Application Of Support Vector Machine Method To Predict Heart Disease

Golfrid Heraldi Simatupang[1]*, Elvis Sastra Ompusunggu*[2]
[1,2]*Universitas Prima Indonesia, Medan-Indonesia*
email: heraldisims@gmail.com, sastraelvis@gmail.com

ABSTRACT

Heart attack disease is when the arteries are blocked by fatty deposits This results in symptoms like chest discomfort and dyspnea. Furthermore, Damage to the heart muscle can result from obstructed or reduced blood flow to the heart. Heart attack disease remains Indonesia's greatest cause of death as of right now. The current problem is that it is very difficult to predict heart disease and identify heart disease. The right method is needed to predict heart disease. The purpose of this study was to calculate the level of accuracy of the Support Vector Machine method in predicting heart attack disease. The research findings and data analysis conducted utilizing the Support Vector Machine algorithm yielded an accuracy rate of 91.8%. Thus, it can be said that in comparison to the K-Nearest Neighbor approach, the support vector machine algorithm is superior in predicting the development of heart attack disease, which achieved an accuracy of 88%, and Logistic Regression, which achieved 83% accuracy.

Keywords: Heart Attack, Support Vector Machine, Prediction.

INTRODUCTION

The World Health The World Health Organization (WHO) says that fat deposits cause clogged arteries, leading to heart disease. This disease can cause breathing difficulties and chest pain [1]. Furthermore, reduced circlulation of blood that inhibits and damages the heart muscle can also lead to heart disease [2]. Other causes of heart disease include plaque build-up in the arteries, which can block blood flow to the heart, infections, valves that are not functioning properly, having unhealthy lifestyle habits, as well as the use of certain medications. There are four chambers in the heart: Two ventricles and Two atria. The right atrium receives blood from the rest of the body, while the left atrium receives blood from the lungs. And the right ventricle pumps blood to the lungs for oxygen, while the left ventricle pumps oxygen-rich blood throughout the body [3]. Between the left and right chambers is a muscular wall called the septum, which prevents the mixing of low-oxygen blood with high-oxygen blood. The heart's main function is to spread oxygen-rich blood throughout the body. [4]. After the oxygen in all organs is depleted, blood containing little oxygen returns to the heart for the process of replenishing oxygen in the lungs [5]. Currently, the problem faced is the difficulty of diagnosing heart disease accurately and in a timely manner is a complex obstacle. Therefore, appropriate techniques are required for more accurate heart disease prediction [6]. Harvard Health Publishing states that heart disease is frequently perceived as an illness that strikes unexpectedly, despite the fast that years of plaque accumulation that clogs the heart arteries cause this disease to manifet[7]. Blockage of the heart's blood arteries occurs due to the accumulation of fat, cholesterol, afactnd plaque formation in arteries. When the plaque ruptures, it forms a blood clot that disrupts the circulation of blood flow and damages the cardiac muscle[8]. Based on study done in 2019 by Ade Riani. Application of Data Mining for heart disease prediction employing the Naive Bayes Technique. The Naive Bayes algorithm, which this study uses for computations, yields an accuracy value of 86%; however, further approaches will need to be used to forecast heart disease with a higher accuracy value [9]. Utilization of Data Mining in Predicting Heart Disease Risk: A Literature Review of the Logistic Regression and k-Nearest Neighbor Algorithms. by Nicholas (2022). Where the calculation uses based on the comparison results, the K-Nearest Neighbor algorithm shows an accuracy of 83% and the Logistic Regression algorithm shows an accuracy of 88% in predicting heart disease. This shows that alternative techniques with higher accuracy are needed to improve the effectiveness of heart disease prediction[10]. The calculation results demonstrating an accuracy of applying the Support Vector Machine algorithm 91.8%. Based on this research, drawing this conclusion, the Support Vector Machine (SVM) algorithm is the most effective in predicting heart attacks compared to the K-Nearest Neighbor (KNN) algorithm which has 88% accuracy and Logistic Regression with 83% accuracy.

