ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2022.0366

COVID-19 DISEASE IDENTIFICATION USING HYBRID ENSEMBLE MACHINE LEARNING APPROACH

Jyotsna Yadav and Richa Bhardwaj
University School of Information, Communication and Technology, Guru Gobind Singh Indraprastha University, India

Abstract
corona viral infected disease 2019 (Covid-19) has created a pandemic in year 2020 taking many lives and affecting millions of people. Due to lack of sufficient testing resources and healthcare systems, many countries and hospitals are not able to test this disease as the workload on the existing laboratories is increasing. In the proposed work, we have used hybrid ensemble machine learning models to predict this disease based on clinical variables and standard clinical laboratory tests. The main motive of the ensemble model is that combination of classifiers will classify the unseen data samples more accurately and chances for misclassification is very less as compared to the classification made by a single classifier. The performance comparison from various classification techniques is also done to show that hybrid ensemble classifier has outperformed decision tree and Support Vector Machine based classification algorithms.

Keywords:
Hybrid Ensemble Learning, Decision Tree, Support Vector Machine

1. INTRODUCTION

Covid-19 started from Wuhan (located in china) in year 2019 and quickly got spread to all other countries in the world. By the year 2020 World Health Organisation (WHO) declared covid-19 or corona as pandemic to the entire human population [1]. Since then, various research work is being conducted to find the vaccine or cure for the covid-19. Since, many individuals across the world are getting mental issues due to this pandemic condition they are committing suicide in fear that they might transfer this disease to their family or loved ones, so the classification between Covid-19 and influenza patients becomes very necessary [2].

In this paper, the grouped power of decision tree and support vector machine learning algorithms is applied to COVID-19 dataset so that it can classify the data samples more accurately. This is the concept of ensemble learning. Ensemble model, which is also called as combination of classifier, is inspired by the human group decision making process. It has been found favourable in a wide variety of application domains. In this approach, different types of machine learning (ML) models are merged together so that the performance of those ML models can be improved. The ML models that take part in ensemble learning are called as ensemble or weak learner. Each of the weak learners or machine learning techniques give their own result and finally, the results of each machine learning (ML) models are grouped for obtaining the final result [3].

This is the theory of hybrid ensemble leaning which is based on machine learning in which weak learners are of different type or we can say that weak learners are heterogeneous. The main element of the success of ensembles is a concept of diversity [4].

Our goal is to show that the hybrid ensembles can give best classification performance based on accuracy as compared to ensembles in which weak learners are of same type i.e., the weak learners are homogeneous. Here, the performance of individual weak learning models and the homogenous collection of weak learners will be calculated and then their performances will be compared with the performance of our hybrid ensemble model.

This paper is organized as follows: section 2 gives the review of literature. Section 3 gives preliminary description; section 4 describes the proposed methodology. In section 5, the experimental results are presented. After that performance comparison of various classification techniques is comprised in section 6 and in section 7, conclusion of proposed work is explained.

2. LITERATURE REVIEW

In previous section, introduction of ensemble models and their contribution to improve the classification performance made by a single classifier is given. Although several machine learning approaches have been used to classify the Covid-19 and influenza patients, but we have used the grouped power of decision tree and support vector machine for overall classification.

Rustam et al. [1] gave a demonstration of the capacity of ML techniques for predicting the number of Covid-19 patients that are coming in future. The ML techniques that were used include linear regression (LR), least absolute shrinkage and selection operator (LASSO), Support Vector Machine (SVM), and exponential smoothing (ES). Among these ML models, ES model gave the best performance according to the proposed work.

In the research work [2], the ML models and deep learning models are exploited for forecasting reach ability of Covid-19 over the nations in future. This was done by adopting the real-time details from the Johns Hopkins dashboard.

Study [3] exploited a novel hybrid ensemble method for clustering the features that are extracted from a pharmaceutical database. Soft clusters were formulated by using unsupervised learning strategies. After that the decisions were merged by adopting the parallel data fusion techniques. The affiliations in the features and blend the decisions produced by machine learning techniques was observed for discovering the strong clusters that can affect overall accuracy of classification.

Qi et al. [4] exploited hybrid ensemble method to ameliorate the slope stability prediction by making use of ensemble classifier and genetic algorithm. Study [5] identified multiple sclerosis by using a classification technique which is based on ensemble learning. The extraction of features is done with the help of an eighteen different Gray Level Co-occurrence Matrix. After that, on the extracted features, an ensemble learning which is based on decision tree is performed. Three distinct types of boosting techniques are used to classify MR image of brain that are healthy, from MR images of brain that are weak. The best performance is achieved by using ensemble machine learning techniques in terms of accuracy, specificity and sensitivity.

