

# *heart* DISEASE

FRANSISKUS ABEL (1102200398)

DON VITO V.M.I.H.R.I (1102201684)

KURNIA MUSTIKA WATI (1102204347)

# *Table of* **CONTENT**

---

LATAR BELAKANG

TUJUAN

METODE

HASIL

PENUTUPAN

# LATAR BELAKANG

Penyakit jantung masih menjadi penyebab utama kematian di Indonesia. Data Riset Kesehatan Dasar (Riskesdas) 2013 dan 2018 menunjukkan tren peningkatan penyakit jantung yakni 0,5% pada 2013 menjadi 1,5% pada 2018. Bahkan penyakit jantung ini menjadi beban biaya terbesar. Berdasarkan data BPJS Kesehatan pada 2021 pembiayaan kesehatan terbesar ada pada penyakit jantung sebesar Rp.7,7 triliun. Oleh sebab itu kami menggunakan artificial intelligence untuk membantu mendeteksi resiko terkena penyakit jantung agar dapat ditindaklanjuti lebih dini.

# TUJUAN

kami menggunakan artificial intelligence untuk membantu mendeteksi resiko penyakit jantung agar dapat ditindaklanjuti lebih dini.



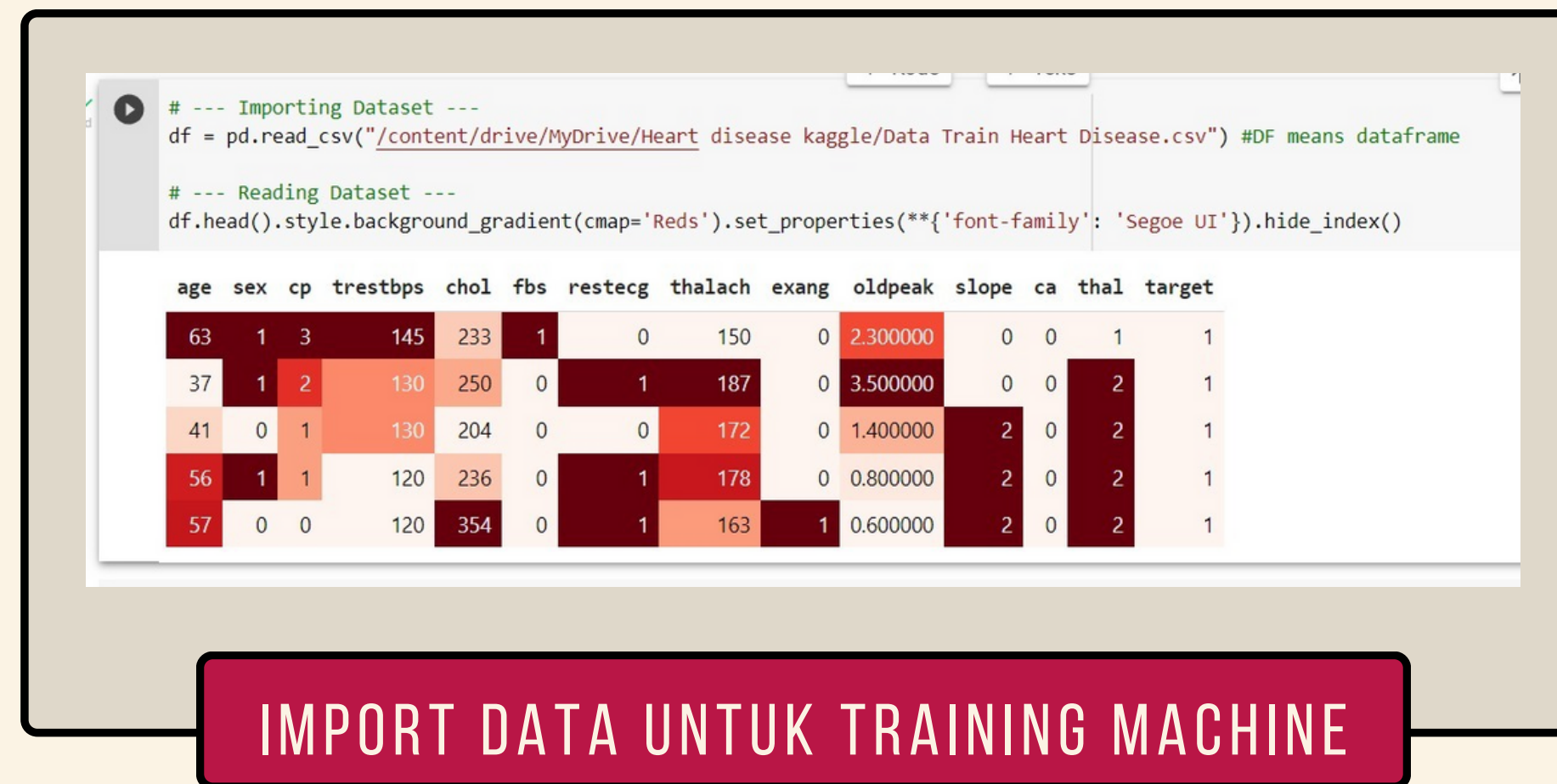
# METODE

Dari beberapa metode supervised learning, kami menggunakan metode Decision Tree

# DECISION TREE

Decision tree adalah algoritma machine learning yang menggunakan seperangkat aturan untuk membuat keputusan dengan struktur seperti pohon yang memodelkan kemungkinan hasil, biaya sumber daya, utilitas dan kemungkinan konsekuensi atau resiko. Konsepnya adalah dengan cara menyajikan algoritma dengan pernyataan bersyarat, yang meliputi cabang untuk mewakili langkah-langkah pengambilan keputusan yang dapat mengarah pada hasil yang menguntungkan.

# TAHAP-TAHAP MACHINE LEARNING



```
