P ISSN : 2618-1630
E ISSN : 2909-6130

Diastolic Dysfunction Prediction with Symptoms Using Machine Learning Approach

Muhammad Shoaib Anjum[1], Omer Riaz[2], Muhammad Salman Latif[1]
[1]Department of Computer Science, The Islamia University of Bahawalpur, Bahawalpur, Pakistan.
[2]Department of Information Technology, The Islamia University of Bahawalpur, Bahawalpur, Pakistan.
* Correspondence: Muhammad Shoaib Anjum Email ID:Shoaib.anjum@hotmail.com

Citation | Anjum. S. M, Riaz.O, Latif, S,M "Diastolic Dysfunction Prediction with Symptoms Using Machine Learning Approach". International Journal of Innovations in Science and Technology. Vol 4, Issue 3, 2022, pp:714-727
Received | May 09, 22; **Revised** |June 26, 22 **Accepted** | June 28, 22 **Published** | June 30, 22 DOI:https://doi.org/10.33411/IJIST/2022040312

Cardiac disease is the major cause of deaths all over the world, with 17.9 million deaths annually, as per World Health Organization reports. The purpose of this study is to enable a cardiologist to early predict the patient's condition before performing the echocardiography test. This study aims to find out whether diastolic function or diastolic dysfunction using symptoms through machine learning. We used the unexplored dataset of diastolic dysfunction disease in this study and checked the symptoms with cardiologist to be enough to predict the disease. For this study, the records of 1285 patients were used, out of which 524 patients had diastolic function and the other 761 patients had diastolic dysfunction. The input parameters considered in this detection include patient age, gender, BP systolic, BP diastolic, BSA, BMI, hypertension, obesity, and Shortness of Breath (SOB). Various machine learning algorithms were used for this detection including Random Forest, J.48, Logistic Regression, and Support Vector Machine algorithms. As a result, with an accuracy of 85.45%, Logistic Regression provided promising results and proved efficient for early prediction of cardiac disease. Other algorithms had an accuracy as follow, J.48 (85.21%), Random Forest (84.94%), and SVM (84.94%).Using a machine learning tool and a patient's dataset of diastolic dysfunction, we can declare either a patient has cardiac disease or not.

Keywords: Diastolic Dysfunction; Random Forest; J.48; Logistic Regression and SVM.

Acknowledgment. We are thankful to "The Heart Center Bahawalpur" for providing the data for experiment purposes.
Project details. Nil.

CONFLICT OF INTEREST: Authors declare there is no conflict of interest for publishing this manuscript in IJIST.

Author's Contribution. All the authors contributed equally

Introduction

Cardiac disease is the major cause of death all over the world, with 17.9 million deaths annually, as per the World Health Organization reports[1].Due to unawareness and some unhealthy activities, the causes of the increase in the risk of heart diseases like diabetes, obesity, hypertension etc.[2]. Heart disease can be predicted through many dysfunctions, like Systolic Dysfunction and Diastolic Dysfunction[3]. Systolic and diastolic function and dysfunction, derived from the echocardiography test and Doppler test as previously defined in[4]. Systolic and diastolic function are the two states of heart pumping. When the heart pumps blood outside the body, this phase is called systolic function, and the relaxing state of the heart when it fills the blood from the body is called diastolic function. Diastolic Dysfunction normally relates to the age factor. More than 50% of adults at the age of 70 years have diastolic dysfunction[5]. Diastolic dysfunction is responsible for nearly half of all heart failure patients [6, 7]. In this study, patients' basic data was collected from the private clinic of the cardiac center "The Heart Center Bahawalpur". A machine learning algorithm was used in this study to predict the diastolic function or dysfunction conclusion predicted by the cardiologist in this private clinic after an echocardiography test.

The cardiovascular system (CVS) consists of the heart and blood vessels. A wide range of dysfunctions can arise in the cardiovascular system. This dysfunction is considered a form of cardiovascular disease or heart disease (HD)[8].Heart disease can be predicted with some common indicators like age, BP, shortness of breath, hypertension, obesity etc. Through data mining or machine learning, it can be determined which features are more important or not to predict the specific disease. Machine learning (ML) is a systematic approach to building a modal and checking whether some features are able to predict some specific value or not.
Both supervised and unsupervised machine learning approaches help in a lot of studies to predict cardiac disease with both their techniques. Supervised machine learning is used when the prediction class is known, but unsupervised machine learning is used when the prediction class is unknown. Some studies were conducted to predict the diastolic function with supervised machine learning and some studies used unsupervised learning to predict the diastolic dysfunction. Unsupervised machine learning was used in the [9-11], which used echocardiography test results to make the clusters to predict diastolic function. Another study[12], also used unsupervised machine learning to predict diastolic function. Some studies use a supervised machine learning approach to predict known class and duke treadmill score that is derived after a stress echocardiography test, but in these studies[13-16],predictions were made with symptoms using supervised machine learning. Another example of a supervised learning technique performed in[17]study, the dataset used the symptoms and predicted whether the patient had heart disease or not.

Echocardiography and the Doppler test can detect diastolic function or dysfunction. In a recent study to predict diastolic dysfunction[18], they used the test data and applied a machine learning approach to get the best prediction of all four categories: normal, mild, moderate, and severe[19].

We want to predict a patient's situation with the basic features or symptoms mentioned early in this section. For this purpose, we used a unique dataset that has not previously been used in any other study. This data was taken from the local clinic of the cardiology center. The data that we got from the cardiology center was in SQL format, and the disease conclusion was in a paragraph. Symptoms that were taken from patients were in text format. We processed the data and formed it into a single table format. After data processing, data mining algorithms were applied to the dataset. The symptoms that were used by the cardiologist were predicted toward specific disease in coordination with the expert opinion of the cardiologist.