METHODS

This research method is carried out through several stages

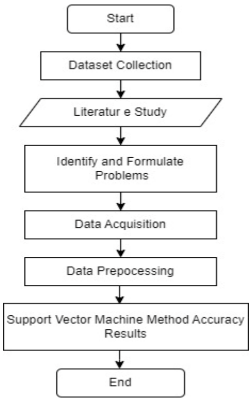

Start
Dataset Collection
Literatur e Study
Identify and Formulate Problems
Data Acquisition
Data Prepocessing
Support Vector Machine Method Accuracy Results
End

Figure 1. Research flow stages

a. Dataset Collection: This research uses *heart attack* diseasedata obtained this data will be processed on the Kaggle platform. used as an important source of information in this study

b. Literature Study: Collection of previous research books and journals related to the research objectives, specifically, heart attack disease [10]
c. Identify and formulate problems: At this stage, it is important to increase efforts to find reliable sources of information regarding heart disease and the reasons behind the issues that those who have it face.
d. Data Acquisition: This study used a heart attack disease dataset taken from the kaggle platform. This dataset is then processed through a series of data mining processes to identify the most common factors that result in heart attack disease [11]
e. Data Prepocessing: The process of transforming raw data into a more structured format is necessary because raw data is often presented in an unorganized form. This aims to make it a source of information that can be processedfurther through an organized data set [12].
f. Support Vector Machine Results of the Method Accuracy: This study will produce a comparison of the accuracy value between Support Vector Machine and previous research. to predicting disease heartattack.

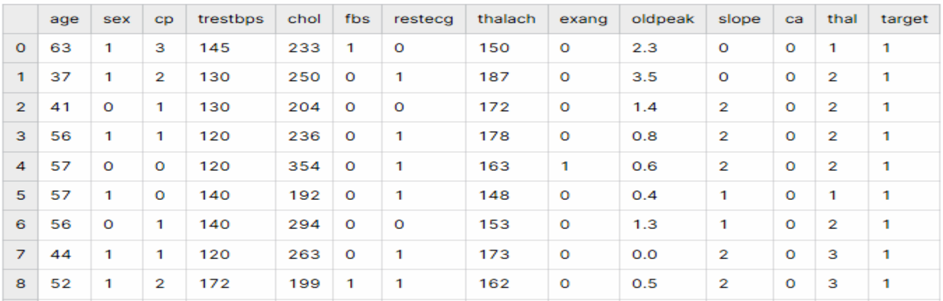
RESULTS AND DISCUSSION

Problem Analysis

Heart disease is a condition where the heart is impaired in performing its function. This disorder can arise due to various factors, such as damage to blood vessels. Heart valves, or the heart muscle itself. Other factors such as infection and birth defects can also cause heart disease[13]. The disease is usually caused by blockages, obstructions, or damage to the blood vessels and heart muscle. This condition inhibits blood flow to the heart, reducing the circulation of oxygen and nutrients tosurrounding tissues and muscles [14]. The highmortality rate from heart disease is triggeredby two main factors: Not many people realize how important it is to undergo regular heart health check-ups and avoid unhealthy lifestyles.

Data Analysis

The dataset used in this study comes from Kaggle. Data regarding heart attacks is used as a sample for data processing.



	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
7	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
8	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1

Figure 2. Heart Disease Dataset

The overall data to be processed consists of 303 samples that have 14 characteristics to be processed. In addition, there are 2 labels, where 0 indicates a lower probability of having heart disease and 1 indicates a higher probabilty of having heart disease. The dataset that the attributes described here.

Table 1. Description of Heart Attack Research Attributes

No.	Name Field	Description
1	Age	Patient Age.
2	Sex	Patient Gender (1= Male, 0 = Female).
3	Cp	Typical Angina 0=Common symptoms of chest pain with possible coronary artery blockage. 1=Atypical Angina: Non- specific symptoms, lower likelihood of arterial blockage than typical angina. 2=Non-anginal pain: A stabbing or aching sensation in the chest, prolonged or prolonged, unrelated to arterial blockage. 3=Asymptomatic: No symptoms of chest disease.
4	Trestbps	The patient's blood pressure at rest in mmHg.
5	Chol	Serum cholesterol level in mgdl.
6	Fbs	Fasting blood sugar level in mg/dl (0=less than 120mg/dl, 1=more than 120mg/dl.
7	Restecg	Electrocardiographic results at rest. (0 = Normal, 1= ST wave increased/decreased by more than 0.5 mV, 2 = Left ventricular hypertrophy).
8	Thalach	Maximum heart rate reached
9	Exang	Chest pain arising from physical activity (0 = no pain, 1 = pain).

