niranjan⁴,
Kuldeep Chaurasia⁵, and Mayank Dixit⁵

¹ Fr. Conceicao Rodrigues Institute of Technology, Vashi, Navi Mumbai, India
srivi2k@gmail.com

² Saintgits College of Engineering, Kottayam, India

³ Vivekanand Education Society's Institute of Technology, Mumbai, India

⁴ Sri Ramakrishna Engineering College, Coimbatore, India

⁵ School of Engineering and Applied Sciences, Bennett University, Delhi, India

Abstract. The H1N1 Flu that came into existence in 2009 had a great impact on the lives of people around the world. It was a life-threatening season to hundreds of people mainly below 65 years old which eventually made the World Health Organization (WHO) to declare it as the greatest pandemic in more than 40 years. To find out the vaccination status National 2009 H1N1 Flu Survey (NHFS) was conducted in U.S. In this paper, the data from the above survey was used to develop a model that predicts how likely people got H1N1 and seasonal flu vaccine. For this purpose, various Machine Learning (ML) and Artificial Neural Network (ANN) models are used to determine the probability of person receiving H1N1 and Seasonal Flu vaccine.

Keywords: Machine learning · H1N1 · Seasonal flu · SVM

1 Introduction

H1N1 or swine flu virus first emerged in 2009, spring season in Mexico and then in the United States and quickly spread across the globe. A distinctive combination or integration of influenza genes was discovered in this novel H1N1 virus which was not identified prior in humans or animals [1]. This contagious novel virus had a very powerful impact on the whole world and spread across the world like a forest fire and as a result on June 11 2009 the World Health Organization (WHO) declared that a pandemic of 2009 H1N1 flu or swine flu had begun [2]. The effects of this novel H1N1 virus were more severe on people below the age of 65. There was significantly high pediatric mortality, and higher rate of hospitalizations for young adults and children [3].

According to Centres for Disease Control and Prevention (CDC) the first and foremost step in protecting oneself of this virus is a yearly flu vaccination [4]. There are various factors such as age, health perceptions of an individual and the similarities or “match” in the vaccine’s virus structure and the virus structure which is affecting the community which affects the ability of the vaccination to provide protection to the person who is vaccinated [5]. Several activities were performed using various social media

platforms and broadcasting networks such as Twitter was used to track the levels of disease activity and the concern of the public towards this pandemic [6]. The social media played an important role to assess the sentiments towards vaccination and the implications for disease dynamics and control [7] etc. The popular among them is the phone survey conducted by the U.S. where they asked respondents whether they had received the H1N1 and seasonal flu vaccines, in conjunction with questions about themselves.

In the present study, we used the data obtained from the National 2009 H1N1 Flu Survey (NHFS) to predict how likely people got H1N1 and seasonal flu vaccines. The NHFS data is used for estimating the probability of a person receiving H1N1 and Seasonal Flu vaccine using various Machine Learning (ML) and Artificial Neural Network (ANN) models. The performance of various ML and ANN techniques are also discussed. In Sect. 2 literature review is presented. Section 3 discusses the data resource i.e. NHFS survey and Sect. 4 presents the methodology used. Section 5 discusses the results obtained and Sects. 6 and 7 presents conclusion and future research scope.

2 Literature Review

Mabrouk et al. [8] “A chaotic study on pandemic and classical (H1N1) using EIIP sequence indicators”, states that the methods such as moment invariants, correlation dimension, and largest Lyapunov exponent which were used to detect H1N1 indicated the differences between the pandemic and classical influenza virus. Chinh et al. [9] “A possible mutation that enables the H1N1 influenza A virus to escape antibody recognition” explained the methods such as phylogenetic analysis of pandemic strains, molecular docking for the predicted epitopes. Huang et al. [10], “Aptamer-modified CNTFET (Carbon NanoTube Field Effect Transistors) biosensor for detecting H1N1 virus in a droplet,” suggested the combination immersed in nanotube which gives CNTFET and thus it acts as a biosensor which is used in the detection of H1N1 virus by droplet.