Narayan et al. [12] has proposed supervised machine learning technique for analyzing sentiments that are associated with unstructured text data. Efficiency is enhanced by doing an extensive study of ensemble machine learning algorithms in the domain of sentiment classification. After that different kind of machine learning algorithms are chosen to achieve better performance in terms of accuracy.

Jia et al. [17] adopted three models – Logistic model, Bertalanffy model and Gompertz model for predicting and analyzing the covid-19 disease. The prediction results of three different mathematical models are distinct for various parameters in various regions. Regression coefficient was used for evaluating the ability to fit the three models.

Yan et al. [19] implemented XG Boost classifier to predict patients at high risks to prioritize them and thus reduce the mortality rate. The accuracy of the algorithm obtained was 90% and the model could predict the mortality rate of patients 10 days in advance thus leading to less mortality rate.

In research work [20], machine learning analysis of genetic variants from asymptomatic, mild and severe covid-19 patients were performed to predict and classify the people based on their vulnerability or resistance to potential covid-19 infection in order to study the effect of covid-19 on patients and for clinical study of patients. However, in this study there was no suitable outcome indicating that there is no common pattern of covid-19 patients among asymptomatic, mild, and severe patients. Thus, in clinical study it is difficult to determine the stage of covid-19 patient.

Study [21] determines the feasibility of the usage of machine learning method in evaluation of prediction results. Prediction of confirmed cases, negative cases, released, and deceased cases of Covid-19 coronavirus was done. For this deep neural network - Recurrent Neural Network method is used. Three models long short-term memory (LSTM), Gated recurrent unit (GRU) and combined model of LSTM+GRU was proposed. Experimental Results showed that the combined approach LSTM-GRU-RNN provides quite better result over LSTM-RNN, GRU-RNN in terms of Accuracy, RMSE metrics. In the research work [23], performance of prediction is ameliorated by exploiting a hybrid ensemble model for credit scoring by combining Gabriel Neighborhood Graph editing (GNG) and Multivariate Adaptive Regression Splines (MARS).

3. PRELIMINARY DESCRIPTION

3.1 ENSEMBLE LEARNING

An ensemble is a machine learning model in which we combine the predictions of two or more distinct models. Ensemble members are counted as the models which contribute to the ensemble that might be of same type or different types also, it may or may not be trained using the same data for training. The predictions we make using ensemble members are often combined with the statistics same way like we do for mode or mean. They can be also combined in some different and better sophisticated methods which can learn, under what conditions, also that how much we can trust each member [5]. The process of ensemble learning is shown in Fig.1.

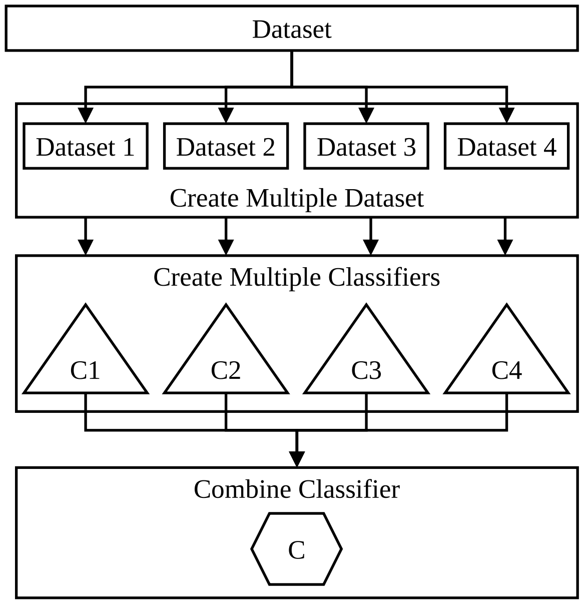


Fig.1. Ensemble Machine Learning

There are two main reasons for using an ensemble when compared with a single model:

- ***Performance***: By making more accurate and better prediction, an ensemble will help achieve better performance when compared to any single contributing model.
- ***Robustness***: An ensemble will also tone down the spreading and also dispersion of many predictions and model performances.

3.1.1 Ensemble Methods:

The ensemble methods are segregated in two groups:

- ***Sequential Ensemble Methods***: These methods can be defined as the methods in which the sequential generation of base learners is achieved. The main motivation behind these methods is to take advantage of the dependence between the base learners. Base learner or weak learner is the machine learning technique that participates in ensemble learning.
- ***Parallel Ensemble Methods***: These methods can be defined as the methods in which the base learners are generated independently from each other. The main motivation of such methods is to capitalize on the inter-dependence among the base learners, because the error can be decreased by averaging [6].