10	Oldpeak	The magnitude of the ST segment at rest is relative to the resting state.
11	Slope	The slope of the ST segment at peak or maximum physical activity conditions. (0 = downsloping, 1 = flat, 2 = upsloping).
12	Ca	Number of blocked main vessels (0-3).
13	Thal	Cardiac status was categorized into 4 including, 0 = unknown, 1 = permanent disability, 2 = normal, 3 = reversible disability.
14	Target	Indication of heart attack. (0 indicates a lower risk of having a heart attack, while 1 indicates a higher risk)

***Support Vector Machine* Flowchart**

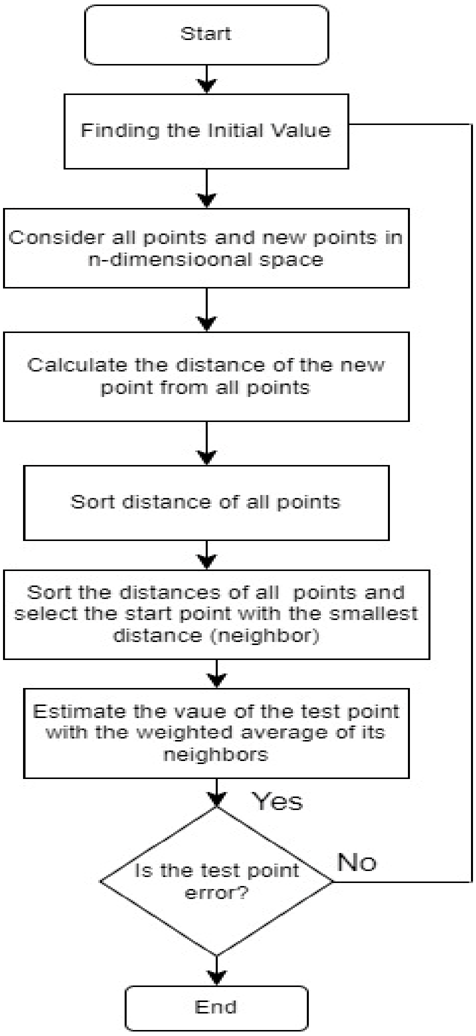

Start
Finding the Initial Value
Consider all points and new points in n-dimensioonal space
Calculate the distance of the new point from all points
Sort distance of all points
Sort the distances of all points and select the start point with the smallest distance (neighbor)
Estimate the vaue of the test point with the weighted average of its neighbors
Yes
Is the test point error?
No
End

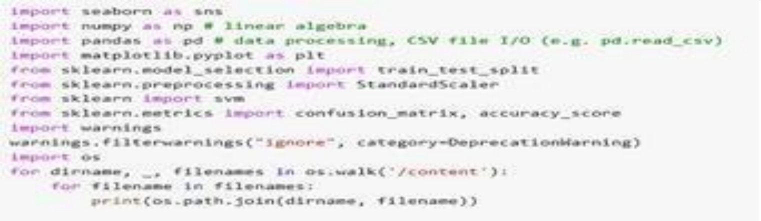
Figure 3. Support Vector Machine Flowchart

Data Processing

Heart disease is like the number one killer that lurks at every age, ready to take lives at any time data processing will be carried out using the *Support Vector Machine* this research method is used to produce accurate value. The following is the data processing process carried out

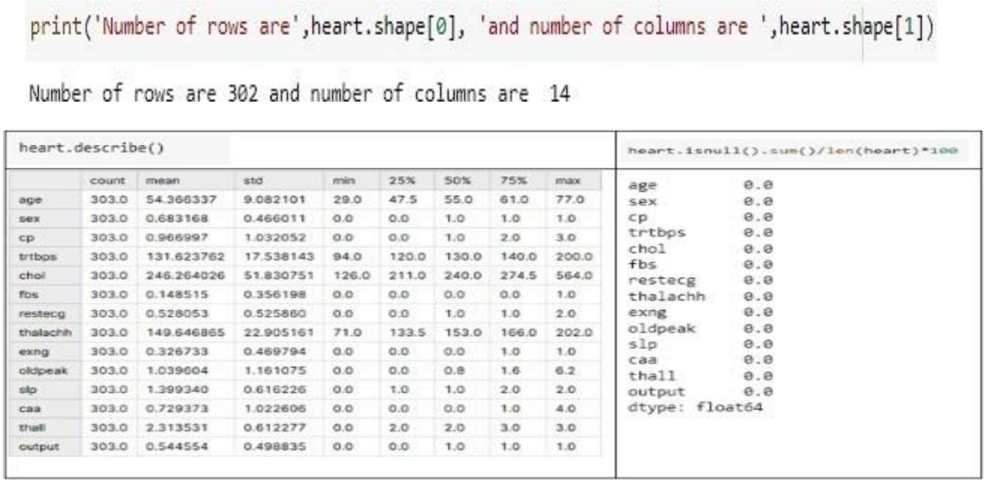
Import Library

Here are the initial steps in the import library data processing process:

Import Library *Code Source:*	

```
import seaborn as sns
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn import svm
from sklearn.metrics import confusion_matrix, accuracy_score
import warnings
warnings.filterwarnings("ignore", category=DeprecationWarning)
import os
for dirname, _, filenames in os.walk('/content'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

After the library import process is carried complete, the next step to check the number of datasets:



```
print('Number of rows are',heart.shape[0], 'and number of columns are ',heart.shape[1])
```

```
Number of rows are 302 and number of columns are  14
```

```
heart.describe()
```

	count	mean	std	min	25%	50%	75%	max
age	303.0	54.366337	9.082101	29.0	47.5	55.0	61.0	77.0
sex	303.0	0.683168	0.466011	0.0	0.0	1.0	1.0	1.0
cp	303.0	0.966997	1.032052	0.0	0.0	1.0	2.0	3.0
trtbps	303.0	131.623762	17.538143	94.0	120.0	130.0	140.0	200.0
chol	303.0	246.264026	51.830751	126.0	211.0	240.0	274.5	564.0
fbs	303.0	0.148515	0.356198	0.0	0.0	0.0	0.0	1.0
restecg	303.0	0.528053	0.525860	0.0	0.0	1.0	1.0	2.0
thalachh	303.0	149.646865	22.905161	71.0	133.5	153.0	166.0	202.0
exng	303.0	0.326733	0.469794	0.0	0.0	0.0	1.0	1.0
oldpeak	303.0	1.039604	1.161075	0.0	0.0	0.8	1.6	6.2
slp	303.0	1.399340	0.616226	0.0	1.0	1.0	2.0	2.0
caa	303.0	0.729373	1.022606	0.0	0.0	0.0	1.0	4.0
thall	303.0	2.313531	0.612277	0.0	2.0	2.0	3.0	3.0
output	303.0	0.544554	0.498835	0.0	0.0	1.0	1.0	1.0

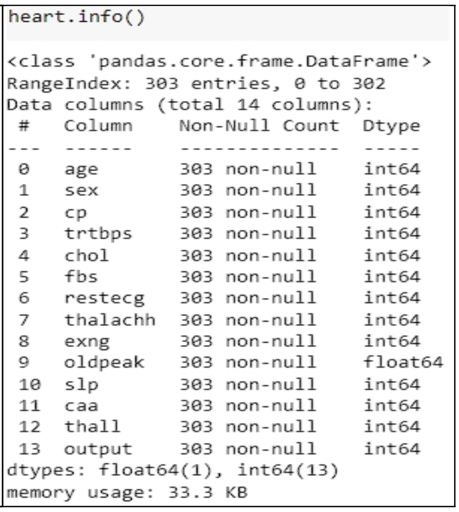
```
heart.isnull().sum()/len(heart)*100
```

```
age         0.0
sex         0.0
cp          0.0
trtbps      0.0
chol        0.0
fbs         0.0
restecg     0.0
thalachh    0.0
exng        0.0
oldpeak     0.0
slp         0.0
caa         0.0
thall       0.0
output      0.0
dtype: float64
```

Figure 4. Dataset Header Initialization

Verifying The Data Type of Attributes

The next step is to verify the data type to be used as an attribute, the data type checking process is a quick way to find a value that matches attributes such as *strings* or*string arrays*.

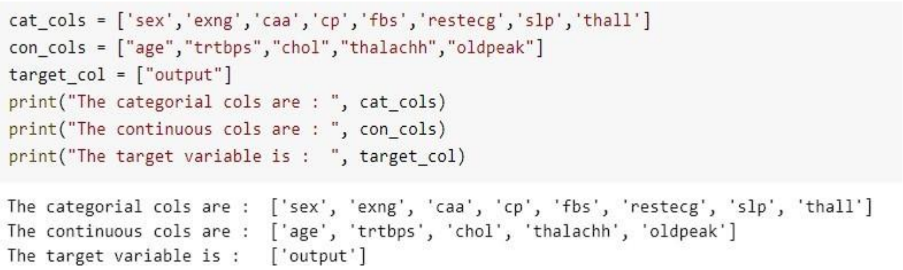