M. S. Ünlü [11], “Optical interference for multiplexed, label-free, and dynamic biosensing: Protein, DNA and single virus detection,” described interferometric reflectance imaging sensor which can be used for label-free, high throughput, high sensitivity and dynamic detection and gives detection of H1N1 virus and nanoparticles and Kamikawa et al. [12] “Pandemic influenza detection by electrically active magnetic nanoparticles and surface plasmon resonance” indicated that the detection consists of several processes such as nanoparticle synthesis, glycans, polyaniline, and sensor modification by means to find H1N1 by nanoparticle and resonance. Jerald et al. [13], “Influenza virus vaccine efficacy based on conserved sequence alignment,” spoke about the vital strain sequence used from National Center for Biotechnology Information (NCBI) and sequence alignment which helps vaccine efficiency for influenza.

Chrysostomou, et al. [14] “Signal-processing-based bioinformatics approach for the identification of influenza A virus subtypes in Neuraminidase genes” discussed the methods used for identification of influenza virus such as neuraminidase genes, signal processing, F-score, Support Vector Machines (SVM) and Wiriyachaiporn et al. [15] “Rapid influenza an antigen detection using carbon nano string as the label for lateral flow immune chromatographic assay,” presented preparation of allantoic fluid infected with influenza A virus conjugation of Central Nervous System (CNS) to antibody and about the evaluation of CBNS-MAb using Lateral Flow Immunoassay (LFIA)

and Ma et al. [16], “An integrated passive microfluidic device for rapid detection of influenza a (H1N1) virus by reverse transcription loop-mediated isothermal amplification (RT-LAMP)” demonstrated the loading of virus and magnetic beads and discussed about virus capture, collection of virus-magnetic beads complexes, removal of excessive wastes, virus particle lysis, RT-LAMP reaction and the coloration steps to detect H1N1 virus.

Nieto-Chaupis, Huber. [17]. “Face To Face with Next Flu Pandemic with a Wiener-Series-Based Machine Learning: Fast Decisions to Tackle Rapid Spread” explained about the Wiener model used in order to increase optimization, efficiency and performance to find the spread of seasonal flu and Stalder et al. [18] “Tracking the flu pandemic by monitoring the social web” related the retrieving data from Twitter and official health reports provides inexpensive and timely information about the epidemic and Motoyama et al. [19] “Predicting Flu Trends using Twitter Data” demonstrated the use of SNEFT model and twitter crawler methods for predicting the flu using twitter data.

Wong et al. [20] “Diagnosis of Response Behavioural Patterns Towards the Risk of Pandemic Flu Influenza A (H1N1) of Urban Community Based on Rasch Measurement Model” presented the source of data and data analysis methodology used for the response behavioral patterns towards H1N1 and Bao et al. [21] “Influenza-A Circulation in Vietnam through data analysis of Hemagglutinin entries” provided NCBI influenza virus resource datasets (2001–2012) which is used for the analysis of influenza virus and Hu et al. [22], “Computational Study of Interdependence Between Hemagglutinin and Neuraminidase of Pandemic 2009 H1N1” explained sequence data and informational spectrum model.

3 Data Resources

Data is one of the most important and vital aspect of any research study. The National Flue Survey (NFS) is being conducted since 2010–11 influenza season [23]. The data for our study is obtained from the National 2009 H1N1 Flu Survey (NHFS) which was carried out for Centres for Disease Control and Prevention (CDC). The main aim of the survey was to monitor and evaluate H1N1 flu vaccination efforts among adults and children. The survey was conducted through telephones, twitter and with the help of various other electronic media in all the 50 states. The survey consists of national random digit dialed telephone survey based on rolling weekly sample of landline and cellular telephone contacted to identify residential households. Various questions about flu related behaviors, opinions about flu vaccine’s safety and effectiveness, medical history like recent respiratory illness and pneumococcal vaccination status were asked apart from the major question about H1N1 and seasonal flu vaccination status. The NHFS data was collected during Oct., 2009 to May, 2010. This data was obtained to get a fair idea about the knowledge of people on the effectiveness and safety of flu vaccines and to learn why some people refrained from getting vaccinated against the H1N1 flu and seasonal flu. Huge amount of data was gathered through this survey which is being commonly used for analysis and research purposes and the data also measures the number of children and adults nationwide who have received vaccinations.