3.2 INTEGRATION OF WEAK LEARNERS

Ensemble learning algorithms work by combining individual predictions derived using multiple models. Thus, these models are designed to perform slightly better than any of the individual contributing ensemble member. First of all, base models are to be selected which is being aggregated to build an ensemble learning model. In homogeneous ensemble model, similar types of weak learners are trained in different ways. Bagging and boosting methods are used to combine those weak learners. On the other hand, different types of weak learners can be combined to build “heterogeneous or hybrid ensembles model”. Weak learners must be chosen in coherent manner with the way we aggregate them

[7]. An aggregating method which reduces the variance should be used whenever the base models which have low bias and high variance is chosen. On the other hand, an aggregating method which reduces bias should be used whenever we the base models which have low variance and high bias is chosen. There exist three kinds of meta-algorithms to combine these weak learners:

3.2.1 Bagging:

In bagging, homogeneous weak learners are used. It learns them in parallel manner being independent from each other. After that, it combines them by following some kind of deterministic averaging process. Bagging commonly referred to as Bootstrap aggregation, can be defined as the ensemble-based machine learning meta-algorithm which is designed to enhance the accuracy and stability of the machine learning algorithms which are used in statistical regression and classification [8]. It curtails the variance and assists in avoiding overfitting. Working of bagging algorithm is shown in Fig.2.

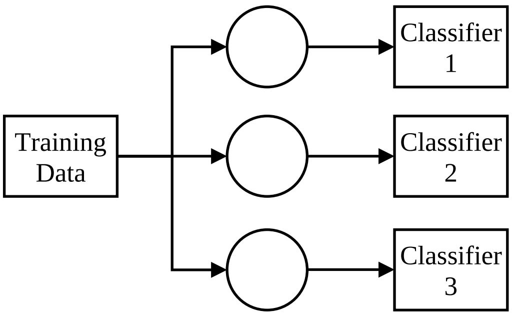


Fig.2. Implementation of Bagging Algorithm (Parallel)

3.2.2 Boosting:

Boosting is a unique technique which is used to generate collection of predictors. In this method, learning is done in a sequence by the learners with the help of early learning fitting models of data after which we examine the data for faults and errors [9]. Boosting can be referred to as the collection of algorithms which are capable of converting the heterogeneous weak learners into one strong learner. The major principle of the process of boosting is to arrange those different types of weak learners so as to fit them in sequence [10]. Working of boosting algorithm is shown in Fig.3.

3.2.3 Stacking:

Stacking can be defined as the ensemble learning technique which associates numerous regression or classification models via a meta-regressor or meta-classifier. Implementation of stacking is shown in Fig.4.

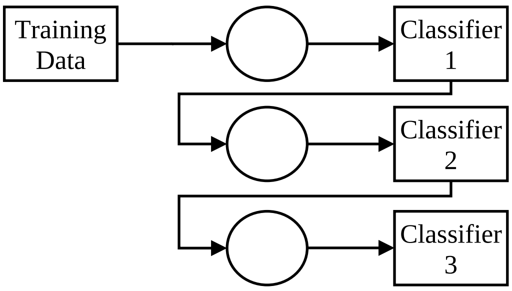


Fig.3. Implementation of Boosting Algorithm (Sequential)

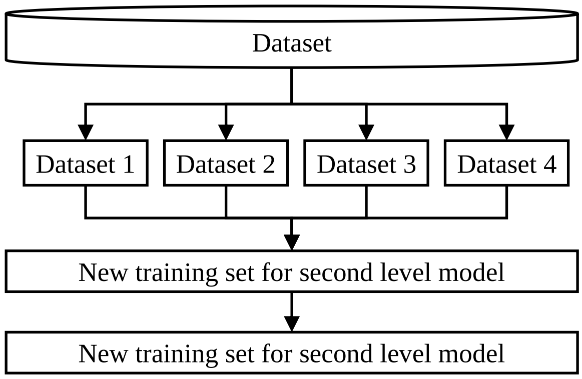


Fig.4. Implementation of Stacking Algorithm

The base level models are trained on the basis of complete training set, and then the meta-model gets trained as features which are also the outputs of the base level model [11]. The base level generally comprises of diverse learning algorithms and accordingly the stacking ensembles are sometimes heterogeneous in nature.

3.3 HYBRID ENSEMBLE MODEL

The ensemble learning focuses more on the collective processing of multiple models, which are logically or strategically yielded and results in the consolidation to solve a specific computational intelligence problem. The learning algorithms that contribute in the result of ensemble learning are called as ensemble members [12]. When same types of weak learners are used as ensemble members, it is called as homogeneous ensemble model. If different types of weak learners are used as ensemble members, it is called as heterogeneous ensemble model or hybrid ensemble model. The combination of several machine learning models is achieved by ensemble learning, due to which it enhances the results of the machine learning algorithms. The approach is regarded as remarkable as it generates improved predictive performance in correlation with a single model [13]. The ensemble learning focuses more on the collective processing of multiple models, which are logically or strategically yielded and results in the consolidation to solve a specific computational intelligence problem. The combination of several machine learning models is achieved by ensemble learning, due to which it enhances the results of the machine learning algorithms. The approach is regarded as remarkable as it generates improved predictive performance in correlation with a single model [14].