```
heart.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 303 entries, 0 to 302
Data columns (total 14 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   age       303 non-null    int64
 1   sex       303 non-null    int64
 2   cp        303 non-null    int64
 3   trtbps    303 non-null    int64
 4   chol      303 non-null    int64
 5   fbs       303 non-null    int64
 6   restecg   303 non-null    int64
 7   thalachh  303 non-null    int64
 8   exng      303 non-null    int64
 9   oldpeak   303 non-null    float64
 10  slp       303 non-null    int64
 11  caa       303 non-null    int64
 12  thall     303 non-null    int64
 13  output    303 non-null    int64
dtypes: float64(1), int64(13)
memory usage: 33.3 KB
```

Separating the columns in categorical and continuous

Step next step, separated the data columns into two types: categorized data and continuous data, making data processing easier

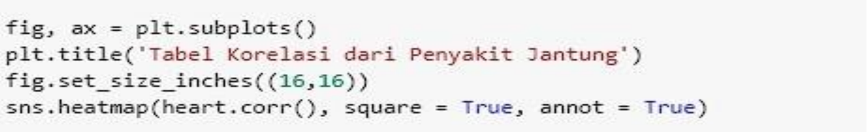


```
cat_cols = ['sex','exng','caa','cp','fbs','restecg','slp','thall']
con_cols = ["age","trtbps","chol","thalachh","oldpeak"]
target_col = ["output"]
print("The categorial cols are : ", cat_cols)
print("The continuous cols are : ", con_cols)
print("The target variable is :  ", target_col)
```

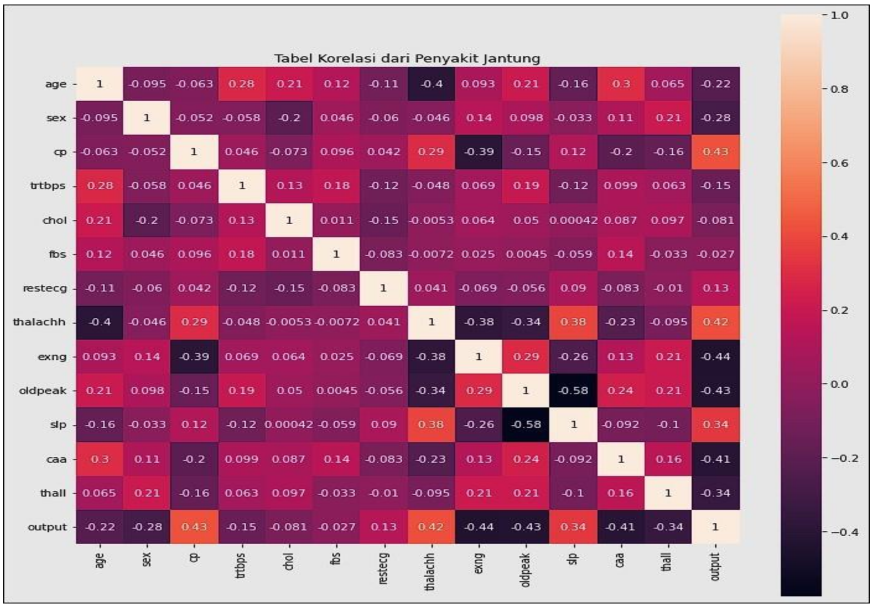
```
The categorial cols are :  ['sex', 'exng', 'caa', 'cp', 'fbs', 'restecg', 'slp', 'thall']
The continuous cols are :  ['age', 'trtbps', 'chol', 'thalachh', 'oldpeak']
The target variable is :   ['output']
```

Data Visualization

The purpose of data visualization is to transform data into something easier to understand and analyze. The following is the *source code for* data visualization:



```
fig, ax = plt.subplots()
plt.title('Tabel Korelasi dari Penyakit Jantung')
fig.set_size_inches((16,16))
sns.heatmap(heart.corr(), square = True, annot = True)
```



	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
age	1	-0.095	-0.063	0.28	0.21	0.12	-0.11	-0.4	0.093	0.21	-0.16	0.3	0.065	-0.22
sex	-0.095	1	-0.052	-0.058	-0.2	0.046	-0.06	-0.046	0.14	0.098	-0.033	0.11	0.21	-0.28
cp	-0.063	-0.052	1	0.046	-0.073	0.096	0.042	0.29	-0.39	-0.15	0.12	-0.2	-0.16	0.43
trtbps	0.28	-0.058	0.046	1	0.13	0.18	-0.12	-0.048	0.069	0.19	-0.12	0.099	0.063	-0.15
chol	0.21	-0.2	-0.073	0.13	1	0.011	-0.15	-0.0053	0.064	0.05	0.00042	0.087	0.097	-0.081
fbs	0.12	0.046	0.096	0.18	0.011	1	-0.083	-0.0072	0.025	0.0045	-0.059	0.14	-0.033	-0.027
restecg	-0.11	-0.06	0.042	-0.12	-0.15	-0.083	1	0.041	-0.069	-0.056	0.09	-0.083	-0.01	0.13
thalachh	-0.4	-0.046	0.29	-0.048	-0.0053	-0.0072	0.041	1	-0.38	-0.34	0.38	-0.23	-0.095	0.42
exng	0.093	0.14	-0.39	0.069	0.064	0.025	-0.069	-0.38	1	0.29	-0.26	0.13	0.21	-0.44
oldpeak	0.21	0.098	-0.15	0.19	0.05	0.0045	-0.056	-0.34	0.29	1	-0.58	0.24	0.21	-0.43
slp	-0.16	-0.033	0.12	-0.12	0.00042	-0.059	0.09	0.38	-0.26	-0.58	1	-0.092	-0.1	0.34
caa	0.3	0.11	-0.2	0.099	0.087	0.14	-0.083	-0.23	0.13	0.24	-0.092	1	0.16	-0.41
thall	0.065	0.21	-0.16	0.063	0.097	-0.033	-0.01	-0.095	0.21	0.21	-0.1	0.16	1	-0.34
output	-0.22	-0.28	0.43	-0.15	-0.081	-0.027	0.13	0.42	-0.44	-0.43	0.34	-0.41	-0.34	1

Figure 5. Correlation Table of Heart Disease

From the Correlation Table of Regarding heart disease, it can be observed that when the color is darker, the risk of developing heart disease is less, otherwise if the color is lighter / brighter, the risk is greater (prone to heart disease).

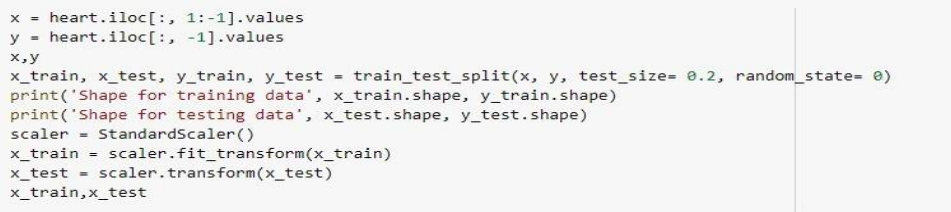
When viewed from the table, thelargest influence value is obtained 0.43 for the influence of the relationship between output and cp, where the output (people with heart attack disease) is indicated. or not) and cp (the type of chest pain that oftenoccurs), so it can be said that people who areindicated by heart attack disease and experience chest pain have an influence on the occurrence of heart disease.

Followed by a value of 0.42 which is influenced by thalach and output, where thalach (maximum heart rate in each person), so that if the heart does not reach a sufficient value and indicates heart attack disease, it can be a factor causing heart disease.

Followed by a value of 0.38 which is influenced by thalach and slope, where slope (The level of steepness of the ST segment during exercise or at maximum stressconditions), so that if people who exercise, the heart rate can trigger the occurrence of heart disease experienced if the slope of the ST segment is large enough.

Data Prepocessing

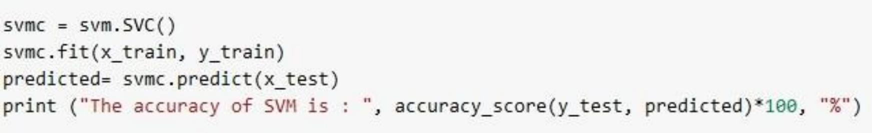
By converting raw data into a simpler format, the information contained in it becomes easier to understand and process, making it easier to analyze and process data.