4 Methodology

A methodology is proposed to determine the probability that a person will receive H1N1 and seasonal Flu vaccination based on many parameters. The data obtained from the National 2009 H1N1 Flu Survey (NHFS) contains 3 CSV files namely the training set features, the training set labels, and the test set features. The data has been obtained from over 53000 people from which around 26000 observations have been considered for the training set and the rest have been considered for the testing set.

We have considered various methodologies and compared different Machine Learning and Artificial Neural Network models to predict the probability. The Machine Learning algorithms such as Multiple Linear regression, Support Vector Regression, Random Forest Regression and Logistic Regression were used. The system architecture of Machine Learning model is presented in Fig. 1.

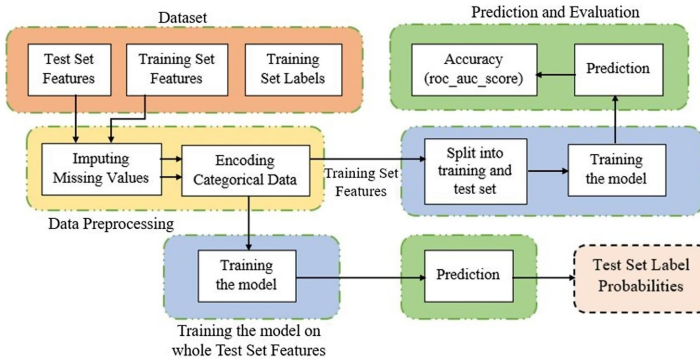


Fig. 1. System architecture of machine learning model

Artificial Neural Network (ANN) with different optimizers such as Adam, RMSprop, SGD were used to predict the probability of the test set features. The system architecture of ANN is presented in Fig. 2.

4.1 Taking Care of Missing Data

The missing data values in the dataset were imputed by univariate feature imputation using the most frequent strategy (statistics) with the help of the Simple Imputer class from the sklearn.impute module. This Simple Imputer class provides fundamental strategies for assigning some value for the missing values in the columns of the dataset [24, 25].

4.2 Encoding Categorical Data

For encoding categorical data the columns are transformed separately and then the features generated during this transformation of columns are concatenated to a single feature space and this process is carried out with the help of Column Transformer class of sklearn.compose module. Heterogeneous or columnar data is most benefited from this

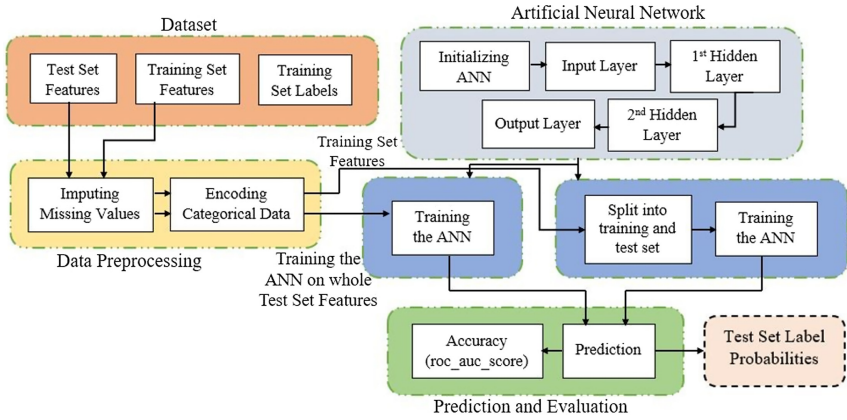


Fig. 2. System architecture of artificial neural network

method as several feature extraction mechanisms can be combined together or it gets transformed into a one transformer [24]. Then the categorical data is encoded using the One Hot Encoder class of the sklearn.preprocessing module. The features are encoded using an encoding scheme. A binary column is created in this method for every category. A sparse matrix or a dense array is returned depending on the sparse parameter [25].

4.3 Splitting the Dataset

The training set features and training set labels have been split into training set (80%) and testing set (20%) using `train_test_split` from `sklearn.model_selection`. This library splits the dataset into training and testing sets.

4.4 Hyperparameter Tuning

Hyperparameter tuning is done to find the most optimal parameter for the model on which the model gives the best results. We have used various Hyperparameter tuning methods such as `GridsearchCV`, `RandomSearchCV` for our machine learning models to obtain better results. K fold cross Validation method has been used to tune hyperparameters for the Artificial Neural Network.