Table 1: Previously studies in term of prediction

Ref	Author	Accuracy (%)	Technique
[20]	Shafenoor et al.	95.00	Used the **K-NN** and **Decision Tree**
[21]	Devansh et al.	90.79	Used the **K-NN**
[22]	Apurb et al.	93.70	Used the **Random Forest**
[23]	Archana et al.	87.00	Used the **K-NN**
[17]	Mamun et al.	100	Used the **Random Forest**
[24]	Kavitha et al.	88	Used the **Decision Tree** & **Random Forest**
[25]	Norma et al.	98.40	Used the **HDPM**
[26]	Paul et al.	80.00	Used the **Neural network with fuzzy logic**
[27]	Verma et al.	80.68	Used the **Decision tree**
[28]	Ismaeel et al.	86.50	Used the **Extreme Machine learning**
[29]	El-Bialy et al.	78.54	Used the **Decision tree**
[30]	Subanya et al.	86.76	Used the **Support Vector Machine**
[31]	Nahar et al.	69.11	Used the **Naïve Bayes**
[32]	Tougui et al.	85.86	Used the **Artificial Neural Network**

MATERIAL & METHODS

Data Set

The dataset we used to conduct this research was collected from "The Heart Center Bahawalpur". There was a dataset of 1285 patients including echocardiography results, out of which 788 male and 497 female patients. The average age of male patients was between 31 and 100 years, and female patients' age was between 30 and 95 years. The BSA of male patients was between 0.77 and 2.92, and for female patients it was between 0.54 and 2.62. BMI of male patients' was between 10.28 and 48.82 and for female patients' it was between 13.84 and 49.58. Two features are: blood pressure (BP) systolic is between 48 and 210; blood pressure (BP) diastolic is between 19 and 180.These features are patients' basic information and initial checkup. Four features are now risk factors in patients. The first risk factor was hypertension, found in 419 patients. Where 215 were male patients and 204 were female patients. Hypertension is a very common risk factor in patients these days. Hypertension has increased mortality in heart disease patients[33, 34]. The second risk factor considered in our experiments is obesity. Obesity risk factors are considered when the patient has more weight than the normal range concerning age and height. 189 patients had an obesity risk factor, where 85 were male patients and 104 were female. Obesity has a lot of effects on diastolic dysfunction as defined in[35, 36]. Shortness of Breath (SOB) starts when the veins of the heart become fatter and blood cannot flow properly. The patient will start breathing issues after a little jogging or even walking. In our case, 426 patients had SOB risk factors, where 239 were male patients and 187 were female. Shortness of breath is the first stage to detect an issue in heart patients[37].

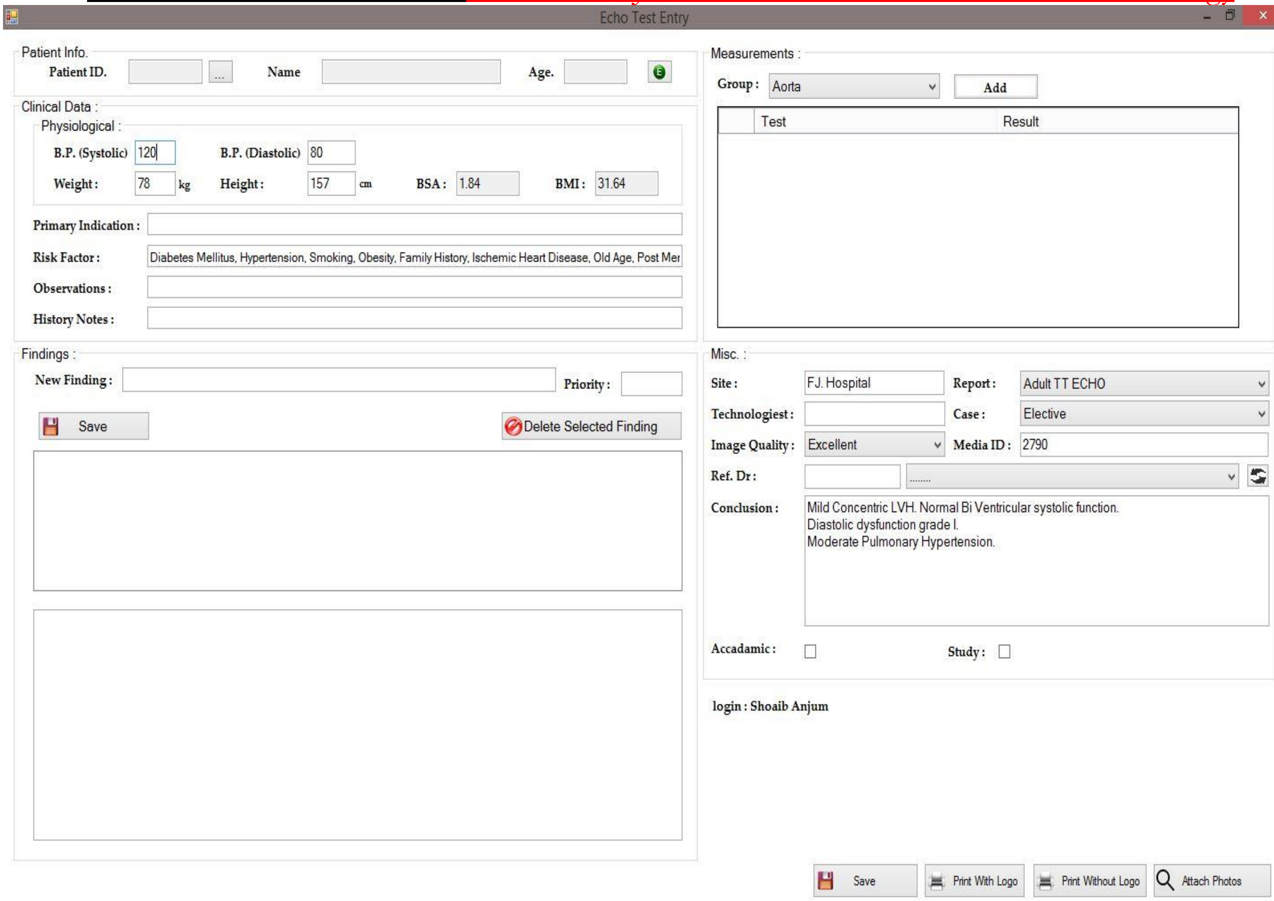


Figure 1: Interface for data collection of Cardiac Disease practically available at Heart Center Bahawalpur