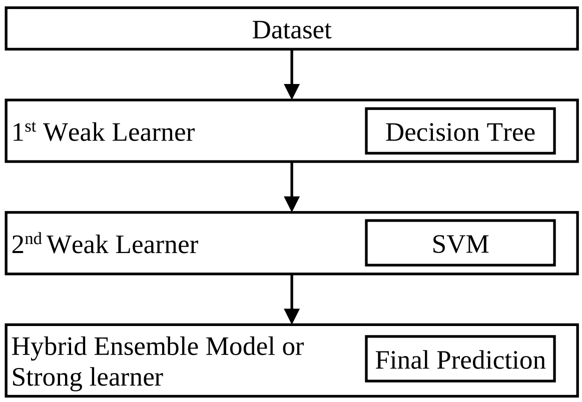


Fig.5. Implementation of Hybrid Ensemble Model

According to the Fig.5, the different machine learning models are sourced with the dataset which serves as the most important initial requirement for a typical machine learning model. Then, there are lots of different machine learning models for implementation purpose including Decision Tree, Support Vector Machine. Hence, according to the results obtained by these models after their successful implementation and execution, the final prediction is made as per the objectives to be accomplished [15].

4. METHODOLOGY USED

The workload for testing the Covid-19 patients is increasing on the existing laboratories. So, it becomes necessary to predict this disease using machine learning model. The dataset which contains several clinical variables, is taken from the following repository: <https://github.com/yoshihiko1218/COVID19ML> to classify covid-19 and influenza patients. There are 1486 instances and 51 attributes in the dataset. Data is pre-processed by taking 32 variables as machine learning input. The target attribute is having two outcomes that is H1N1 and COVID19. Where H1N1 describes that the patient has influenza and COVID19 describes that the patient is COVID-19 positive. We have converted categorical data to dummy variables using the LabelEncoder function in pre-processing from sklearn because machine learning algorithms do not allow non-numerical data. After that missing value is handled by using mean value imputation method. After that dataset is splitted into training and testing data having 297 instances as testing data and remaining instances are of training data. 10-fold cross validation was then performed and fed into different types of classification techniques to improve accuracy. These techniques include:

Decision Tree.

- SVM.
- Bagging of Decision Tree (DT) having 100 estimators.
- Bagging of SVM having 100 estimators.
- Hybrid Ensemble Learning which includes DT and SVM.

The steps for the methodology that we have used are shown in Fig.6.

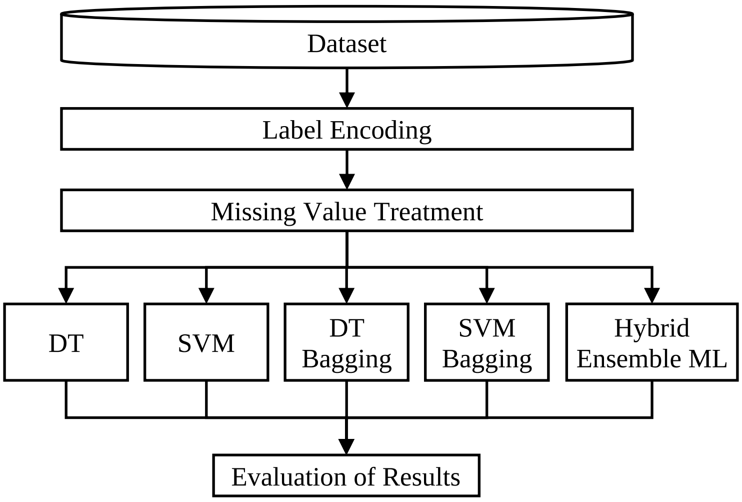


Fig.6. Process Flow Diagram of Proposed Methodology

In Hybrid Ensemble Learning, at first, we applied decision tree classification three times and after that we applied SVM three times. The output of first decision tree classification is given as the input to the second decision tree classification and so on. Finally, the output of third decision tree is fed-up as the input to the first SVM classification. The output of first SVM is given as the input to the second SVM and so on. Then output of third SVM is the final output. Finally, the results of classification techniques were compared by plotting receiver operating characteristic (ROC) curve and calculating area under curve from the corresponding ROC curve. After that, accuracy, specificity and sensitivity were also calculated.

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 DIFFERENT CLINICAL VARIABLES SUMMARY

There are a total of 51 distinct clinical variables which includes 27 continuous and 24 categorical variables. Categorical variables like gender, which has 55.59% (826 patients) male and 44.41% (660 patients) female and CReactiveProteinLevelsCategorical, of which 60 patients (4.03%) have normal CReactive protein levels, and 99 patients (6.66%) have high CReactive protein levels.