4.5 Training the Model on the Dataset

An open-source Python library `scikit-learn` is used which uses a unified interface to implement many machine learning, preprocessing, cross-validation, and visualization algorithms [26]. `Scikit-learn` can be used for both supervised and unsupervised learning, using a consistent, task-oriented interface [24].

In `scikit-learn` all supervised estimators implement a `fit(X, y)` method to fit the model i.e. to train a model and a `predict(X)` is a method that gives predicted labels y for given unlabelled observations X .

4.6 Predicting the Results

Given the training model, we have predicted the label output for the test set features using the predict function of the model. Probabilities of the labels of test set features have also been predicted using the predict_proba method. In this predict_proba method, the highest probability is returned [24, 27].

4.7 Evaluation of the Results

There are various evaluation methods available to measure the performance and the quality of the prediction made by the model such as roc_auc_score(), r2_score(), Confusion Matrix, etc.

In our implementation, we have used the roc_auc_score() method from the sklearn.metrics library. roc_auc_score metric is essentially defined for binary classification tasks. In this by default the positive class labeled is 1 and only the positive labelled class is evaluated. Then the roc_auc_score function computes the area under the curve by this the curve information can be summarized in one number which is denoted as AUC or AUROC [24, 27].

5 Results and Discussion

5.1 H1N1 and Seasonal Flu Vaccination

In the default models the best performing method on the dataset has been the Artificial Neural Network method with 2 hidden layers and activation function being selu and the optimizer being Stochastic Gradient Descent (SGD) optimizer. The sigmoid function is used in the output layer for activation function. The accuracy obtained with ANN is shown in Table 1. Other Machine Learning algorithms have also yielded comparatively good results except logistic regression which has been the worst performing model with accuracy less than 70% in both H1N1 flu and seasonal flu vaccination prediction. Comparison of all the methods used during implementation are presented in Table 1.

The Results are also plotted using the ROC AUC curve. In Figs. 3, 4, 5 and 6 we can observe the performance of various models on the dataset and it can be concluded that Artificial Neural Network method has performed the best with accuracy over 82% in H1N1 flu vaccination prediction and 86% in Seasonal flu vaccination prediction.

5.2 H1N1 and Seasonal Flu Vaccination in ML Models with Hyperparameter Tuning

To obtain better results tuning of parameters has been carried out. Various methods such as GridSearchCV, RandomSearchCV, kfold method were used for hyperparameter tuning of the Machine Learning models. It is learnt that Support Vector Machine with Radial Basis Function (RBF) kernel and C:20 using the GridSearchCV method yields the best result for H1N1 vaccination prediction as shown in Table 2. The optimal parameters for

Table 1. Results for H1N1 flu and Seasonal flu vaccination prediction

Model	H1N1 score	Seasonal flu score	Parameters
svm	0.8085	0.8596	{ 'kernel': 'rbf' }
random_forest	0.8154	0.8494	{ 'n_estimators': 100 }
logistic_regression	0.6792	0.5949	{ max_iter=1000 }
ANN	0.8257	0.8601	{ '1st hidden layer: 'units=60, activation='selu' '2nd hidden layer: 'units=30, activation='selu' output layer: 'units=1, activation='sigmoid' 'optimizer='SGD', loss='binary_crossentropy' }

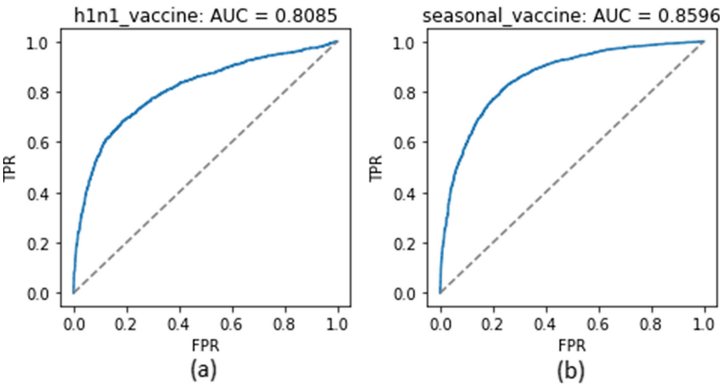


Fig. 3. ROC AUC Curve using Support Vector Machine: RBF Kernel for (a) h1n1 vaccine and (b) seasonal flu vaccine