```
x = heart.iloc[:, 1:-1].values
y = heart.iloc[:, -1].values
x,y
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size= 0.2, random_state= 0)
print('Shape for training data', x_train.shape, y_train.shape)
print('Shape for testing data', x_test.shape, y_test.shape)
scaler = StandardScaler()
x_train = scaler.fit_transform(x_train)
x_test = scaler.transform(x_test)
x_train,x_test
```

Support Vector Machine

After with the completion of data processing, the next steps are to find the accuracy value using the algorithm *Support Vector Machine in* order to know the value generated using the algorithm. Thefollowing is the *source code for* prediction using *Support Vector Machine*:

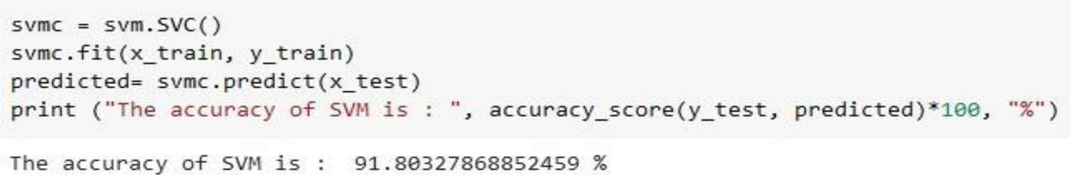


```
svmc = svm.SVC()
svmc.fit(x_train, y_train)
predicted= svmc.predict(x_test)
print ("The accuracy of SVM is : ", accuracy_score(y_test, predicted)*100, "%")
```

RESULTS

This research aims to predict the symptoms of heart attack using the Support Vector Machine (SVM) algorithm. SVM was chosen because of its ability to classify data with a high level of accuracy.

Source Code Support Vector Machine:



```
svmc = svm.SVC()
svmc.fit(x_train, y_train)
predicted= svmc.predict(x_test)
print ("The accuracy of SVM is : ", accuracy_score(y_test, predicted)*100, "%")
```

