Comparative Study on Machine Learning Techniques ANN and SVM for Lymphoma Malady Precision and Hazard Factors

N. Sivaranjini
Asst. Professor,
Department of Computer Science,
Auxilium College (Autonomous),
Vellore, Tamil Nadu, India
email: sathiya.siva5@gmail.com

M. Gomathi
Asst. Professor,
Department of Computer Science,
Government Arts and Science College, Kumarapalayam,
Erode, Tamil Nadu, India
email: mdgomathi@gmail.com

***Abstract* - As indicated by the pervasiveness of coronary illness and the cost of conclusion and treatment strategies other than having such huge numbers of symptoms, presence of certain techniques, for example, information mining which makes doctors ready to have an exact expectation about the danger of an infection among patients with various characteristics is so significant, in light of the fact that it very well may be efficient, affordable and furthermore decline determination focuses clog. Another impact of applying these sorts of techniques is anticipation of patients from being affected by reactions brought about by some finding strategies, for example, angiography. Today information mining is being utilized in numerous fields of science including clinical science and one of its applications is infections expectation dependent on past encounters and datasets. Right now, are going to address a kind of coronary illness named Lymphoma Artery Disease (LAD) brought about by cholesterol sedimentation in primary veins and making a deterrent for blood development in those vessels. So as to address this issue, we initially have a concise presentation about LAD then we present a few patient order and malady expectation techniques. In the accompanying, we enroll a portion of these models to have an expectation on a dataset having a place with Cleveland which is a city in US and afterward we contrast these models with pick the most able one for forecast utilizing ROC bend and measurable techniques. At last, we figure their precision to pick the best model for this issue.**

Keywords - Lymphoma artery disease (LAD), Data mining algorithms, artificial neural network (ANN), Support vector machine (SVM).

1. INTRODUCTION

Lymphoma diseases are among the common diseases in both developed and creating nations and viewed as the primary driver of death all through the world (Healthy environment, healthy heart (Internet). Jakarta: Ministry of Health, Republic of Indonesia, 2014) Truth be told, any condition or infection that influences the heart, its vessels (Montalescot G, 2013), and the blood circulatory framework can be identified with lymphoma vascular illnesses (LADs) (Kelly BB, 2010). As a rule, the clinical range of LADs ranges from asymptomatic ischemia to interminable stable angina pectoris, insecure angina (UA), intense myocardial dead tissue (AMI), ischemic cardiomyopathy and abrupt demise (Sanchis-Gomar F, 2016). They are now and again connected with conditions, for example, hypertension, stroke, lymphoma supply route sicknesses, ceaseless cardiovascular breakdown, inborn coronary illness, musicality issue, sub clinical atherosclerosis, valvar malady, and fringe blood vessel ailment (Mozaffarian D, 2016). Lately, notwithstanding the fundamental hazard factors, different factors, for example, contamination, incendiary and interminable infections have been talked about as other hazard components of Lymphoma illnesses (Bergh C, 2017).

Toward the start of the twentieth century, 10% of the considerable number of passing was credited to Lymphoma ailments. Toward the finish of this century, the mortality brought about by LADs expanded to 25%. It is evaluated that, thinking about the present expanding pattern, more than 35–60% of passing worldwide would be because of Lymphoma maladies by 2025 (Longo D, 2011). In light of the report by WHO, in 2017, the greater part (54%) of the passing around the globe were brought about by 10 driving causes, and Lymphoma maladies which prompted 15 million passing in 2015 comprised the biggest gathering of deadly sicknesses (world Health Organization. Physical activity and older adults, 2018). Lymphoma infections execute a great many individuals every year and this worth might be expanded up to 24.8 million by 2020 if preventive measures are not taken (The top ten causes of death. Geneva: World Health Organization, 2017).