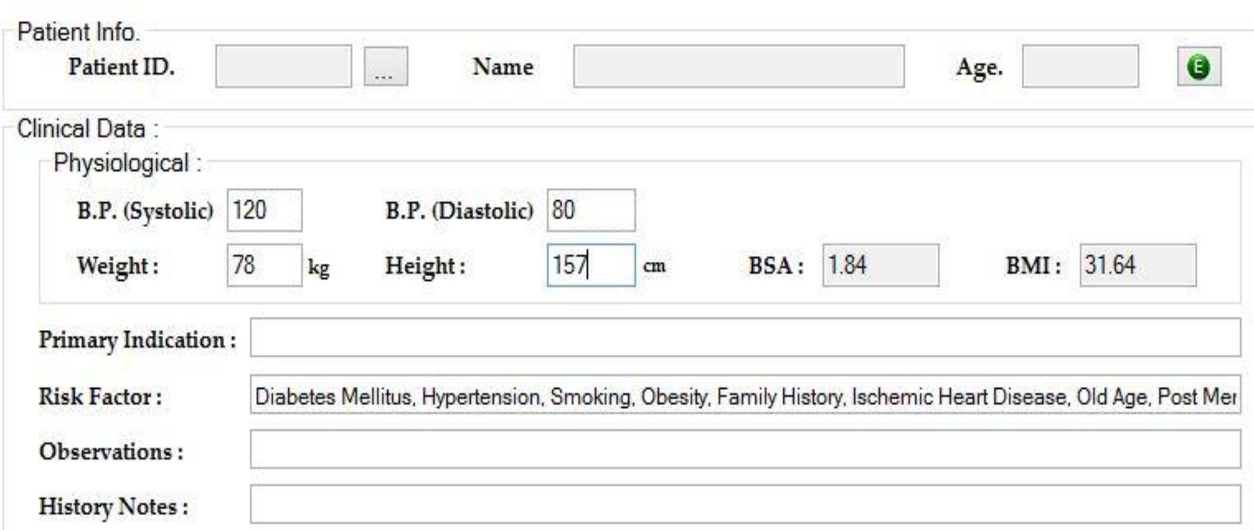


Figure 2: Patient features collection part

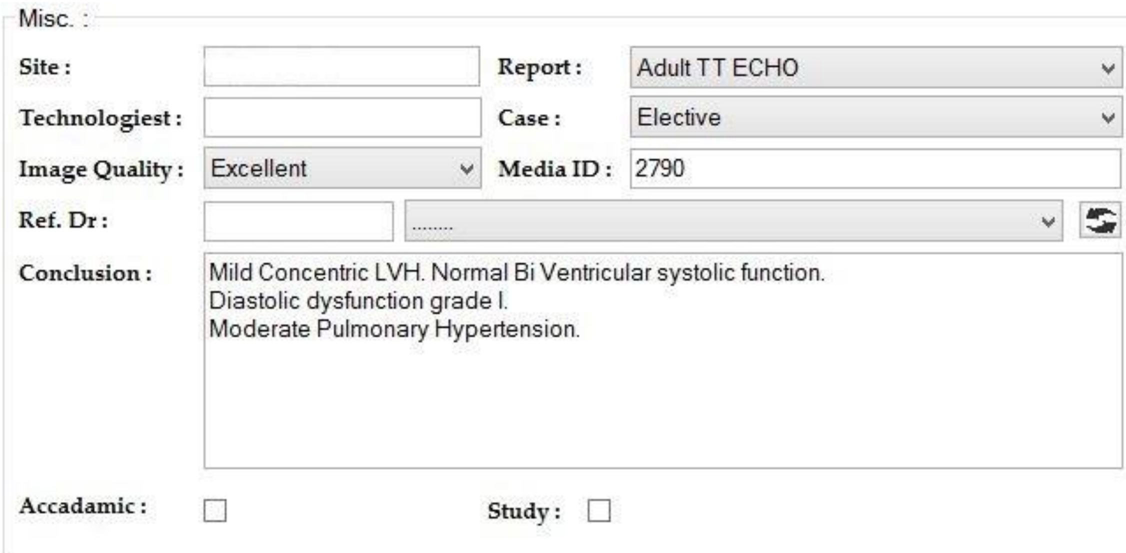


Figure 3: Patient test conclusion part

Table 2: Attribute detail

Attribute	Detail	Instances	Male	Female
Gender	Gender feature show the patient is either male or female. This features is in binary format. Male: 1; Female: 0	1285	788	497
Age	This feature represent patients' age (In years)	1285	31-100	30-95
BP Systolic	Patients' Upper bound BP	1285	59-210	48-210
BP Diastolic	Patients' lower bound BP	1285	19-180	30-154
BSA	This feature show the patients' BSA	1285	0.77-2.92	0.54-2.62
BMI	This feature shows the patient's BMI	1285	10.28-48.82	13.84-49.58
Hypertension	This Risk factor show in binary format in dataset. Either 1 or 0	419	215	204
Obesity	This feature is in binary format. Either 1 or 0	189	85	104
SOB	This risk factor is also in binary format. Either 1 or 0	426	239	187

Data Preparation

We used the WEKA tool for experiments. Initially, data was provided in SQL database format. Where risk factors and primary indications were multi-value. The class variable has been marked manually with two values, which are diastolic function and diastolic dysfunction, based on the echocardiography test conclusion provided by the cardiologist. With all these above-mentioned features, risk factors, and primary indications, we were trying to predict patients' diastolic function without performing echocardiography.

