

How to Cite:

Sruthi, M. S., Nivetha, A., Pavithra, M., & Moulyasree, M. (2022). Cardio vascular disease prediction using machine learning web application. *International Journal of Health Sciences*, 6(S8), 1–13. <https://doi.org/10.53730/ijhs.v6nS8.10405>

Cardio vascular disease prediction using machinelearning web application

Ms. Sruthi M S

Department of Computer Science Engineering Sri Krishna College of Technology
Coimbatore, TamilNadu, India

Corresponding author email: sruthi.ms@skct.edu.in

Nivetha A

Department of Computer Science Engineering Sri Krishna College of Technology
Coimbatore, TamilNadu, India

Email: nivetha.annadurai2018@gmail.com

Pavithra M

Department of Computer Science Engineering Sri Krishna College of Technology
Coimbatore, TamilNadu, India

Email: pavithramv01@gmail.com

Moulyasree M

Department of Computer Science Engineering Sri Krishna College of Technology
Coimbatore, TamilNadu, India

Email: moulya2310@gmail.com

Abstract---The Principal Target of this application to be create and make a strategy model that can forecast a patient's status of health based on a large amount of data. Coronary illness is the most common cause of mortality among humans. ML provides many classification techniques to divine the chance of a patient having HD based on the specified task. This project tries to recognize at-risk traits and classify data based on those aspects, allowing for exact prediction. To arrive at efficient findings, the dataset is first pre-handled and investigated against different calculations, with their accuracies evaluated. The efficient method is then employed in the back-end and connected with the front-end as a web application to predict whether or not a person has disease based on the input values. As a result, this web application would be the efficient need of the hour for obtaining precise health outcomes.

Keywords---machine learning, heart disease prediction, feature selection, prediction model, classification algorithms, cardiovascular disease (CVD).

Introduction

Coronary illness, otherwise called cardiovascular sickness, is one of the most widely recognized medical problems that individuals are confronting today. Whenever the heart can't push the required volume of blood to different pieces of the body, this condition creates. Coronary illness, frequently known as cardiovascular sickness, has stayed the main source of mortality overall throughout the last ten years. In the field of clinical information investigation, foreseeing cardiovascular sickness is a significant trouble.. Heart disease diagnosis must be accurate and timely in order to prevent and cure heart failure.

This system integrates many features required for prediction, performs a deep examination of a range of parameters, assures the finesse of the output with several method comparisons, and establishes correct outcomes. Heart sickness has certainly stood out enough to be noticed in clinical examination among the various perilous issues. The finding of heart sickness is a troublesome undertaking that can give a robotized visualization of the patient's heart wellbeing, taking into consideration more fruitful treatment. Smoking, elevated cholesterol, a family background of coronary illness, corpulence, hypertension, and an absence of actual work are largely factors that increment the gamble of coronary illness.

The assortment and investigation of information in the medical care industry to create bits of knowledge and improve navigation is known as medical services examination. Medical care examination can be used on both full scale and miniature levels to effectively work on activities, work on tolerant consideration, and cut complete expenses in basic regions like clinical consumptions, clinical information, patient way of behaving, and meds.

If heart disease and diabetes are detected early enough, lifestyle adjustments may be sufficient to prevent the disease from progressing further and necessitating treatment. Diagnosis at a later stage may result in higher costs and more complicated conditions that are harder to manage. Since the idea of cardiovascular illness is mind boggling, it should be treated with alert. Powerlessness to cause so may harm the heart or result in surprising passing. Clinical science and data mining are utilized to reveal different kinds of metabolic issues. In the expectation of coronary illness and information request, information mining with grouping assumes a significant part. As a result, a predictive online application might be utilized to identify the results and administer the required meds at an earlier stage.

Related works

Gharehchopogh, F. S., & Khalifelu, Z. A. has proposed., Patient hep is keys to effectively overseeing heart illnesses and to assisting the patients with keeping up with personal satisfaction. As of late, with information mining strategies has been separated significant secret data through clinical information, so these techniques confect helpful data for clinical examination and wellbeing focuses. At long last, the contextual investigation is demonstrated by Neural Network (NN). Preparing information gathers of 40 patients clinical records in wellbeing place

search about heart issues in people. The outcomes show that our NN model produces correct predictions for 85% of experiments [1].