In Iran, the Ministry of Health revealed that 39.9% of the death rate in the nation is because of Lymphoma illnesses and their hazard factors, among them LAD is the most pervasive sort and is incredibly expanding (Mann DL, 2014). Fellow is a multi-causal ailment, in which a progression of hazard factors, for example

expanded cholesterol, hypertension, diabetes and smoking ought to be considered (U.S. National Library of medicine (NLM). Lymphoma artery disease (LAD) medlineplus, 2017). As indicated by the consequences of an epidemiological examination with the point of inspecting lymphoma vein sickness death rate, 63 out of 6537 passing cases were because of LAD in 2015 (Kamiya K, 2015). Fellow is more predominant among men than ladies, and the side effects of the ailment may show up in ladies 10 years later than in men (Charchar FJ, 2012). Along these lines, considering the extraordinary increment in Lymphoma maladies which forces a substantial budgetary weight on the general public, clinical networks endeavor to discover a route for the precise and convenient forecast of LAD by utilizing new factual methods, for example, information mining (Rezaei-hachesu P, 2013). It is vital that the medicinal services area is loaded up with information. Be that as it may, the information required for compelling dynamic and disclosure of shrouded designs are not separated. By extricating helpful information and finding information from the enormous volume of clinical information, the reasons for rate, development or the spread of maladies can be distinguished and doctors can be furnished with significant data for better dynamic. Along these lines, numerous social insurance habitats are looking for down to earth answers for information disclosure by methods for information mining strategies (Sudhakar K, 2014) These strategies can assist with perceiving the examples and components affecting infections (Yeh DY, 2011). The tale study of information mining is among the 10 creating sciences which have made the following LAD face tremendous innovative advancements. Utilizing specific information, it will have broad applications in the area of medication. (Bellazzi R, 2011) (Amato F, 2013)

The writing audit demonstrated that various calculations, for example, bunching, orders, relapse and affiliation rules, choice trees, Bayesian system, neural system, multi-layer perceptron with blunder back engendering calculation, scaled conjugate inclination (SCG) and bolster vector machine (SVM) have been utilized for anticipating LAD (Sivagowry S, 2013), (Sufi F, 2011), (Amin SU, 2013), (Desai Shrinivas D., 2018), (Kausar N, 2016), (Abawajy JH, 2013), (Zhou X, 2010), (Guner LA, 2010), (Orphanou K, 2016), (Kim J, 2015), (Karaolis MA, 2010), (Verma L, 2016), (Das R, 2009). Be that as it may, the examination between the calculations has not gotten sufficient consideration. Among these calculations, fake neural system has a few points of interest, for example, rapid, straightforwardness and capacity of unraveling complex connections among factors and extricating the non-direct connections by methods for preparing information. Another calculation is bolster vector machine which is the most widely recognized and successful AI calculation. SVMs have an amazing hypothetical foundation that utilized in various exercises, for example, grouping, acknowledgment and forecast in administered learning (Acharya UR, 2017), (Dolatabadi AD, 2017). Along these lines, the present examination meant to look at the PPV of LAD utilizing fake neural system (ANN) and SVM calculations and their differentiation as far as anticipating LAD in the chose emergency clinics.

2. LITERATURE SURVEY

2.1. Study structure and setting

The present research was directed utilizing information mining systems. The exploration setting was three chosen emergency clinics subsidiary to AJA University of Medical Sciences.

2.2. Members and examining

Right now, clinical records of patients with lymphoma conduit sickness who were hospitalized in three showing medical clinics between March 2016 and March 2017 were utilized (n = 1324). Different infections, for example, arrhythmia, angina pectoris, intense myocardial dead tissue, ceaseless rheumatic heart sicknesses, inherent coronary illness, enlarged cardiomyopathy, cardiovascular breakdown, hypertrophic cardiomyopathy, hypertensive heart maladies, ischemic heart ailments, myocardial localized necrosis, mitral spewing forth, mitral valve prolapse, pneumonic stenosis, and aspiratory coronary illness were avoided. A special dataset including similar LAD anticipating factors was utilized for both SVM and ANN methods.

2.3. Instruments

The information assortment instrument was an agenda structured dependent on the factors utilized in the rule of the Cleveland coronary illness dataset strategy in UCI (University of California) store. (Janosi A, 1988) The agenda included 25 factors for foreseeing LAD. These factors were sexual orientation, age, weight, conjugal status, occupation, address, family ancestry, smoking, comorbidity, diabetes, beat rate, T.S.T waves, hypertension (HBP), cholesterol, triglyceride (TG), hemoglobin (Hgb), blood glucose level, creatinine, systolic circulatory strain, diastolic pulse, chest torment, low thickness lipoprotein (LDL), high thickness lipoprotein (HDL), LAD finding, and the length of hospitalization. The gathered information was constrained by utilizing various techniques, for example, information readiness, coordination, cleaning, standardization and decrease.