5.2 EVALUATION OF DIFFERENT CLASSIFICATION TECHNIQUES

Firstly, the ROC curve of prediction results for decision tree, SVM, bagging of decision tree, bagging of SVM and hybrid ensemble learning was plotted and we obtained an AUC of 88%, 95%, 88%, 94%, 95% respectively as shown in Fig.7. We have plotted a confusion matrix by using the prediction results from the decision tree. We calculated accuracy of 96.12%, a sensitivity of 77.21% and a specificity of 99.08% by using this confusion matrix. After that the result of SVM is used to plot confusion matrix to calculate accuracy, sensitivity and specificity which is 97.97%, 89.87% and 100% respectively. Now the accuracy of 93.71%, 92.38%, 98.23% and sensitivity of 77.21%, 88.60%, 89.87% and specificity of 99.08%, 100%, 100% are calculated by plotting confusion matrix from the result of decision tree bagging, SVM bagging, hybrid ensemble learning respectively as shown in Fig.8.

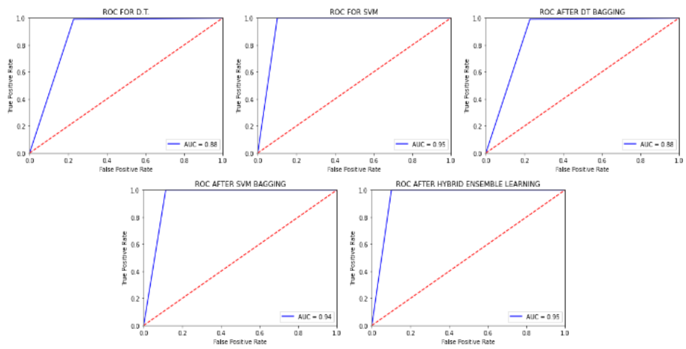


Fig.7. ROC Curve for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning

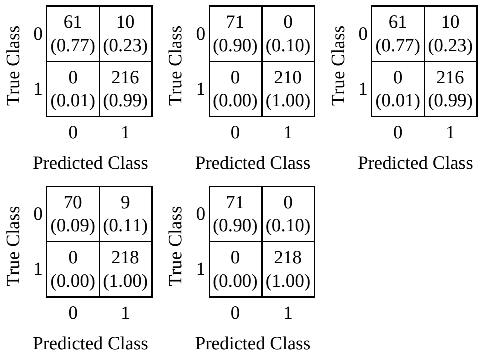


Fig.8. Confusion Matrix for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning

5.3 CLASSIFICATION PERFORMANCE ON BASIS OF AGE

We have applied our classification models to four distinct groups and those groups are based on age: Children (1-18), Young (19-39), Middle aged (40-65), old (age>65).

5.3.1 Evaluation of Different Machine Learning Model Outcomes for Children:

The ROC Curve is plotted for the different classification models from the prediction results. Using Decision Tree (DT), SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for children only, we have obtained an AUC of 88%, 92%, 88%, 92% and 100%, respectively (shown in Fig.9), while distinguishing influenza and COVID-19 patients.

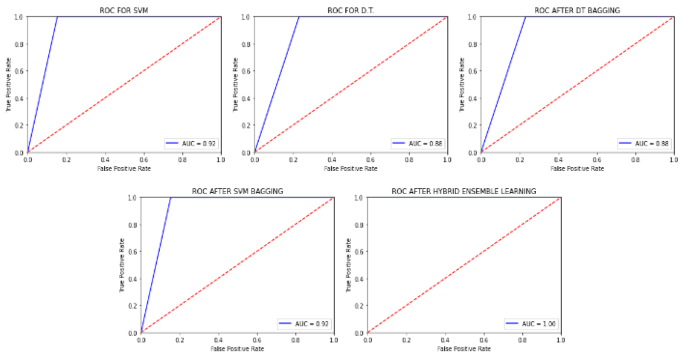
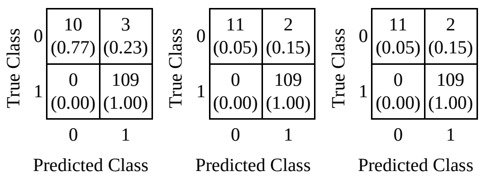


Fig.9. ROC Curve for DT, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Children

The prediction results for child patients from DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning were utilized to plot a confusion matrix (Fig.10). We have calculated accuracy of 98.15%, 97.74%, 90%, nan and 99.59% respectively by using those confusion matrices. The sensitivity of DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for children is of 100% for all models and specificities are of 76.92%, 84.61%, 76.92%, 84.61% and 100% respectively.



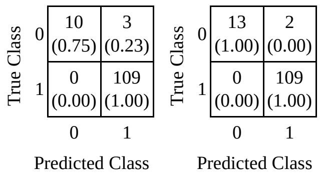


Fig.10. Confusion Matrix for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Children

5.3.2 Evaluation of Various Machine Learning Model Outcomes for Young Patients:

The ROC Curve is plotted for the different classification models from the prediction results. Using DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for young patients only, we have obtained an AUC of 98%, 98%, 98%, 98% and 100%, respectively (shown in Fig.11), while distinguishing influenza and COVID-19 patients.