Nag, P., Mondal, S., Ahmed, F., More, A., & Raihan, M. has proposed., Intense Myocardial Infarction is a Cardio Disease prompts Ischemic Diseases(IHD) is considered one of the significant executioners around the world. An able system In this paper, it is suggested that the potential can be expected. outcomes of coronary disappointment when an individual carry body desolation indistinguishable after effects. Succeed a model by combining apathetic statistics of victim treated at various clinical sites for Acute Myocardial Infarction (AMI). Chest pain, palpitation, shortness of breath, illness, sweat, and spew are the identifiable indication of coronary failure [2].

Raihan, M., Islam, M. M., Ghosh, P., Shaj, S. A., Chowdhury, M. R., Mondal, S., & More, A. has proposed., Intense Coronary Artery Disease (ACS) is obligated in order to abrupt passing. The cause of tachycardia is ongoing medication use, polygenic hyperpiesia strife. From clinical benefits element, A dataset of Coronary Artery Disease victims been accumulated. Data has been pre-processing the potential outcomes of the necessity of tachycardia by having AI approaches are explored. SVM, AdaBoost, Capture, K-NN, Unplanned holt philosophies pre-owned to the extent that expect the unfairness of Coronary Artery Disease. The top - quality accuracy with AdaBoost and Capture are 75.49% & 76.28%. The Accuracy and survey for AdaBoost are 0.741; 0.75 and 0.755; 0.763 for Capture procedures exclusively [3].

Sen, A.K., Patel, S.B., and Shukla, D.P. has proposed., Cardiovascular sickness stays the greatest reason for passages overall furthermore, the Heart Disease Prediction beginning phase is significance. As huge measure of information is created in clinical associations (emergency clinics, clinical focuses) yet as this information isn't as expected utilized. There is an abundance of stowed away data available during word processing file. This information imaginably changed over in the direction of supportive features. Thus we are able to utilize various information excavate frameworks. As a result of document, portrayed a multi layered perspective for perceiving disorder believability. [4].

Kunwar, V., Chandel, K., Sabitha, A. S., & Bansal, A. has proposed., Information mining has been a latest thing for achieving indicative results. Gigantic proportion of unmined information is assembled clinical benefits steadiness track down concealed away particulars for reasonable end and route. Data mining is the technique engaged with removing hidden away information from immense dataset, orchestrating significant and extraordinary models in data. The target of my peal is to anticipate Chronic Kidney Disease(CKD) utilizing game-plan strategies like Independence Bayes and Unnatural Sensual Meshwork [5].

Ishtake, M. has proposed., Today clinical benefits have made considerable progress to treat patients with different infections. Among the most deadly one is the coronary sickness issue in this way ought not be noticeable with an independent eye & comes rapidly at the point when its limits are reached. Today diagnosing patients precisely and controlling convincing drugs have become a lot of a test. Low deatched choices could end to forbearing end and can't be

administered the clinical office no longer have its standing. The expense to manage a victim with a cardiac infarction is incredibly reasonable not entirely set in stone. To execute a right and astute medicaments PC found particulars or possibly decision genuinely strong organizations can be made to do the endeavor[6].

Bhatla, N., & Jyoti, K. has proposed., Coronary illness is a term that allocates to an enormous number of ailments connected with heart. This infirmities depict uncommon clinical issue that clearly sway the heart and all of its parts. Coronary ailment is a huge ailment in right now. This paper targets analyzing the different data digging procedures introduced recently for coronary ailment assumption. The insights uncover the Sensual associations in the company of 15 credits has outmaneuvered any leftover excavation methodology. One more end from an origin investigation is that structural outline has moreover appear incredible perfection in addition to assist historical estimation and part subdivison assurance [7].

Existing system

The human heart is the main organ in the body. The primary function of the heart is to keep oxygen-rich blood moving throughout the body. Because the heart is so important for living, any irregularities in the heart might create problems in different districts of the body. Coronary illness is characterized as any break in the typical working of the navel. Coronary illness be associated with the fundamental wellsprings of demise in the current society. Coronary illness can be brought about by an unfortunate way of life, for example, smoking, drinking an excess of liquor, or eating a lot of fat, which can prompt hypertension.