2.4. Measurable investigation

After standardization, handling and purging, information were gone into SPSS (V23.0) and Microsoft Excel 2013. In addition, R 3.3.2 was utilized for measurable figuring. The dataset was isolated into preparing and

testing sets and to do as such, the standard randomized distribution strategy was utilized. Therefore, 70% of the records were utilized for preparing and 30% was utilized for testing the models.

2.5. Moral thought

The examination convention was endorsed by the Ethical Clearance Committee of AJA University of Medical Sciences. The information was utilized namelessly and was kept secret.

2.6. Socio-segment indicators of LAD

The recurrence dissemination of LAD and the kind of occupation indicated that there was a critical distinction among having and not having the occupation. What's more, the recurrence appropriation of LAD and the spot of home proposed that most of the patients ($n = 1082$, 98%) dwelled in urban areas. So also, the recurrence dissemination of LAD and a family ancestry demonstrated that there was a huge distinction among having and not having a family ancestry of LAD ($n = 1049$, 79.3%) ($p < 0.001$). Also, the outcomes demonstrated that 77.3% ($n = 1024$) of patients were non-smokers and there was a critical contrast ($p < 0.001$) among the emergency clinics as far as smoking and LAD.

2.7. The anticipating factors

The principle target of this examination was to decide the PPV of LAD utilizing ANN calculation and contrast the outcomes and the consequences of the SVM model. Along these lines, 25 anticipating factors were separated from the database of Lymphoma patients in the chose clinics and were utilized as the info factors and the heaviness of every wa determined by running calculations so as to fit the multi-layer ANN model (Fig) Based on the determined loads, the accompanying factors were chosen as LAD foreseeing factors: sex, occupation, spot of habitation, family ancestry, smoking status, comorbidity, mean estimation of heartbeat rate, TST waves status, hypertension history, chest torment, cholesterol, triglyceride, blood glucose level and creatinine level.

In the present examination, 70% of the information was utilized for preparing and 30% was utilized for testing the ANN model. The outcomes uncovered that the integrity of fit was suitable in ANN model with the PPV (Available from: https://en.wikipedia.org/wiki/Positive_and_negative_predictive_values) of 0.798, the littler mean squared blunder (MSE) and relative mistake in the test dataset is shown in equation (1).

$$PPV = \frac{number\ of\ true\ positive}{number\ of\ true\ positives + number\ of\ false\ positives} = \frac{number\ of\ true\ positives}{number\ of\ positive\ calls} \quad (1)$$

Table 1. MSE predictive value

	Sample	MSE	Relative error	Positive Predictive value
Training	70%	5.39	0.002	0.798
Testing	30%	3.84	0.002	0.202

Table 1 shows the MSE predictive value. Figure 1 delineates the recipient working trademark (ROC) bend for LAD patients. The PPV of the model relies upon the degree, to which the test has effectively recognized LAD patients (affectability). This PPV is determined by processing the region under the ROC bend. The closer this worth is to 1, the higher the PPV of the model. Additionally, the closer the estimation of this proportion is to one side corner, the bigger the territory under the bend would be. The outcomes indicated that the ANN model had high PPV while anticipating LAD.