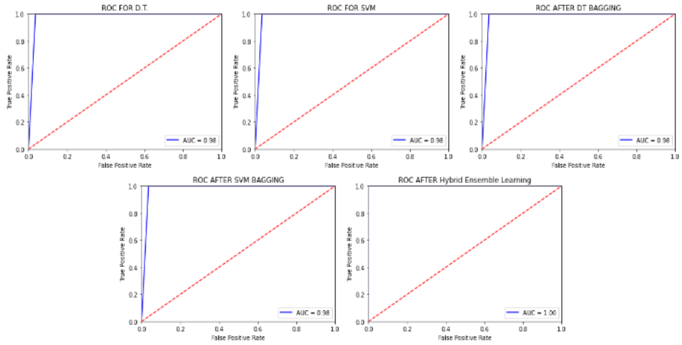


Fig.11. ROC Curve for DT, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Young Patients

The prediction results for young patients from DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning were utilized to plot confusion matrix (Fig.12). From those confusion matrices, the accuracy of 98.25%, 98.25%, 95.31%, 94.3% and 98.25% respectively was calculated. The sensitivity of DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for young patients is of 96.42%, 96.42%, 96.42%, 96.42% and 100% respectively and specificity is of 100% for each classification models.

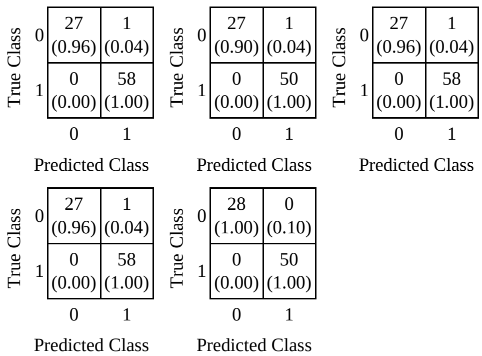


Fig.12. Confusion Matrix for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Young Patients

5.3.3 Evaluation of Various Machine Learning Model Outcomes for Middle Aged Patients:

The ROC Curve is plotted for all the classification models from the prediction results. Using decision tree, SVM, bagging of decision tree, bagging of SVM and hybrid ensemble learning for middle aged patients only, we have obtained an AUC of 96%, 96%, 96%, 96% and 98%, respectively (shown in Fig.13), while distinguishing between influenza and COVID-19 patients.

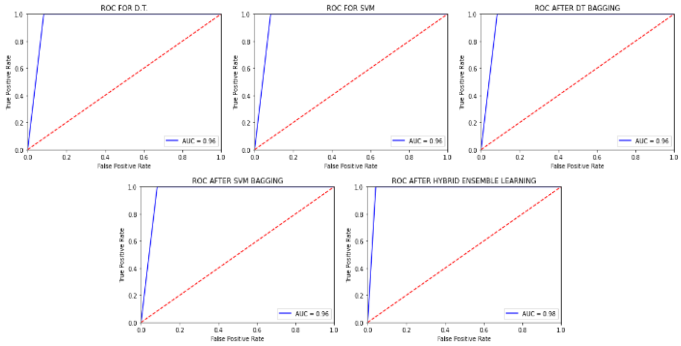


Fig.13. ROC Curve for DT, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Middle Aged Patients only

The prediction results for middle aged patients from DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning were used to plot corresponding confusion matrices (Fig.14).

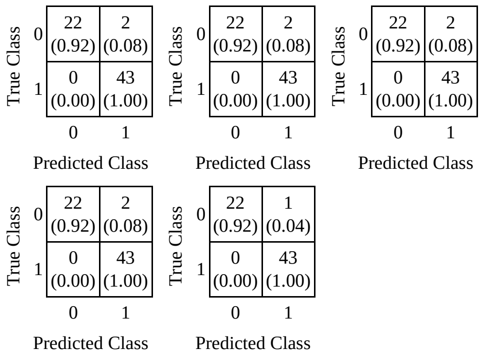


Fig.14. Confusion Matrix for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Middle Aged Patients

From those confusion matrices (shown in Fig.14), the accuracy of 94.74%, 96.58%, 93.74%, 89.80% and 97.74% respectively have calculated. The sensitivity of DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for middle aged patients is of 91.66%, 91.66%, 91.66%, 91.66% and 95.83% respectively and specificity is of 100% for each classification models.

5.3.4 Evaluation of Various Machine Learning Model Outcomes for Old Patients:

The ROC Curve is plotted for all the classification models from the prediction results. Using DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for old patients only, we have obtained an AUC of 95%, 100%, 95%, 100% and 100%, respectively (shown in Fig.15), while distinguishing between influenza and COVID-19 patients.