Clinical decisions in India not set in stone based on a specialist's experience instead of information on rich realities put away in a data set. In the present medical care, the key trouble is to offer excellent types of assistance and exact analysis. Despite the fact that coronary illness has been distinguished as the main source of death overall lately, additional one of the sickness might be really controlled and made due.

Drawbacks

- In the ongoing framework, there is a great deal of administrative work (as clinical reports) that takes a ton of time and cash.
- There is a chance that the sickness will be misdiagnosed. This would be extremely inconvenient for both management and patients.
- Once calculated, the results may not be consistent (i.e., opinions may differ from one doctor to the next or between hospitals).
- Clinical Misdiagnoses are a not kidding hazard to our medical services calling if, despite everything that they continue, people will fear going to the center for therapy..We can stop clinical misdiagnosis by illuminating people in general and documenting cases and suits against the clinical specialists to blame.

Proposed system

The plan expects to identify heart diseases from the get-go to limit disastrous impacts. The objective of this strategy is to precisely foresee whether a patient has heart illness. The data is integrated into a K-Nearest Neighbor Classifier model that estimates the probability of coronary illness. The model is prepared on a major dataset to foresee the pertinent result contingent upon the factors. The Python programming language was used to create the complete application. Google Colab was the platform on which the applications were built. The above-mentioned algorithms analysed the entire dataset, and factors such as accuracy, precision score, and roc-auc score were calculated and compared.

The proposed solution is divided into two phases:

- Application Development for the Web
- AI for Disease Prediction

The infection expectation framework is utilized to foresee the illness in light of the given information values, while the web application permits the client to associate with the framework. The outfit classifier procedure utilizing K-NN is utilized in the forecast framework. The backend was made utilizing the Python programming language. The proposed strategy utilizes the KNN calculation, the Flask web structure, and the Application Platform (Heroku).

The web application is created for the module so the client can communicate with the stage. The web application comprises of two significant parts:

- Front-End (significantly created in HTML and CSS)
- Back-End (significantly created in Python)

The front-end offers a U/I with insignificant intricacies, though the back-finish of the site is fostered that empowers the correspondence between the frontend and the back-end.

- The infection conjecture system uses unique returns to from front-beginning to end to make the outcome by taking the secondary effects from the client without creating another requesting.
- Illness Prediction: This gives a calculation to anticipate the infection in view of the given side effects utilizing cutting edge AI order calculations.
- Online intuitive stage: Also introducing an intelligent stage coordinated with the sickness forecast framework alongside other significant elements for a more vigorous medical services the board framework

Block diagram



Figure 4.1. Proposed methodology Flow Diagram

Module description

Data Collection

The datasets for our investigations on cardiovascular illness, then again, were utilized in this review. This dataset had likewise been utilized by various different analysts. The dataset is from the Cleveland data set of the UCI store of coronary illness patients, and it is openly accessible on the Kaggle stage as a component of a continuous cardiovascular review. There are 304 events and 76 properties in the assortment. Just 14 of the 76 characteristics are tried, regardless of their significance in demonstrating the viability of various calculations. The reason for the characterization is to decide whether the patient has an decade risk of coronary artery disease (CHD). This collection has 303 occurrences, 14 ascribes, and one objective quality.

Pre-Processing

The first step is to take the bounding box coordinates from the YOLOv4 detection phase and simply place the sub image region within the box's borders. Data preparation is a method of transforming raw data into a format that may be easily understood. The raw data is processed by data manipulation. Customer relationship management and rule-based applications both employ data preparation (like neural networks). During pre-processing, data passes through a variety of steps. Filling in missing values, deleting corrupted or inconsistent data, smoothing noisy data, and addressing data inconsistencies are all examples of data cleansing. It is critical for ML algorithms to cope with noisy outcomes in a

seamless manner. Data can be "filtered" by dividing it into equal halves or by employing "linear regression" or "cluster analysis" techniques (clustering).

Due to human mistake, there are data abnormalities (the data was put away in an off-base field). To hold the worth back from acquiring an edge, never copy information base qualities (inclination). Information mix is the most common way of joining data from an assortment of sources in various arrangements. The information is cleaned up and standardised. Data normalisation ensures that all data and relationships are saved in a single location. The numbers make sense. As the amount of data grows larger, queries will slow down. The intent of the information excavate stage to imply extricate sourced data frameworks. There are a few techniques for diminishing information. A dataset is frequently too vast or complex to work with. Given that the data has the same qualities, sampling techniques can be used to choose and function with only a subset of it.