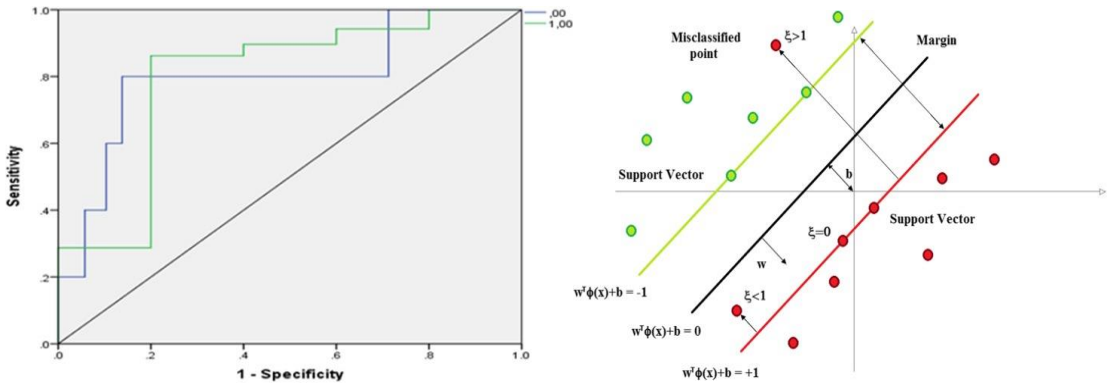

Figure 1. Sensitivity Information

In this phase, 70% of the data were considered as training data and the remaining 30% was used as test data to run the SVM algorithm. Then, PPV of the model is presented in Table 2.

$$Cohen'skappa\ coefficient = (Accuracy - expAccuracy/(1 - expAccuracy) \quad (2)$$

$$F - measure = 2 * ((PPV * Recall)/(PPV + Recall)) \quad (3)$$

Table 2. Measure calculation

	Sample	F.Measure	Kappa Coefficent	Positive Predictive value
Traning	70%	0.761	0.706	0.871
Tesiting	30%	0.696	0.636	0.129

Table 2 shows, F-measure and Cohen's Kappa coefficients were utilized to decide the PPV of the SVM model. The outcome demonstrated that the SVM model had a moderate to high power and affectability for anticipating LAD patients. Additionally, the SVM model had higher PPV in grouping and foreseeing LAD. Moreover, correlation between the precision files indicated that, the SVM model had higher exactness contrasted with the ANN model and introduced better characterization. That the discoveries likewise indicated that the zone under the ROC bend was bigger in the SVM model than in the ANN model (Fig. 4). Thus, SVM would do well to execution in anticipating patients with LAD.

In light of the outcomes, the most significant variables influencing the occurrence of LAD were sexual orientation, occupation, family ancestry, smoking, co-dismalness, mean estimation of pulse, TST wave status, hypertension, chest torment, cholesterol, triglyceride, blood glucose level and creatinine. Likewise, past examinations presented various components influencing the malady and the advancement of Lymphoma illnesses. These variables were partitioned into six general gatherings: natural elements, day by day propensities, chance components, fundamental illnesses, mental-character elements and social elements (Steenman M, 2017). Other basic hazard factors related with LAD incorporate hypertension, way of life (Bayturan O, 2010), elevated level of cholesterol (Nicholls SJ, 2010), diabetes (Murthy VL, 2012), heftiness (Perk J) and smoking (Frey P, 2011).

The after effects of the present investigation demonstrated that the occurrence of the infection was higher in men than ladies, and the danger of LAD could increment by an expansion in age and weight. Correspondingly, as indicated by another investigation, age, sexual orientation (male) and smoking had critical connections with LAD. In the examination led by Masethe and Masethe, a framework was proposed for foreseeing coronary failure and incorporated the factors of sexual orientation, age, sort of chest torment, pulse, cholesterol, smoking, blood glucose level, circulatory strain, diet and liquor utilization.

The discoveries uncovered that the danger of LAD was higher among the utilized patients contrasted with the jobless and the resigned ones. Also, the consequences of a companion study spoke to that the danger of Lymphoma infections was about 40%, due to work strain, and an expansion in remaining task at hand multiplied the danger of these ailments. In this way, the sort of employment can be a hazard factor for Lymphoma sicknesses (Kausar N, 2016). In addition, the occurrence of the infection was higher among the individuals who were living in the urban than the country zones. In another investigation, kermani et al. analyzed the connection between the death rate brought about by Lymphoma and constant obstructive pneumonic infections (COPD) because of nitrogen dioxide air poisons in Tehran and revealed a critical connection between these hazard factors. Another examination researched the connection between spatial scattering of particulate issue and mortality among patients with Lymphoma ailments in Beijing and announced that an expansion in particulate issue expanded the pace of death among those dwelling in urban communities. In another task, scientists assessed the danger of death via air contamination in 10 urban communities in Canada and found that there were huge connections between mortality among patients with Lymphoma respiratory sicknesses, urban habitation and urban air toxins. In any case, the aftereffects of another examination indicated that Lymphoma programs have not been executed in the provincial regions; along these lines, the death rate brought about by Lymphoma maladies were expanded in the rustic regions contrasted with the enormous urban areas.