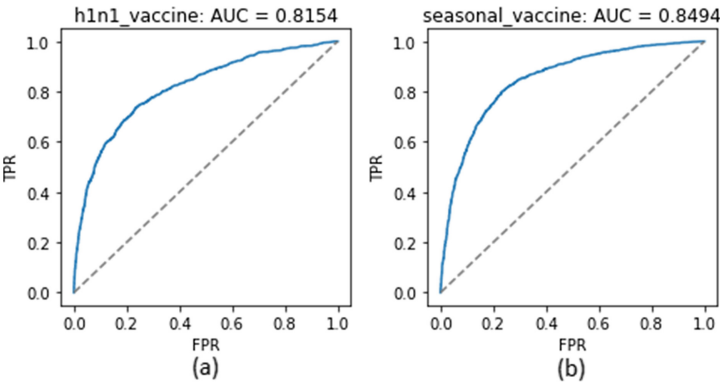


Fig. 4. ROC AUC Curve using Random Forest Regressor for (a) h1n1 vaccine and (b) seasonal flu vaccine

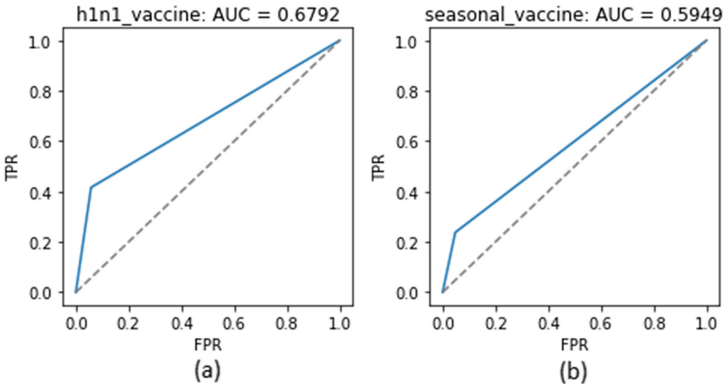


Fig. 5. ROC AUC Curve using Logistic Regression for (a) h1n1 vaccine and (b) seasonal flu vaccine

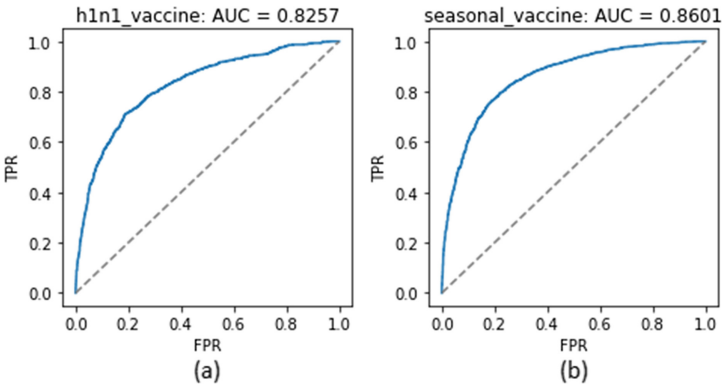


Fig. 6. ROC AUC Curve using Artificial Neural Network for (a) h1n1 vaccine and (b) seasonal flu vaccine

random forest regression are training the model with ‘10’ $n_estimators$, and the optimal parameters for logistic regression is C:5. All these results are presented in tabulated form in Table 2 and Table 3. It is observed that the results of Seasonal flu vaccination prediction have not been upto the mark using hyperparameter tuning, they were better predicted using the default models.