Rescaling is used to scale the dataset when it contains attributes with different scales. To convert the value into 0 and 1, the binary transformation was used. Every attribute's value is interpreted as 1 if it is above the threshold and 0 if it is below the threshold. Each attribute in the standardised method having a 0 mean and a 1 standard deviation

Feature Selection

Highlight determination is expected for each AI classifier that has been prepared since the outcomes might be impacted on the off chance that superfluous elements are not taken out from the dataset. The classifier calculation with include choice further develops execution and abbreviates the model's execution time. Since this informational index is very basic, the MLP classifier is used, and the exactness score is close to

We're presently prepared to rank the importance of elements. To reiterate, the contrast between the gauge score s and the typical score produced by permuting the comparing section of the examination set decides pertinence a component. The prototype is uncaring toward stages of the component assuming the thing that matters is small, consequently its significance is insignificant. On the off chance that the thing that matters is huge, then again, the element's significance is high. The quantity of stages per highlight is constrained by the boundary n . With more permutations, better estimates are obtained at the expense of computing time. It's worth noting that you've picked a different scoring metric than accuracy, such as the F1 score.

Classification Algorithms For Prediction

A key component of supervised learning is the classification technique. Classifiers learn from the training dataset and use it to locate the target attribute on the testing dataset.

KNN

The KNN ascription approach look through the total informational collection for the k number of most similar cases, or neighbors, that show similar examples as the line with missing information, utilizing the KNN calculation. For each missing worth, the KNNs were used to compute a normal of missing information factors. One disadvantage of KNN is that every variable requires a satisfactory number of values from current information, which in this study is five. K-Nearest Neighbor (KNN) is a notable AI calculation that has been applied to huge scope information mining projects. The idea is that a huge measure of preparing information is utilized, with every information point being characterized by an assortment of elements.

Each point is thoughtfully addressed in a high -layered space, with every hub comparing to a solitary variable. The square foundation of N, the all out number of focuses in the preparation informational index, is regularly utilized as the number k. (Assuming N is 400, K equivalents 20). KNN is an essential idea with the additional advantage of being nonparametric. That is, the methodology can be used with class factors, though assuming you are utilizing numeric factors, it is ideal to standardize them to dispose of scale errors

Classifier Validation Methods

The Raw Data secluded into two regions: getting ready and testing. The preparation dataset contains half of the information, while the excess information is alluded to as testing information. Every order technique got a presentation assessment grid.

Performance Evaluation Measure

The precision score

The quantity of right expectations partitioned by the complete number of information tests is the proportion. It evaluates a calculation's proclivity for properly arranging information.

Classification accuracy

The right pace of expectation findings is shown by classification accuracy. The The confusion matrix is a tool for determining calculate it. Accuracy of classification has been discovered.

Error Classification

The compute of inaccurate expectation outcomes is shown by the classification error. The confusion matrix is used to calculate it. Equation is used to determine the categorization error.

Precision

Accuracy is a critical rule for assessing model execution. It's the magnitude of related cases in of recovered things. A guage esteem is positive. In condition, the accuracy is determined as follows:

Recall

Review is a significant basis for assessing model execution. It's the extent of related examples in the general number of occasions recovered. In condition, the not entirely settled as follows:

F-measure

F Score is another name for it. The F-measure is used to determine the correctness of a test. Equation is used to calculate it from precision and recall.

Flask

Cup is a Python-based web structure that is smaller than expected. It's called a microframework contemplating the manner in which because it does not necessitate the use of specific instruments or libraries. It doesn't have any informational collection reflection layer, structure approval, or anything that different parts libraries for outcasts ordinary restraint. Anyway, Flask maintains developments that can add application features as though they were completed in itself. Object-social mappers, structure endorsement, move management, and other expansions are all available. Different open affirmation developments, as well as a few others typical framework related gadgets. In sending this cardiovascular breakdown expectation model into creation, a web application system called Flask is utilized. Flagon makes it simple to compose applications, and furthermore gives an assortment of decisions for creating web applications.