The outcomes likewise uncovered that there was a critical contrast between family ancestry and LAD. 193(20.2%) and 215(22.5%) patients had fatherly and maternal positive family ancestry (father, mother and kin) of LAD; there was a likelihood to be determined to have the illness before 55 and 65 years old in people, separately. As referenced previously, family ancestry of the sickness and other hazard factors, for example, blood glucose level, HDL, LDL, cholesterol, systolic and diastolic pulse just as age and sexual orientation have been featured in the writing.

As indicated by the aftereffects of the present examination, just 58 of the members were smokers and 142 were non-smokers. The aftereffects of a Chi-squared (X2) test indicated that mean plasma levels of NO was altogether lower in smoker patients ($P = 0.004$). As indicated by the writing, smoking has an expanding pattern in Asian nations contrasted with the remainder of the world. Essentially, another examination demonstrated that heftiness, family ancestry, co-morbidities and smoking can build the danger of LAD. As smoking is a solid and free hazard factor for Lymphoma maladies, all patients with these illnesses must quit smoking. Specialists accentuate that the danger of LAD can be impressively decreased in future by constraining smoking.

Accordingly, the status of smoking must be deliberately assessed in patients with Lymphoma maladies (Perk J). As indicated by the consequences of a clinic based observational investigation, there is an immediate relationship between the smoking status and LAD among the youthful grown-ups. As a rule, the occurrence of LAD had a higher mean an incentive among smokers and the time of patients was lower than or equivalent to 35 years old (Charchar FJ, 2012).

As per the outcomes, 28.6% of the patients had one or numerous co-morbidities. In another examination, the outcomes demonstrated that patients with ischemic coronary illness (IHD) and ceaseless obstructive aspiratory sickness had the most extreme difficulties contrasted with those with just one of the prominent maladies (Montalescot G, 2013). Besides, as indicated by the consequences of different investigations; stoutness, hypertension, diabetes mellitus, metabolic disorder, elevated levels of LDL, low degree of HDL, high fat eating regimen, absence of ordinary exercise and dyslipidemia are the hazard factors for the referenced maladies (Kamiya K, 2015), (Charchar FJ, 2012), (Rezaei-hachesu P, 2013), (Sudhakar K, 2014).

As far as the connection between the mean estimation of pulse and the frequency of LAD, a noteworthy relationship was seen which appeared, the danger of LAD increments by expanding the mean estimation of pulse. The discoveries of the present investigation demonstrated that lone 45% of patients had strange TST waves and there was no connection between TST waves' status and LAD. In an exploration on the determination of ventricular cardiomyopathy utilizing ANN calculations, the outcomes indicated that a decrease in the components of heart signals positively affected the cardiovascular sound characterization (Nicholls SJ, 2010).

Another finding of the present examination was identified with the connection between the frequency of LAD and level of triglyceride and creatinine. Truth be told, the danger of LAD could increment because of an expansion in these factors. Besides, a noteworthy relationship was seen between the chest agony and LAD (Perk J). Be that as it may, the consequences of the investigation led by wertli et al. indicated that there was no huge connection between these factors. The chest torment has an abstract nature which can't be utilized for foreseeing LAD and frenzy issue ought to be considered in perceiving sorts of chest torment.

As per the writing survey, various examinations have been led to foresee LAD by utilizing information mining calculations. For example, Kurt et al. utilized calculated relapse, choice trees, order and neural systems lastly, the multi-layer perceptron ANN with the PPV of 78.8% was presented as the best model. In another examination, Sajja utilized a straightforward Bayesian calculation, choice tree and multi-layer perceptron ANN on a dataset. The outcomes demonstrated that the accuracy of the multi-layer perceptron ANN calculation was 91.75, showing the best execution (Janosi A, 1988). In the present examination, LAD was chosen as the yield variable and 25 factors were utilized as info factors. The outcomes indicated that the ANN model could be suitable for fitting these information with the absolute PPV of 0.798. Then again, the SVM calculation fitted the information with littler MAPE and blunder. The bigger estimation of Hosmer-Lemeshow integrity of-fit test additionally demonstrated the predominant exhibition of the SVM model on the information and gave better forecast to LAD analysis. Moreover, the SVM calculation anticipated LAD patients with higher PPV and affectability than the ANN model.