```
The accuracy of SVM is :  91.80327868852459 %
```

The findings of this study show that SVM is the best algorithm for predicting heart attack disease. The ability of *Support Vector Machine (SVM)* to achieve 91.8% accuracy makes it a very useful tool in the field of medicine.

CONCLUSION

Heart attack has been proven to be one of the most common diseases dangerous in the world. Accurate and precise prediction of the risk of heart attack disease in needed in the research conducted. Based on the results of reseacrh and data analysis, the Support Vector Machine algorithm obtained a 91.8% accuracy rate. Therefore, the conclusion is that the Support Vector Machine algorithm is more effective in predicting heart attack disease when compared to the K-Nearest Neighbor and Logistic Regression algorithms From previous research.

REFERENCES

Supriyatna, H. A., Away, Y., & Zulhelmi, Z. "Desain sistem Internet of Things (IoT) untuk pemantauan dan prediksi gejala serangan jantung." Jurnal Komputer, Informasi Teknologi, dan Elektro, vol. 4, no. 1, 2019.

Dhany, H. W. "Performa Algoritma K-Nearest Neighbour dalam Memprediksi Penyakit Jantung." In Seminar Nasional Informatika (SENATIKA), pp. 176-179, 2021, June.

Al Azhima, S. A. T., Darmawan, D., Hakim, N. F. A., Kustiawan, I., Al Qibtiya, M., & Syafei, N. S. "Hybrid Machine Learning Model untuk memprediksi Penyakit Jantung dengan Metode Logistic Regression dan Random Forest." Jurnal Teknologi Terpadu, vol. 8, no. 1, pp. 40-46, 2022.

Annisa, R. "Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung." JTIK (Jurnal Teknik Informatika Kaputama), vol. 3, no. 1, pp. 22-28, 2019.

Putra, P. D., & Rini, D. P. "Prediksi Penyakit Jantung dengan Algoritma Klasifikasi." In Annual Research /Seminar (ARS), vol. 5, no. 1, pp. 95-99, 2020, February.

Pangaribuan, J. J., Tanjaya, H., & Kenichi, K. "Mendeteksi Penyakit Jantung Menggunakan Machine Learning Dengan Algoritma Logistic Regression." Journal Information System Development (ISD), vol. 6, no. 2, pp. 1-10, 2021.

Nawawi, H. M., Purnama, J. J., & Hikmah, A. B. "Komparasi Algoritma Neural Network dan Naive Bayes Untuk Memprediksi Penyakit Jantung." Jurnal Pilar Nusa Mandiri, vol. 15, no. 2, pp. 189-194, 2019.

Gunawan, M. I., Sugiarto, D., & Mardianto, I. "Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Seacrh pada Algoritma Logistic

Regression." JEPIN (Jurnal Edukasi Dan Penelitian Informatika), vol. 6, no. 3, pp. 280- 284, 2020.

Achmad, A. I. "Metode Regresi Probit Biner untuk Pemodelan Faktor- Faktor yang Mempengaruhi Diagnosis Penyakit Jantung." Jurnal Riset Statistika, pp. 27-33, 2022.

Rahayu, S., Subekhi, A., Astuti, D., Widaningsih, I., Sartika, I., Nurhayani, N., ... & Rafidah, R."UPAYA MEWASPADAI SERANGAN JANTUNG MELALUIPENDIDIKAN KESEHATAN."JMM (Jurnal Masyarakat Mandiri), vol. 4, no. 2, pp. 163-171, 2020.

Majid, A. M., & Miharja, M. N. D. "PENERAPAN METODE DISCRETIZATION DAN ADABOOST UNTUK MENINGKATKAN AKURASI ALGORITMA KLASIFIK ASI DALAM MEMPREDIKSI PENYAKIT JANTUNG." Indonesian Journal of Business Inrelligence (IJUBI), vol. 5, no. 2, pp. 70-75, 2022

Riani, A., Susianto, Y., & Rahman, N. "Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Mengunakan Metode Naive Bayes." Journal of Innovation Information Technology and Application (JINITA), vol. 1, no. 1, pp. 25-34, 2019.

Andiani, L., Sukemi, S., & Rini, D. P. "Analisis Penyakit Jantung Menggunakan Metode KNN Dan Random Forest." In Annual Research Seminar (ARS), vol. 5, no. 1, pp. 165- 169, 2020, February.

Karyatin, K. "Faktor-Faktor Yang Berhubungan Dengan Kejadian Penyakit Jantung Koroner." Jurnal Ilmiah Kesehatan, vol. 11, no. 1, pp. 37-43.

Bianto, M. A., Kusrini, K., & Sudarmawan, S. "Perancangan Sistem Klasifikasi Penyakit Jantung Mengunakan Naïve Bayes." Creative Information Technology Journal, vol. 6, no. 1, pp. 75-83, 2020.

Qomariyah, N., Hamzah, N., & Mustika, W. P. "PENERAPAN METODE ARTIFICIAL NEURAL NETWORK UNTUK MENDETEKSI SERANGAN JANTUNG DI RS AWAL BROS BEKASI." INTI Nusa Mandiri, vol. 13, no. 1, pp. 27-32, 2019.

Apriyatmoko, R., & Aini, F. "Remaja Mengenali Serangan Jantung Koroner." INDONESIAN JOURNAL OF COMMUNITY EMPOWERMENT (IJCE), vol. 2, no. 2, 2020.

Derisma, D. "Perbandingan Kinerja Algoritma untuk Prediksi Penyakit Jantung dengan Teknik Data Mining." Journal of Applied Informatics and Computing, vol. 4, no. 1, pp. 84-88, 2020.

Utomo, D.P., & Mesran, M. "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung." Jurnal Media Informatika Budidarma, vol. 4, no. 2, pp. 437-444, 2020.

Aripin, H. A. "OUTCOME PREDICTION UNTUK PENYAKIT JANTUNG DENGAN ALGORITMA ARTIFICIAL NEURAL NETWORK." INFOKOM (Informatika &Komputer), vol. 9, no. 1, pp. 30-45, 2021.

Sitanggang ,D., Nicholas, Wilson ,V., Sinaga ,A.R.A, dan Simanjuntak ,A.D. “IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI PENYAKIT JANTUNG MENGGUNAKAN METODE K-NEAREST NEIGHBOR DAN LOGISTIC REGRESSION” Jurnal Tekinkom (Teknik Informasi dan Komputer), vol 5, no. 2, pp. 493-501 ,2022