Table 2. Results with Hyperparameter tuning (GridSearchCV) for H1N1 flu vaccination prediction

Model	H1N1 score	Seasonal flu score	Parameters
svm	0.8397	0.7836	{ 'C':20, 'Kernel': 'rbf' }
random_forest	0.8213	0.7504	{ 'n_estimators': 10 }
logistic_regression	0.8363	0.7799	{ 'C': 5 }

Table 3. Results with Hyperparameter tuning (RandomSearchCV) for H1N1 flu vaccination prediction

Model	H1N1 score	Seasonal flu score	Parameters
svm	0.8367	0.7836	{ 'Kernel': 'rbf', 'C':20 }
random_forest	0.8205	0.7467	{ 'n_estimators': 10 }
logistic_regression	0.8363	0.7799	{ 'C':5 }

5.3 H1N1 and Seasonal Flu Vaccination in Artificial Neural Network with Hyperparameter Tuning

Kfold method is used to fine tune hyperparameters in the Artificial Neural Network method. The obtained results are more or less equal to the default method but a marginal increase in performance is noted which can be clearly seen in Table 4. The most optimal parameters obtained for ANN with kfold method are 1st hidden layer with selu as activation function and having 60 units, the 2nd hidden layer with selu as activation function and having 3 units, and the output layer with sigmoid as activation function and having 2 units. All the results are presented in Table 4.

Table 4. Results with Hyperparameter tuning (kfold method) for H1N1 flu and Seasonal vaccination prediction

Model	Best score	Parameters
ANN	0.8323	{ '1st hidden layer: 'units=60, activation='selu' '2nd hidden layer: 'units=3, activation='selu' output layer: 'units=2, activation='sigmoid' 'n_splits'=10, 'shuffle'='True', 'optimizer'='SGD', loss='binary_crossentropy' }

6 Conclusion

In this paper, prediction of H1N1 and seasonal flu vaccination are carried out using the data source given by the National 2009 H1N1 flu survey (NHFS) for Center of Disease Control and Prevention (CDC). Various ML and ANN models are used for prediction of H1N1 and Seasonal Flu vaccination. The model studies are improved using several techniques such as taking care of missing data, encoding categorical data, hyperparameter tuning and splitting of data set for training and testing purposes. The results obtained from various models are compared and evaluation. The results indicated that prediction of H1N1 vaccination is done best by the help of SVM model with RBF kernel with the help of hyperparameter tuning using GridSearchCV which yielded an accuracy of 83.97% and seasonal flu vaccination prediction is done best with Artificial Neural Network which yielded an accuracy of 86.10%.

7 Future Research Scope

Although we have achieved promising results, our study has many limitations. Mainly, the use of Twitter for data collection is not uniform across time and geography. Due to this inconsistency in the data our model's performance may vary and suffer. The accuracy difference can be noticed among the regional level and national level pertaining to the fact that people of the same region usually tend to have the same behavioural aspects hence lack of proper data could tamper the implementation. In future with advancements in technology the quality and the quantity of data could increase which could result in better performance and analysis of the issue. More information about the seasons, especially non-pandemic seasons could be very helpful for analysis of this project. In future we also look forward to exploring more Machine Learning algorithms, methods and deep learning techniques to obtain more optimal results.

Acknowledgement. The work presented in this paper is carried out as part of Internship project at Bennett University, Noida, India. Success of our Internship Project involving such high technical proficiency requires patience and massive support of guides. We take this opportunity to express our gratitude to those who have been instrumental in the successful completion of this work. Big thanks to Dr. Madhushi Verma for all the encouragement, timely details and guidelines given to our team. We would also like to thank Dr. Deepak Garg, HOD of Computer Science Engineering Department and Dr. Sudhir Chandra, Dean, School of Engineering & Applied Sciences, Bennett University for giving us the opportunity and the environment to learn and grow.

References

1. CDC. <https://www.cdc.gov/flu/pandemic-resources/2009-h1n1-pandemic.html>. Accessed 21 June 2020
2. CDC. <https://www.cdc.gov/h1n1flu/qa.htm>. Accessed 22 May 2020
3. CDC. <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5929a2.htm>. Accessed 22 May 2020
4. CDC. <https://www.cdc.gov/h1n1flu/>. Accessed 22 May 2020