Likewise, the consequences of the past examinations indicated that the utilization of the SVM calculation predicts the sickness and recognizes patients from non-patients with higher precision (Amin SU, 2013), (Desai Shrinivas D., 2018), (Kausar N, 2016). Different investigations have additionally affirmed the predominant presentation and accuracy of SVM. In any case, there are not many examinations which don't affirm the proficiency of this calculation and propose different techniques, for example, paired molecule swarm enhancement (BPSO) and hereditary calculation as the best model of decision for LAD assurance (Kausar N, 2016). In spite of the fact that input factors were chosen dependent on the writing audit and related rules, there may be other hazard factors which can be concentrated later on to increase a greater image of the sickness chance variables. Besides, right now, consequences of two calculations were analysed. The information can be utilized to test different calculations, for example, hereditary calculation to perceive the best execution model.

3. CONCLUSION

The procedure of illness forecast in clinical sciences is as a significant procedure for dynamic and doctors need to realize the hazard factors for various maladies. This procedure can be encouraged by utilizing coherent and deliberate strategies, for example, AI techniques and information mining calculations. Presently, because of the impressive increment in Lymphoma illnesses and the overwhelming budgetary weight forced by them on the general public, social insurance networks are looking for approaches to anticipate, analyze, and treat these ailments adequately. The aftereffects of the present investigation indicated that the utilization of information mining calculations, for example, the SVM model can be valuable in anticipating LAD. Be that as it may, more research is expected to look at the exhibition of various calculations and to locate the best execution model.

REFERENCES

Abawajy JH, K. A. (2013). Multistage approach for clustering and classification of ECG data. *Comput Methods Prog Biomed, 112*(3), 720–730.

Acharya UR, F. H. (2017). Automated detection of lymphoma artery disease using different durations of ECG segments with convolutional neural network. *Knowl Based Syst*, 62–71.

Amato F, L. A.-M. (2013). Artificial neural networks in medical diagnosis. *Journal of Appl Biomed*, 47–58.

Amin SU, A. K. (2013). Genetic neural network based data mining in prediction of heart disease using risk factors. *International conference on IEE of Information & Communication Technologies (ICT).*

Bayturan O, K. S. (2010). Clinical predictors of plaque progression despite very low levels of low-density lipoprotein cholesterol. *Journal of the American Col of Cardio, 55*(24), 2736–2742.

Bellazzi R, F. F. (2011). Predictive data mining in clinical medicine: a focus on selected methods and applications. *Wiley interdisciplinary reviews: WIREs data mining and knowledge discovery, 1*(5), 416–430.

Bergh C, F. K. (2017). Severe infections and subsequent delayed Lymphoma disease. *Eur J Prev Cardiol., 24*(18), 1958–1966.

Charchar FJ, B. L. (2012). Inheritance of lymphoma artery disease in men: an analysis of the role of the Y chromosome. *Lancet, 379*(9819), 915–922.

Das R, T. I. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert Syst Appl., 36*(4), 7675–7680.

Desai Shrinivas D., G. S. (2018). Back-Propagation Neural Network Versus Logistic Regression in Heart Disease Classification. *Advanced Computing and Communication Technologies*, (pp. 133-144).

Dolatabadi AD, K. S. (2017). Automated diagnosis of lymphoma artery disease (LAD) patients using optimized SVM. *Comput Methods Prog Biomed*, 117–126.

Frey P, W. D. (2011). Impact of smoking on Lymphoma events in patients with lymphoma disease receiving contemporary medical therapy (from the treating to new targets (TNT) and the incremental decrease in end points through aggressive lipid lowering (IDEAL) trials). *Am J Cardiol., 7*(2), 145–150.

Guner LA, K. N. (2010). An open-source framework of neural networks for diagnosis of lymphoma artery disease from myocardial perfusion SPECT. *Journal of Nucl Cardiol, 17*(3), 405–413.

(2014). *Healthy environment, healthy heart (Internet). Jakarta: Ministry of Health, Republic of Indonesia.* Retrieved from http://www.depkes.go.id/article/view/201410080002/lingkungan-sehat-jantung-sehat.html.