Logistic Regression Algorithm. Logistic Regression Classifier Algorithm are supervised learning algorithms. LR is an algorithm that is based on statistical techniques. LR creates a modal based on input and output variables. LR works with binary values which are specified as some are dependent and independent variables. Logistic regression is used in lots of studies in healthcare for classification purposes[38, 39].

SVM. Support Vector Machine (SVM) is a well-known supervised classification algorithm. Algorithm is useful in data analysis and pattern recognition. SVM is a mathematically based algorithm for creating a model for data analysis. The SVM algorithm is rapidly used in machine learning studies for better classification [40-42].

J.48 Algorithm. J.48 Algorithm is widely used in medical data analysis. algorithm has been previously used in many studies to predict disease using symptoms[43, 44].

Random Forest. Random Forest algorithm is a supervised algorithm widely used in data science for classification purposes. RF is a supervised classification algorithm. RF built a tree of features for decision-making [45-47].

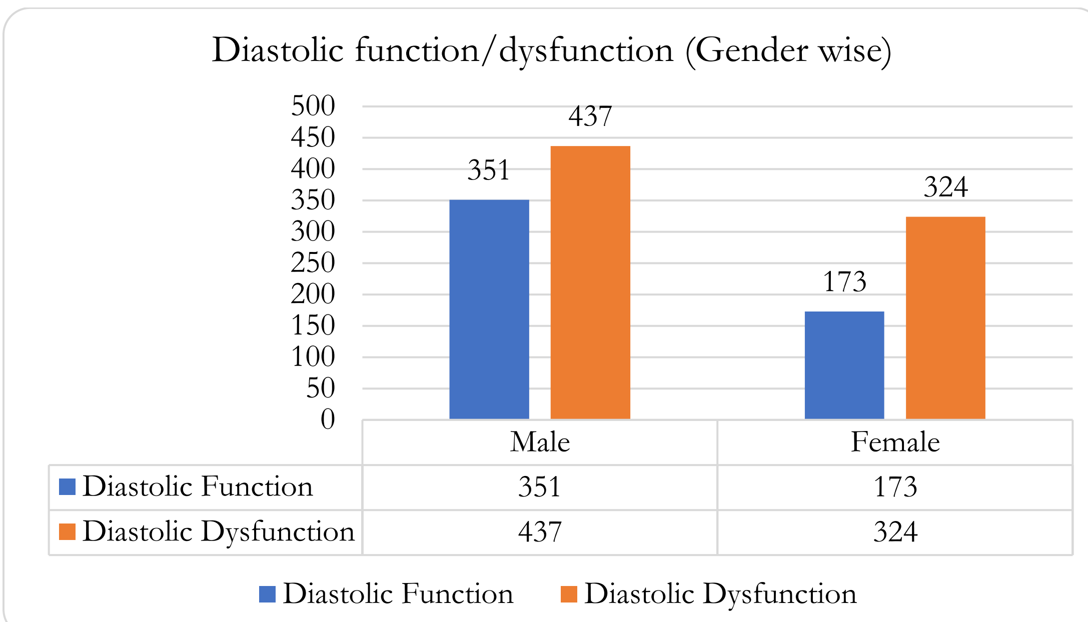
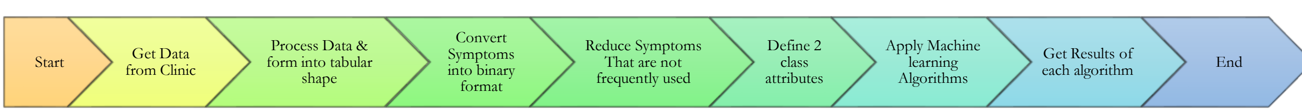


Figure 4. Diastolic function/dysfunction

Process Flow

We use the following sequence to conduct this study.



Results

Two techniques were used, split percentage and cross-validation. The split percentage technique divides the data into a given percentage of training and testing data. Train data is used for building the modal and test data is used to apply the already built modal and get the accuracy to check if the modal is worthy for further classification or not. Cross-validation

creates the folds and divides the data into each fold, checking each fold. We apply the "Replace Missing Values" filter on the data before applying any algorithm.
The results are attributed in terms of "Precision", "Recall", "F-Measure", "MCC" and "ROC Area".

Accuracy = (TP + TN) / (TP+TN+FP+FN)

Precision = TP / (TP+FP)

Recall = TP / (TP+FN)

F-Measure = 2 * Precision * Recall / (Precision + Recall)[48]

ROC Area define the test performance guide for classifications accuracy of diagnostic test based on: Excellent (90-100), Good (80-89), Fair (70-79), Poor (60-69), Fail (50 – 59).