Heroku

Heroku is a multi-language application stage that licenses specialists to send, and manage their applications. It is versatile and easy to use, offering engineers the most direct approach to getting their applications to publicize. Heroku is a cloud stage as a help (PaaS) supporting a couple of programming tongues. One of the chief cloud stages, Heroku has been being created since June 2007, when it maintained only the Ruby programming language, but as of now maintains Java, Node.js, Scala, Clojure, Python, PHP, and Go..

Implementation & Result

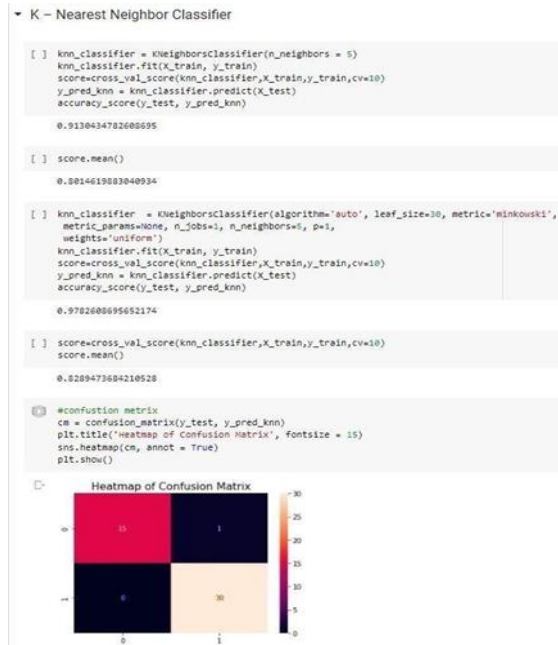


Figure 6.1. Results obtained for KNN Classifier

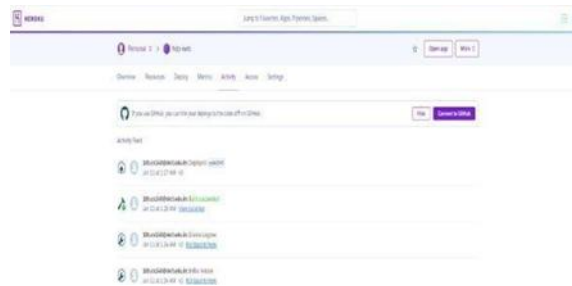
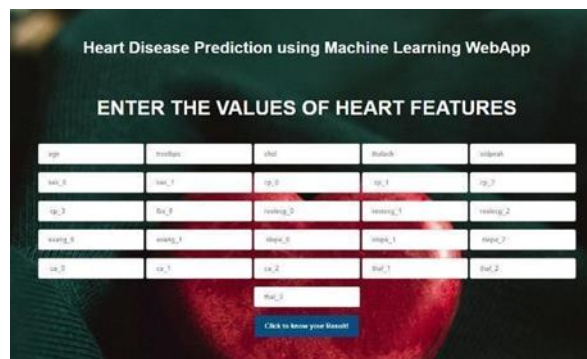


Figure 6.2. Project Upload in Heroku



The screenshot shows the front-end of the 'Heart Disease Prediction using Machine Learning WebApp'. The page has a dark background with a heart image. The title 'Heart Disease Prediction using Machine Learning WebApp' is at the top. Below the title, there's a section titled 'ENTER THE VALUES OF HEART FEATURES'. This section contains a grid of input fields for various heart features: age, sex, chest pain, blood pressure, cholesterol, fasting blood sugar, resting blood sugar, exercise induced angina, oldpeak, slope, max heart rate, and rest heart rate. Each feature has a corresponding input field with a label and a unit. At the bottom of the grid, there's a 'Click to know your Result' button.

Figure 6.3. Web Application Front-end

Heart Disease Prediction using Machine Learning WebApp

ENTER THE VALUES OF HEART FEATURES

45	145	400	187	123
0	1	0	0	1
0	1	0	1	0
1	1	1	0	0
1	1	0	0	1

Click to know your Result!

Figure 6.4. Input Values entered into the Fields

Heart Disease Prediction using Machine Learning WebApp

ENTER THE VALUES OF HEART FEATURES

age	height	wt	temp	cholesterol
sex_0	sex_1	sex_2	sex_3	sex_4
chol_1	chol_2	cholesterol_1	cholesterol_2	cholesterol_3
cholesterol_4	cholesterol_5	cholesterol_6	cholesterol_7	cholesterol_8
chol_9	chol_10	chol_11	chol_12	chol_13

Click to know your Result!