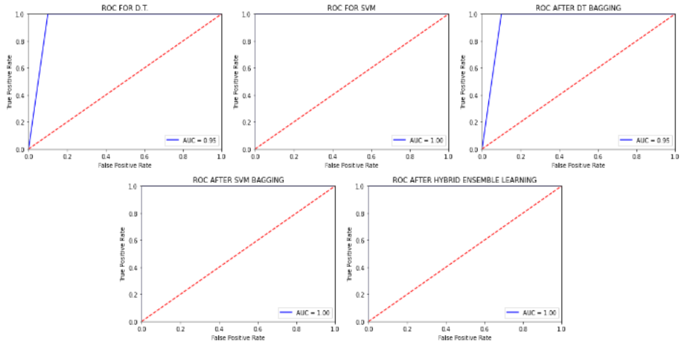


Fig.15. ROC Curve for DT, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Old Patients

The prediction results for old patients from DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning were used to plot confusion matrices (Fig.16).

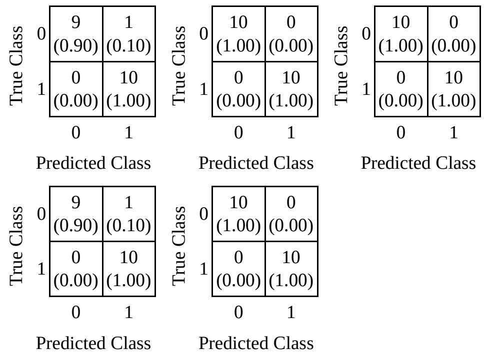


Fig.16. Confusion Matrix for Decision Tree, SVM, DT Bagging, SVM Bagging, Hybrid Ensemble Learning for Old Patients

From those confusion matrices shown in Fig.16, the accuracy of 77.5%, 87.5%, 76%, 82% and 87.5% respectively have calculated. The sensitivity of DT, SVM, bagging of DT, bagging of SVM and hybrid ensemble learning for middle aged patients is of 90%, 100%, 90%, 100% and 100% respectively and specificity is of 100% for each classification model.

5.4 PERFORMANCE COMPARISON OF VARIOUS CLASSIFICATION TECHNIQUES

The results of all the classification techniques are compared to know the improvement of accuracy (shown in Table.1) and area under curve (shown in Table.2) for all the patients as well as the different groups of patients on the basis of age.

Table.1. Accuracy of Classification between COVID-19 and influenza patients in different Age groups

Accuracy	DT	SVM	DT Bagging	SVM Bagging	Hybrid Ensemble Learning
Overall	96.12%	97.97%	93.71%	92.38%	98.33%
Children (1-18)	98.15%	97.74%	90%	Nan	99.59%
Young (19-39)	98.25%	98.25%	95.31%	94.3%	98.25%

Middle (40-65)	94.74%	96.58%	93.74%	89.80%	97.74%
Old (Age>65)	77.5%	87.5%	76%	82%	87.5%

Table.2. AUC of Classification between COVID-19 and influenza patients in different Age groups

Area Under Curve	DT	SVM	DT Bagging	SVM Bagging	Hybrid Ensemble Learning
Overall	88%	95%	88%	94%	95%
Children (1-18)	88%	92%	88%	92%	100%
Young (19-39)	98%	98%	98%	98%	100%
Middle (40-65)	96%	96%	96%	96%	98%
Old (Age>65)	95%	100%	95%	100%	100%

6. CONCLUSION AND FUTURE WORK

In this paper, different types of machine learning algorithms are utilized for predicting presence of COVID-19 using clinical variables. Those machine learning methods are also applied to different age groups patients to know the presence of COVID-19. COVID-19 patients have successfully distinguished Covid-19 patients from influenza patients by using these clinical variables for all the age groups. Also, it has been observed that hybrid ensemble machine learning model gives the best accuracy and area under curve among different machine learning models that is exploited in this work. So, a strong learner is successfully generated by the integration of two weak learners that is Decision Tree and SVM. In future, some other machine learning techniques will be exploited with the decision tree and SVM to classify the data sample in order to acquire better results.

REFERENCES

[1] F. Rustam, A. Reshi, A. Mehmood, S. Ullah and G. Choi, "COVID-19 Future Forecasting using Supervised Machine Learning Models", *IEEE Access*, Vol. 8, pp. 101489-101499, 2020.

[2] N. Punn, S. Sonbhadra and S. Agarwal, S., 2020. COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms", *Medrxiv*, Vol. 23, No. 1, pp. 1-14, 2020.

[3] V. Maheshwari, M.R. Mahmood and V.P. Sundramurthy, "Nanotechnology-Based Sensitive Biosensors for COVID-19 Prediction using Fuzzy Logic Control", *Journal of Nanomaterials*, Vol. 2021, pp. 1-8, 2021.