TP = True Positive TN = True Negative

FP = False Positive FN = False Negative

Random Forest (RF)

The Random Forest algorithm was used to split the data by 70% and 30% for training and testing, respectively. The RF algorithm shows the accuracy of both class variables separately. Diastolic Dysfunction correctly identifies 196 patients and incorrectly identifies 36 patients as having diastolic dysfunction. Diastolic function was correctly identified in 131 patients, and 22 patients were identified with Diastolic Dysfunction, but these patients had diastolic function. The overall accuracy of this algorithm is 84.94%. The below table shows the detailed accuracy by class:

Table 3: Random Forest with split percentage

	RF Evaluation scores in terms of classes				
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	89.9%	84.5%	87.1%	69.2%	92.7%
D-Function	78.4%	85.6%	81.9%	69.2%	92.7%

The RF algorithm test with 10-fold cross-validation and seed value 2. As a result, diastolic dysfunction was correctly identified in 633 patients and incorrectly identified in 128 patients as diastolic function, while 412 patients with diastolic dysfunction were correctly identified and incorrectly identified in 112 patients as diastolic dysfunction. The RF algorithm with cross-validation gives us 1045 patients correctly identified out of 1285 patients. The overall performance for this algorithm is 81.32%. The below table shows the detailed accuracy by class:

Table 4: Random Forest with cross-validation

	RF Evaluation scores in terms of classes				
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	85.0%	83.2%	84.1%	61.5%	90.6%
D-Function	76.3%	78.6%	77.4%	61.5%	90.6%

J.48.

A J.48 supervised classifier was used for classification. The J.48 algorithm was used to split the data by 80% and 20% for train and test, respectively. J.48 shows the accuracy of both class variables separately. Diastolic dysfunction correctly identified 134 patients and incorrectly identified 24 patients as having diastolic dysfunction. Diastolic function was correctly identified in 85 patients, and 14 patients were incorrectly identified as diastolic dysfunction, but these are diastolic functions. Overall accuracy for the J.48 algorithm split percentage is 85.21%. The below table shows the detailed accuracy by class:

Table 5: J.48 with split percentage

J. 48 Evaluation scores in terms of classes					
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	90.5%	84.8%	87.6%	69.6%	91.20%
D-Function	78.0%	85.9%	81.7%	69.6%	91.20%

The J.48 algorithm was tested using cross-validation with 5-Fold. As a result, diastolic dysfunction was correctly identified in 633 patients and incorrectly identified in 128 patients as diastolic function, and 432 patients with diastolic dysfunction were correctly identified and incorrectly identified 92 patients as having diastolic dysfunction. The overall performance for this algorithm is 82.88%. The below table shows the detailed accuracy by class:

Table 6: J.48 with cross-validation

J. 48 Evaluation scores in terms of classes					
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	87.3%	83.2%	85.2%	65%	88.3%
D-Function	77.1%	82.4%	79.7%	65%	88.3%

Logistic Regression

A Logistic Regression supervised classifier was used to split the data by 70% and 30% for training and testing, respectively. LR shows the accuracy of both class variables separately. Diastolic dysfunction was correctly identified in 194 patients and incorrectly identified in 38 patients. Diastolic function was correctly identified in 135 patients and incorrectly identified in 18 patients. Overall accuracy for the LR algorithm is 85.45%. The below table shows the detailed accuracy by class:

Table 7: Logistic Regression with split percentage

LR Evaluation scores in terms of classes					
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D Dysfunction	91.5%	83.6%	87.4%	70.7%	93.9%
D Function	78.0%	88.2%	82.8%	70.7%	93.9%

The LR algorithm test uses cross-validation with 5-Fold. As a result, diastolic dysfunction was correctly identified in 621 patients and incorrectly identified in 140 patients. In 448 patients, diastolic function was correctly identified in 76 patients and incorrectly identified in another 76. The overall performance for this algorithm is 83.19%. The below table shows the detailed accuracy by class:

Table 8: Logistic Regression with cross-validation

LR Evaluation scores in terms of classes					
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	89.1%	81.6%	85.2%	66.2%	91.6%
D-Function	76.2%	85.5%	80.6%	66.2%	91.6%

Support Vector Machine. A SVM supervised classifier was used to split 70% and 30% for training and testing, respectively. The SVM shows the accuracy in both class variables separately. Diastolic dysfunction was correctly identified in 194 patients and incorrectly identified in 38 patients. Diastolic function was correctly identified in 133 patients and incorrectly identified in 20 patients. Overall accuracy for the SVM algorithm is 84.94%. The below table shows the detailed accuracy by class:

Table 9: Support Vector Machine with split percentage

	SVM Evaluation scores in terms of classes				
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	90.7%	83.6%	87.0%	69.5%	85.3%
D-Function	77.8%	86.9%	82.1%	69.5%	85.3%

The SVM algorithm was used for the cross-validation with 5-Fold and random seed 3. As a result, diastolic dysfunction was correctly identified in 618 patients and incorrectly identified in 143 patients. In 444 patients, diastolic function was correctly identified in 80 patients and incorrectly identified in 80 others. The overall performance for this algorithm is 82.65%. The below table shows the detailed accuracy by class:

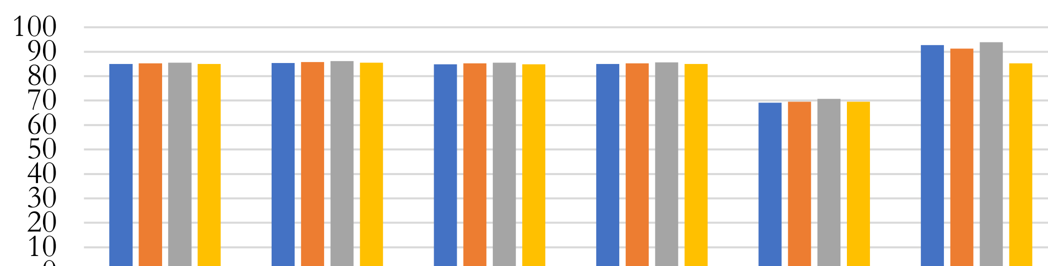
Table 10: Support Vector Machine with cross-validation

	SVM Evaluation scores in terms of classes				
	Precision	**Recall**	**F-Measure**	**MCC**	**ROC Area**
D-Dysfunction	88.5%	81.2%	84.7%	65.1%	83.0%
D-Function	75.6%	84.7%	79.9%	65.1%	83.0%

DISCUSSION

Using both techniques, split percentage and cross-validation, the best performance of each algorithm is listed below in Table 12 and Table 13 respectively. Each algorithm generates approximately equal results, but the highest accuracy originates from the Logistic Regression Algorithm using split percentage, which is 85.45%. We compared every result in two tables. First with the split percentage of each algorithm with detail accuracy and second with cross-validation with class detail accuracy.