Patient has **** HEART DISEASE ****

Figure 6.5. Results Obtained with Detection of Heart Disease

Heart Disease Prediction using Machine Learning WebApp

ENTER THE VALUES OF HEART FEATURES

age	height	wt	temp	cholesterol
sex_0	sex_1	sex_2	sex_3	sex_4
chol_1	chol_2	cholesterol_1	cholesterol_2	cholesterol_3
cholesterol_4	cholesterol_5	cholesterol_6	cholesterol_7	cholesterol_8
chol_9	chol_10	chol_11	chol_12	chol_13

Click to know your Result!

Patient has **NO HEART DISEASE**

Figure 6.6 Results Obtained with Detection of No Heart Disease

Conclusion

Since heart illnesses are a significant executioner in India and all through the world, a use of a promising innovation like AI to the underlying forecast of heart sicknesses will significantly affect the general public. In this way, the application would assist the end clients with getting a fundamental forecast about the state of their heart. This will let the client know if they are at a gamble and assuming that they need to visit the specialist.. This will help with lessening the end rate due to respiratory disappointments. A contamination conjecture system is a valuable device in moving the clinical consideration plan of the country.

Consequently, the task plans to anticipate the chances of having coronary illness and with exact investigation dodges the probabilities of blunder because of manual mediation. An electronic application is moreover developed to help people team up and use the disorder assumption system. Henceforth by utilizing the above approach effective examination of heart sicknesses of the individual was performed and the outcome was acquired which anticipated the gamble of coronary illness.

References

- [1] Gharehchopogh, F. S., & Khalifelul, Z. A. (2011, July). Neural network application in diagnosis of patient: a case study. In *International Conference on Computer Networks and Information Technology* (pp. 245-249). IEEE.
- [2] Nag, P., Mondal, S., Ahmed, F., More, A., & Raihan, M. (2017, December). A simple acute myocardial infarction (Heart Attack) prediction system using clinical data and data mining techniques. In *2017 20th International Conference of Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE.
- [3] Raihan, M., Islam, M. M., Ghosh, P., Shaj, S. A., Chowdhury, M. R., Mondal, S., & More, A. (2018, December). A comprehensive analysis on risk prediction of acute coronary syndrome using machine learning approaches. In *2018 21st International Conference of Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE.
- [4] Sen, A.K., Patel, S.B., & Shukla, D.P. (2013). A Data Mining Technique for Prediction of Coronary Heart Disease Using Neuro-Fuzzy Integrated Approach Two Level.
- [5] Kunwar, V., Chandel, K., Sabitha, A. S., & Bansal, A. (2016, January). Chronic Kidney Disease analysis using data mining classification techniques. In *2016 6th International Conference-Cloud System and BigData Engineering (Confluence)* (pp. 300-305). IEEE.
- [6] Ishtake, M. " Intelligent Heart Disease Prediction System Using Data Mining Techniques ".
- [7] Bhatla, N., & Jyoti, K. (2012). An Analysis of Heart Disease Prediction using Different Data Mining Techniques. *International journal of engineering research and technology*, 1.
- [8] Xie, J., Wu, R., Wang, H., Chen, H., Xu, X., Kong, Y., & Zhang, W. (2021). Prediction of cardiovascular diseases using weight learning based on density information. *Neurocomputing*, 452, 566-575.
- [9] Sultana, M., Haider, A., & Uddin, M. (2016). Analysis of data mining techniques for heart disease prediction. *2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, 1-5.
- [10] Rinatha, K., & Suryasa, W. (2017). Comparative study for better result on query suggestion of article searching with MySQL pattern matching and Jaccard similarity. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-4). IEEE.
- [11] Rinatha, K., Suryasa, W., & Kartika, L. G. S. (2018). Comparative Analysis of String Similarity on Dynamic Query Suggestions. In *2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS)* (pp. 399-404). IEEE.
- [12] Kadir, Y. R., Syarif, S., Arsyad, M. A., Baso, Y. S., & Usman, A. N. (2021).

Female's reproductive health application design on the school teacher knowledge: an android-based learning media. *International Journal of Health & Medical Sciences*, 4(2), 189-195.
<https://doi.org/10.31295/ijhms.v4n2.1686>