[4] C. Qi and X. Tang, "A Hybrid Ensemble Method for Improved Prediction of Slope Stability", *International Journal for Numerical and Analytical Methods in Geomechanics*, Vol 42, No. 15, pp. 1823-1839, 2018.

[5] S. Jain, N. Rajpal and J. Yadav, "Multiple Sclerosis Identification Based on Ensemble Machine Learning Technique", *SSRN Electronic Journal*, Vol. 12, No. 2, pp. 1-13, 2020.

[6] S. Kannan, P. Arulprakash and P. Dhinakar, "Analysis of Convolutional Recurrent Neural Network Classifier for COVID-19 Symptoms over Computerised Tomography Images", *International Journal of Computer Applications in Technology*, Vol. 66, No. 3-4, pp. 427-432, 2021.

[7] A. Khadidos, A.O. Khadidos, S.N. Mohanty and G. Tsaramirsis, "Analysis of Covid-19 Infections on a CT Image using Deepsense Model", *Frontiers in Public Health*, Vol. 30, pp. 751-757, 2020.

[8] P. Kumar, V. Kumar and K. Kumar, "An Analysis on Ensemble Classifiers in Ensemble Classification Problems", *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, No. 3, pp. 1-8, 2019.

[9] Ensemble Learning to Improve Machine Learning Results, Available at https://blog.statsbot.co/ensemble-learning-d1dcd548e936, Accessed at 2021.

[10] A. Torrent, X. Llado and A. Torralba, "A Boosting Approach for the Simultaneous Detection and Segmentation of Generic Objects", *Pattern Recognition Letters*, Vol. 34, No. 13, pp. 1490-1498, 2013.

[11] Xueyi Wang, "A New Model for Measuring the Accuracies of Majority Voting Ensembles", *Proceedings of IEEE International Joint Conference on Neural Networks*, pp. 103-108, 2012.

[12] R. Narayan, M. Roy and S. Dash, "Ensemble based Hybrid Machine Learning Approach for Sentiment Classification-A Review", *International Journal of Computer Applications*, Vol. 146, No. 6, pp. 31-36, 2016.

[13] H. Anwar, U. Qamar and A. Muzaffar Qureshi, "Global Optimization Ensemble Model for Classification Methods", *The Scientific World Journal*, Vol. 2014, pp. 1-9, 2014.

[14] D. Kumar, "A Hands-on Guide to Hybrid Ensemble Learning Models, With Python", Available at https://analyticsindiamag.com/a-hands-on-guide-to-hybrid-ensemble-learning- models-with-python-code, Accessed at 2021.

[15] R. Yang, S. Xu and L. Feng, "An Ensemble Extreme Learning Machine for Data Stream Classification", *Algorithms*, Vol. 11, No. 7, pp. 107-116, 2018.

[16] B. Verma and A. Rahman, "Cluster-Oriented Ensemble Classifier: Impact of Multicluster Characterization on Ensemble Classifier Learning", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24, No. 4, pp. 605-618, 2012.

[17] Y. Tominaga, "Comparative Study of Class Data Analysis with PCA-LDA, SIMCA, PLS, ANNs, and k- NN", *Chemometrics and Intelligent Laboratory Systems*, Vol. 49, No. 1, pp. 105-115, 1999.

[18] Z. Chen and Y. Lu, "Improving Relevance Feedback via using Support Vector Machines", *Advanced Materials Research*, Vol. 13, No. 2, pp. 2028-2032, 2011.

[19] L. Yan, H. Zhang and Y. Yuan, "An Interpretable Mortality Prediction Model for Covid-19 Patients", *Nature Machine Intelligence*, Vol. 2, pp. 283-288, 2020.

[20] A. Alimadadi, S. Aryal and I. Manandhar, "Artificial Intelligence and Machine Learning to Fight COVID-19", *Physiological Genomics*, Vol. 52, No. 4, pp. 200-202, 2020.

[21] S.K. Bandyopadhyay and S. Dutta, "Machine Learning Approach for Confirmation of Covid-19 Cases: Positive, Negative, Death and Release", *Ibero American Journal of Medicine*, Vol. 2, No. 3, pp. 172-177, 2020.

[22] L. Jia, K. Li, Y. Jiang and X. Guo, "Prediction and Analysis of Corona Virus Disease 2019", *Proceedings of IEEE International Conference on Neural Networks*, pp. 1-11, 2020.

[23] M. Alaraj and M. Abbod, "A New Hybrid Ensemble Credit Scoring Model based on Classifiers Consensus System Approach", *Expert Systems with Applications*, Vol. 64, pp. 36-55, 2016.

[24] Covid and Fludata, Available at https://github.com/yoshihiko1218/COVID19ML/blob/master/COVIDandFLUdata, Accessed at 2021.