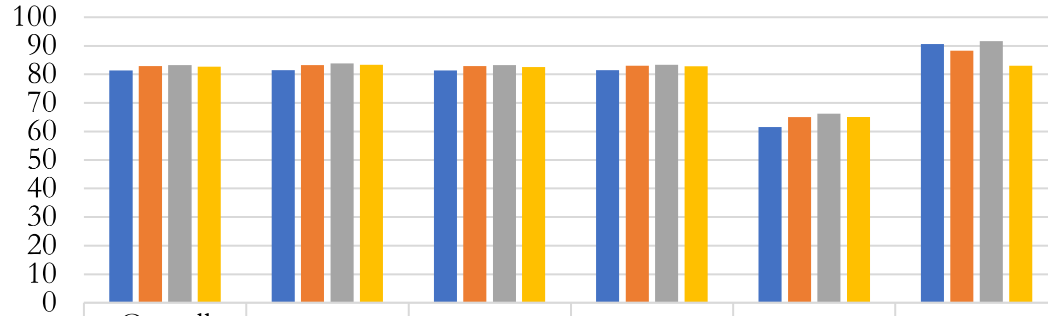
Table 11: All algorithms accuracy by split percentage



	Overall Accuracy	Precision	Recall	F-Measure	MCC	ROC Area
RF	84.94	85.4	84.9	85	69.2	92.7
J.48	85.21	85.7	85.2	85.3	69.6	91.2
LR	85.45	86.2	85.5	85.6	70.7	93.9
SVM	84.94	85.5	84.9	85	69.5	85.3

RF J.48 LR SVM

Table 12: All algorithm accuracy with cross-validation



	Overall Accuracy	Precision	Recall	F-Measure	MCC	ROC Area
RF	81.32	81.4	81.3	81.4	61.5	90.6
J.48	82.88	83.2	82.9	83	65	88.3
LR	83.19	83.8	83.2	83.3	66.2	91.6
SVM	82.65	83.3	82.6	82.8	65.1	83

RF J.48 LR SVM

Now, the statistics of each algorithm are shown in the below tables. Table 13 shows the stats of each algorithm's split percentage technique used. Table 14 shows the stats of each algorithm stats performed with cross-validation:

Table 13: Algorithm stats with split percentage

All classifiers' stats with split percentage technique		
	Kappa Statistic	**Mean Absolute Error**
RF	0.6903	0.2101
J.48	0.6936	0.2162
LR	0.7029	0.2107
SVM	0.6916	0.1506

Table 14: Algorithm stats with cross-validation

All classifiers' stats with cross validation technique		
	Kappa Statistic	**Mean Absolute Error**
RF	0.6151	0.2269
J.48	0.6493	0.2291
LR	0.6585	0.2274
SVM	0.6473	0.1735

CONCLUSION

This work has analyzed the role of machine learning in the medical field for disease prediction (diastolic function or dysfunction) before performing the medical test. In this study, we take some results of an echocardiography test, which has some features, and lastly, a conclusion which is suggested by the cardiologist. Using a machine learning approach, we check whether the patient's described symptoms are related to the resultant disease or not. In this process, we first get the data from the cardiologist, process it, and then form it into our useful shape. Some features related to patient basic data like age, gender, BP, BSA, BMI, and the patient's description of some symptoms or history like hypertension, obesity, and shortness of breath. We have converted these features into columns and then replaced the value with binary. Then, machine learning algorithms were used to examine the relationship between features and their outcomes. After applying machine learning to patients' symptoms and conclusion data of diastolic function and diastolic dysfunction classification, the logistic regression algorithm gives us the best accuracy, up to 85%. It means cardiologists asking for symptoms from patients are 85% enough toward the conclusion. In a further study, features and some other algorithms' usage will increase this accuracy, which will be more beneficial for cardiologists and be time-efficient.

REFERENCES

[1] Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., and Singh, P.: "*Prediction of heart disease using a combination of machine learning and deep learning*", "Computational intelligence and neuroscience", 2021, 2021

[2] Akinosun, A.S., Polson, R., Diaz-Skeete, Y., De Kock, J.H., Carragher, L., Leslie, S., Grindle, M., and Gorely, T.: "*Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis*", "JMIR mHealth and uHealth", 2021, 9, (3), pp. e21061

[3] Dandel, M., and Hetzer, R.: "*Ventricular systolic dysfunction with and without altered myocardial contractility: clinical value of echocardiography for diagnosis and therapeutic decision-making*", "International Journal of Cardiology", 2021, 327, pp. 236-250

[4] Biondi, B., Fazio, S., Palmieri, E.A., Carella, C., Panza, N., Cittadini, A., Bonè, F., Lombardi, G., and Saccà, L.: "*Left ventricular diastolic dysfunction in patients with subclinical hypothyroidism*", "The Journal of Clinical Endocrinology & Metabolism", 1999, 84, (6), pp. 2064-2067

[5] https://www.upmc.com/services/pulmonology/conditions/diastolic-dysfunction#:~:text=When%20the%20muscles%20of%20the,back%20up%20in%20the%20organs., accessed 06-Mar-2022