5. CDC. <https://www.cdc.gov/flu/pastseasons/0910season.htm>. Accessed 22 May 2020
6. Signorini, A., Segre, A.M., Polgreen, P.M.: The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS ONE* **6**(5), e19467 (2011)
7. Marcel, S., Khandelwal, S.: Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comput. Biol.* **7**(10), e1002199 (2011)
8. Mabrouk, M.S., Marzouk, S.Y.: A chaotic study on pandemic and classical (H1N1) using EIIP sequence indicators. In: *Proceedings of 2nd International Conference on Computer Technology and Development*, Cairo, pp. 218–221 (2010)
9. Chinh, T.T.S., Stephanus, D.S., Kwoh, C., Schönbach, C., Li, X.: A possible mutation that enables H1N1 influenza a virus to escape antibody recognition. In: *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Hong Kong, pp. 81–84 (2010)
10. Huang, J., Lin, T., Chang, W., Hsieh, W.: Aptamer-modified CNTFET biosensor for detecting H1N1 virus in droplet. In: *Proceedings of 4th IEEE International NanoElectronics Conference*, Tao-Yuan, pp. 1–2 (2011)
11. Ünlü, M.S.: Optical interference for multiplexed, label-free, and dynamic biosensing: Protein, DNA and single virus detection. In: *Proceedings of XXXth URSI General Assembly and Scientific Symposium*, Istanbul, pp. 1–2 (2011)
12. Kamikawa, T., et al.: Pandemic influenza detection by electrically active magnetic nanoparticles and surface plasmon resonance, nanotechnology. *IEEE Trans. Nanotechnol.* **11**, 88–96 (2012)
13. Baby Jerald, A., Gopalakrishnan Nair, T.R.: Influenza virus vaccine efficacy based on conserved sequence alignment. In: *Proceedings of International Conference on Biomedical Engineering (ICoBE)*, Penang, pp. 327–329 (2012)
14. Chrysostomou, C., Seker, H.: Signal-processing-based bioinformatics approach for the identification of influenza A virus subtypes in Neuraminidase genes. In: *Proceedings of 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3066–3069 (2013)
15. Wiriyaichaiorn, N., Sirikett, H., Dharakul, T.: Rapid influenza a antigen detection using carbon nanostrings as label for lateral flow immunochromatographic assay. In: *Proceedings of 13th IEEE International Conference on Nanotechnology (IEEE-NANO 2013)*, Beijing, pp. 166–169 (2013)
16. Ma, Y., Chang, W., Wang, C., Ma, H., Huang P., Lee, G.: An integrated passive microfluidic device for rapid detection of influenza a (H1N1) virus by reverse transcription loop-mediated isothermal amplification (RT-LAMP). In: *Proceedings of 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, Kaohsiung, pp. 722–725 (2017)
17. Nieto-Chaupis, H.: Face to face with next flu pandemic with a wiener-series-based machine learning: fast decisions to tackle rapid spread, pp. 0654–0658 (2019). <https://doi.org/10.1109/CCWC.2019.8666474>
18. Stalder, F., Hirsh, J.: Open source intelligence. *First Monday* **7**(6), 416 (2002)
19. Motoyama, M., Meeder, B., Levchenko, K., Voelker, G.M., Savage, S.: Measuring online service availability using Twitter. In: *Proceedings of Workshop on online social networks*, Boston, Massachusetts, USA (2010)
20. Wong, L.P., Sam, I.C.: Behavioral responses to the influenza A (H1N1) outbreak in Malaysia. *J. Commun. Health* **34**, 23–31 (2011)
21. Bao, Y., et al.: The influenza virus resource at the national center for biotechnology information. *J. Virol.* **82**(2), 596–601 (2008)

22. Hu, W.: Molecular features of highly pathogenic Avian and Human H5N1 Influenza a viruses in Asia. *Comput. Mol. Biosci.* **2**(2), 45–59 (2012)
23. Smith, P.J., Wood, D., Darden, P.M.: Highlights of historical events leading to national surveillance of vaccination coverage in the United States. *Public Health Rep.* **126**(Suppl 2), 3–12 (2011)
24. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2012)
25. Buitinck, L., et al.: API design for machine learning software: experiences from the scikit-learn project (2013)
26. Dubosson, F., Bromuri, S., Schumacher, M.: A python framework for exhaustive machine learning algorithms and features evaluations. In: *Proceedings of IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, Crans-Montana, pp. 987–993 (2016)
27. Virtanen, P., Gommers, R., Oliphant, T.E., et al.: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261–272 (2020)

Author Queries

Chapter 11

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	Please check and confirm if the authors Given and Family names have been correctly identified.	
AQ3	As Refs. [24] and [27] are same, we have deleted the duplicate reference and renumbered accordingly. Please check and confirm.	