Janosi A, S. W. (1988). Heart disease data set,. *UCI machine learning repository*. Retrieved from http://archive.ics.uci.edu/ml/datasets/heart+disease

Kamiya K, M. T. (2015). Quadriceps strength as a predictor of mortality in lymphoma artery disease. *128*(11), 1212–1219.

Karaolis MA, M. J. (2010). Assessment of the risk factors of lymphoma heart events based on data mining with decision trees. *IEEE Trans Inf Technol Biomed, 14*(3), 559–566.

Kausar N, A. A. (2016). Ensemble clustering algorithm with supervised classification of clinical data for early diagnosis of lymphoma artery disease. *Journal Med Imaging Health Inform, 6*(1), 78–87.

Kelly BB, F. V. (2010). Promoting Lymphoma health in the developing world: a critical challenge to achieve

global health. *National ALADemies Press*.

Kim J, L. J. (2015). Data-mining-based lymphoma heart disease risk prediction model using fuzzy logic and decision tree. *Healthc Inform Res, 21*(3), 167–174.

Longo D, F. A. (2011). *Harrison's principles of internal medicine.*

Mann DL, Z. D. (2014). Braunwald's heart disease e-book: a textbook of Lymphoma medicine. *Elsevier health sciences*, 861–871.

Montalescot G, S. U. (2013). ESC guidelines on the management of stable lymphoma artery disease: the task force on the management of stable lymphoma artery disease of the European society of cardiology. Euro Heart J. *34*(38), 2949–3003.

Mozaffarian D, B. E. (2016). Executive summary: heart disease and stroke statistics-2016 update: a report from the American Heart Association. *133*(4), 447–454.

Murthy VL, N. M. (2012). Association between lymphoma vascular dysfunction and cardiac mortality in patients with and without diabetes mellitus. *Circulation, 126*(15), 1858–1868.

Nicholls SJ, H. A. (2010). Intravascular ultrasound-derived measures of lymphoma atherosclerotic plaque burden and clinical outcome. *Journal of Am Coll Cardiol., 55*(21), 2399–2407.

Orphanou K, S. A. (2016). DBN-extended: a dynamic Bayesian network model extended with temporal abstractions for lymphoma heart disease prognosis. *IEEE J Biomed Health Inform, 20*(3), 944–952.

Perk J, D. B. (n.d.). European Association for Lymphoma Prevention & rehabilitation (EACPR), European guidelines on Lymphoma disease prevention in clinical practice.

Rezaei-hachesu P, A. M. (2013). Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthc Inform Res., 19*(2), 121–129.

Sanchis-Gomar F, P.-Q. C. (2016). Epidemiology of lymphoma heart disease and acute lymphoma syndrome. *Ann Transl Med., 4*(13).

Sivagowry S, D. M. (2013). An empirical study on applying data mining techniques for the analysis and prediction of heart disease. *International conference of Information communication and embedded systems (ICICES)*, (pp. 265–270).

Steenman M, L. G. (2017). Cardiac aging and heart disease in humans. *Biophys Rev, 9*(2), 131–137.

Sudhakar K, M. D. (2014). Study of heart disease prediction using data mining. *Int Adv Res Comput Sci Soft Eng, 4*(1), 1157-1160.

Sufi F, K. I. (2011). Diagnosis of Lymphoma abnormalities from compressed ECG: a data mining-based approach. *IEEE Trans Inf Technol Biomed, 15*(1), 33–39.

The top ten causes of death. Geneva: World Health Organization. (2017).

U.S. National Library of medicine (NLM). Lymphoma artery disease (LAD) medlineplus. (2017).

Verma L, S. S. (2016). A hybrid data mining model to predict lymphoma artery disease cases using non-invasive clinical data. *Journal of Med Syst., 40*(7), 178.

world Health Organization. Physical activity and older adults. (2018). Retrieved from https://www.who.int/dietphysicalactivity/factsheet_olderadults/en/

Yeh DY, C. C. (2011). A predictive model for cerebrovascular disease using data mining. *Expert Syst Appl., 38*(7), 8970–8977.

Zhou X, C. S. (2010). Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support. *Artifi Intell Med, 48*(2-3), 139–15.