[6] Gaasch, W.H., and Zile, M.R.: "*Left ventricular diastolic dysfunction and diastolic heart failure*", "Annu. Rev. Med.", 2004, 55, pp. 373-394

[7] Jung, Y.-H., Ren, X., Suffredini, G., Dodd-o, J.M., and Gao, W.D.: "*Right ventricular diastolic dysfunction and failure: a review*", "Heart Failure Reviews", 2021, pp. 1-14

[8] Lopez, E.O., Ballard, B.D., and Jan, A.: 'Cardiovascular disease': 'StatPearls [Internet]' (StatPearls Publishing, 2021)

[9] Pandey, A., Kagiyama, N., Yanamala, N., Segar, M.W., Cho, J.S., Tokodi, M., and Sengupta, P.P.: "*Deep-learning models for the echocardiographic assessment of diastolic dysfunction*", "Cardiovascular Imaging", 2021, 14, (10), pp. 1887-1900

[10] Burrello, J., Burrello, A., Vacchi, E., Bianco, G., Caporali, E., Amongero, M., Airale, L., Bolis, S., Vassalli, G., and Cereda, C.W.: "*Supervised and unsupervised learning to define the cardiovascular risk of patients according to an extracellular vesicle molecular signature*", "Translational Research", 2022

[11] Flores, A.M., Schuler, A., Eberhard, A.V., Olin, J.W., Cooke, J.P., Leeper, N.J., Shah, N.H., and Ross, E.G.: "*Unsupervised Learning for Automated Detection of Coronary Artery Disease Subgroups*", "Journal of the American Heart Association", 2021, 10, (23), pp. e021976

[12] Ma, K., Canepa, M., Strait, J.B., and Shatkay, H.: 'Using unsupervised learning to determine risk level for left ventricular diastolic dysfunction', in Editor (Ed.)^(Eds.): 'Book

Using unsupervised learning to determine risk level for left ventricular diastolic dysfunction' (IEEE, 2014, edn.), pp. 344-347

[13] Anjum, M.S., Mumtaz, S., Riaz, O., and Sharif, W.: "*Heart Attack Risk Prediction with Duke Treadmill Score with Symptoms using Data Mining*", "International Journal of Innovations in Science & Technology", 2021, 3, (4), pp. 174-185

[14] Al-Yarimi, F.A.M., Munassar, N.M.A., Bamashmos, M.H.M., and Ali, M.Y.S.: "*Feature optimization by discrete weights for heart disease prediction using supervised learning*", "Soft Computing", 2021, 25, (3), pp. 1821-1831

[15] Patro, S.P., Nayak, G.S., and Padhy, N.: 'Heart disease prediction by using novel optimization algorithm: A supervised learning prospective', in Editor (Ed.)^(Eds.): 'Book Heart disease prediction by using novel optimization algorithm: A supervised learning prospective' (Elsevier, 2021, edn.), pp. 100696

[16] Patro, S.P., Padhy, N., and Chiranjevi, D.: "*Ambient assisted living predictive model for cardiovascular disease prediction using supervised learning*", "Evolutionary Intelligence", 2021, 14, (2), pp. 941-969

[17] Ali, M.M., Paul, B.K., Ahmed, K., Bui, F.M., Quinn, J.M., and Moni, M.A.: "*Heart disease prediction using supervised machine learning algorithms: performance analysis and comparison*", "Computers in Biology and Medicine", 2021, 136, pp. 104672

[18] Singh, A., Sun, D., Mor-Avi, V., Addetia, K., Patel, A.R., DeCara, J.M., Ward, R.P., and Lang, R.M.: "*Can echocardiographic assessment of diastolic function be automated?*", "The International Journal of Cardiovascular Imaging", 2021, pp. 1-10

[19] Kameshima, H., Uejima, T., Fraser, A.G., Takahashi, L., Cho, J., Suzuki, S., Kato, Y., Yajima, J., and Yamashita, T.: "*A Phenotyping of Diastolic Function by Machine Learning Improves Prediction of Clinical Outcomes in Heart Failure*", "Frontiers in cardiovascular medicine", 2021, 8

[20] Amin, M.S., Chiam, Y.K., and Varathan, K.D.: "*Identification of significant features and data mining techniques in predicting heart disease*", "Telematics and Informatics", 2019, 36, pp. 82-93

[21] Shah, D., Patel, S., and Bharti, S.K.: "*Heart disease prediction using machine learning techniques*", "SN Computer Science", 2020, 1, (6), pp. 1-6

[22] Rajdhan, A., Agarwal, A., Sai, M., Ravi, D., and Ghuli, P.: "*Heart disease prediction using machine learning*", "International Journal of Research and Technology", 2020, 9, (04), pp. 659-662

[23] Singh, A., and Kumar, R.: 'Heart disease prediction using machine learning algorithms', in Editor (Ed.)^(Eds.): 'Book Heart disease prediction using machine learning algorithms' (IEEE, 2020, edn.), pp. 452-457

[24] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y.R., and Suraj, R.S.: 'Heart disease prediction using hybrid machine learning model', in Editor (Ed.)^(Eds.): 'Book Heart disease prediction using hybrid machine learning model' (IEEE, 2021, edn.), pp. 1329-1333

[25] Fitriyani, N.L., Syafrudin, M., Alfian, G., and Rhee, J.: "*HDPM: an effective heart disease prediction model for a clinical decision support system*", "IEEE Access", 2020, 8, pp. 133034-133050

[26] Paul, A.K., Shill, P.C., Rabin, M.R.I., and Akhand, M.: 'Genetic algorithm based fuzzy decision support system for the diagnosis of heart disease', in Editor (Ed.)^(Eds.): 'Book Genetic algorithm based fuzzy decision support system for the diagnosis of heart disease' (IEEE, 2016, edn.), pp. 145-150

[27] Verma, L., Srivastava, S., and Negi, P.: "*A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data*", "Journal of medical systems", 2016, 40, (7), pp. 1-7

[28] Ismaeel, S., Miri, A., Sadeghian, A., and Chourishi, D.: 'An extreme learning machine (ELM) predictor for electric arc furnaces' vi characteristics', in Editor (Ed.)^(Eds.): 'Book An extreme learning machine (ELM) predictor for electric arc furnaces' vi characteristics' (IEEE, 2015, edn.), pp. 329-334
[29] El-Bialy, R., Salamay, M.A., Karam, O.H., and Khalifa, M.E.: "*Feature analysis of coronary artery heart disease data sets*", "Procedia Computer Science", 2015, 65, pp. 459-468
[30] Subanya, B., and Rajalaxmi, R.: 'Feature selection using Artificial Bee Colony for cardiovascular disease classification', in Editor (Ed.)^(Eds.): 'Book Feature selection using Artificial Bee Colony for cardiovascular disease classification' (IEEE, 2014, edn.), pp. 1-6
[31] Nahar, J., Imam, T., Tickle, K.S., and Chen, Y.-P.P.: "*Computational intelligence for heart disease diagnosis: A medical knowledge driven approach*", "Expert systems with applications", 2013, 40, (1), pp. 96-104
[32] Tougui, I., Jilbab, A., and El Mhamdi, J.: "*Heart disease classification using data mining tools and machine learning techniques*", "Health and Technology", 2020, 10, (5), pp. 1137-1144
[33] Ginelli, P., and Bella, J.: "*Treatment of diastolic dysfunction in hypertension*", "Nutrition, Metabolism and Cardiovascular Diseases", 2012, 22, (8), pp. 613-618
[34] Schillaci, G., Pasqualini, L., Verdecchia, P., Vaudo, G., Marchesi, S., Porcellati, C., de Simone, G., and Mannarino, E.: "*Prognostic significance of left ventricular diastolic dysfunction in essential hypertension*", "Journal of the American College of Cardiology", 2002, 39, (12), pp. 2005-2011
[35] Powell, B.D., Redfield, M.M., Bybee, K.A., Freeman, W.K., and Rihal, C.S.: "*Association of obesity with left ventricular remodeling and diastolic dysfunction in patients without coronary artery disease*", "The American journal of cardiology", 2006, 98, (1), pp. 116-120
[36] Berkalp, B., Cesur, V., Corapcioglu, D., Erol, C., and Baskal, N.: "*Obesity and left ventricular diastolic dysfunction*", "International journal of cardiology", 1995, 52, (1), pp. 23-26
[37] Ilia, R., Carmel, S., Carlos, C., and Gueron, M.: "*Relation between shortness of breath, left ventricular end diastolic pressure and severity of coronary artery disease*", "International journal of cardiology", 1995, 52, (2), pp. 153-155
[38] Li, J.P., Haq, A.U., Din, S.U., Khan, J., Khan, A., and Saboor, A.: "*Heart disease identification method using machine learning classification in e-healthcare*", "IEEE Access", 2020, 8, pp. 107562-107582
[39] Issitt, R.W., Cortina-Borja, M., Bryant, W., Bowyer, S., Taylor, A.M., and Sebire, N.: "*Classification Performance of Neural Networks Versus Logistic Regression Models: Evidence From Healthcare Practice*", "Cureus", 2022, 14, (2)
[40] Razzaghi, T., Roderick, O., Safro, I., and Marko, N.: "*Multilevel weighted support vector machine for classification on healthcare data with missing values*", "PloS one", 2016, 11, (5), pp. e0155119
[41] Ismail, A., Abdlerazek, S., and El-Henawy, I.M.: "*Development of smart healthcare system based on speech recognition using support vector machine and dynamic time warping*", "Sustainability", 2020, 12, (6), pp. 2403
[42] Lai, C.Q., Ibrahim, H., Abdullah, J.M., Azman, A., and Abdullah, M.Z.: "*Convolutional neural network utilizing error-correcting output codes support vector machine for classification of non-severe traumatic brain injury from electroencephalogram signal*", "IEEE Access", 2021, 9, pp. 24946-24964
[43] Huang, F., Wang, S., and Chan, C.-C.: 'Predicting disease by using data mining based on healthcare information system', in Editor (Ed.)^(Eds.): 'Book Predicting disease by using data mining based on healthcare information system' (IEEE, 2012, edn.), pp. 191-194

[44] Zhang, Y., Zhang, Y., Lo, B., and Xu, W.: "*Wearable ECG signal processing for automated cardiac arrhythmia classification using CFASE-based feature selection*", "Expert Systems", 2020, 37, (1), pp. e12432

[45] Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., and Yan, J.: "*Study of cardiovascular disease prediction model based on random forest in eastern China*", "Scientific reports", 2020, 10, (1), pp. 1-8

[46] Elola, A., Aramendi, E., Irusta, U., Del Ser, J., Alonso, E., and Daya, M.: "*ECG-based pulse detection during cardiac arrest using random forest classifier*", "Medical & biological engineering & computing", 2019, 57, (2), pp. 453-462

[47] Javeed, A., Zhou, S., Yongjian, L., Qasim, I., Noor, A., and Nour, R.: "*An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection*", "IEEE Access", 2019, 7, pp. 180235-180243

[48] Alaoui, S.S., Labsiv, Y., and Aksasse, B.: "*Classification algorithms in data mining*", "Int. J. Tomogr. Simul", 2018, 31, pp. 34-44