# --- Importing Dataset ---  
df = pd.read_csv("/content/drive/MyDrive/Heart disease kaggle/Data Train Heart Disease.csv") #DF means dataframe  
  
# --- Reading Dataset ---  
df.head().style.background_gradient(cmap='Reds').set_properties(**{'font-family': 'Segoe UI'}).hide_index()
```

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	2.300000	0	0	1	1
37	1	2	130	250	0	1	187	0	3.500000	0	0	2	1
41	0	1	130	204	0	0	172	0	1.400000	2	0	2	1
56	1	1	120	236	0	1	178	0	0.800000	2	0	2	1
57	0	0	120	354	0	1	163	1	0.600000	2	0	2	1

IMPORT DATA UNTUK TRAINING MACHINE

# TAHAP-TAHAP MACHINE LEARNING

```
[256] # Scaling data atas sampe bawah
cols = ['age', 'trestbps', 'chol', 'thalach', 'oldpeak']
for col in cols:
    minimum = min(df[col])
    maximum = max(df[col])
    df[col] = (df[col] - minimum) / (maximum - minimum)
# Data values now range from 0 to 1 after min-max normalization.
print(minimum)
print(maximum)
df.head()
```

0.0  
6.2

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	0.708333	1	3	0.481132	0.244292	1	0	0.603053	0	0.370968	0	0	1	1
1	0.166667	1	2	0.339623	0.283105	0	1	0.885496	0	0.564516	0	0	2	1
2	0.250000	0	1	0.339623	0.178082	0	0	0.770992	0	0.225806	2	0	2	1
3	0.562500	1	1	0.245283	0.251142	0	1	0.816794	0	0.129032	2	0	2	1
4	0.583333	0	0	0.245283	0.520548	0	1	0.702290	1	0.096774	2	0	2	1

**SCALING AGAR MACHINE LEARNING  
MUDAH MEMPROSES**

# TAHAP-TAHAP MACHINE LEARNING

```
[260] # --- Applying Decision Tree ---
      DTCclassifier = DecisionTreeClassifier(max_depth=3, min_samples_leaf=5, criterion='entropy', min_samples_split=5,
                                             splitter='random', random_state=1)

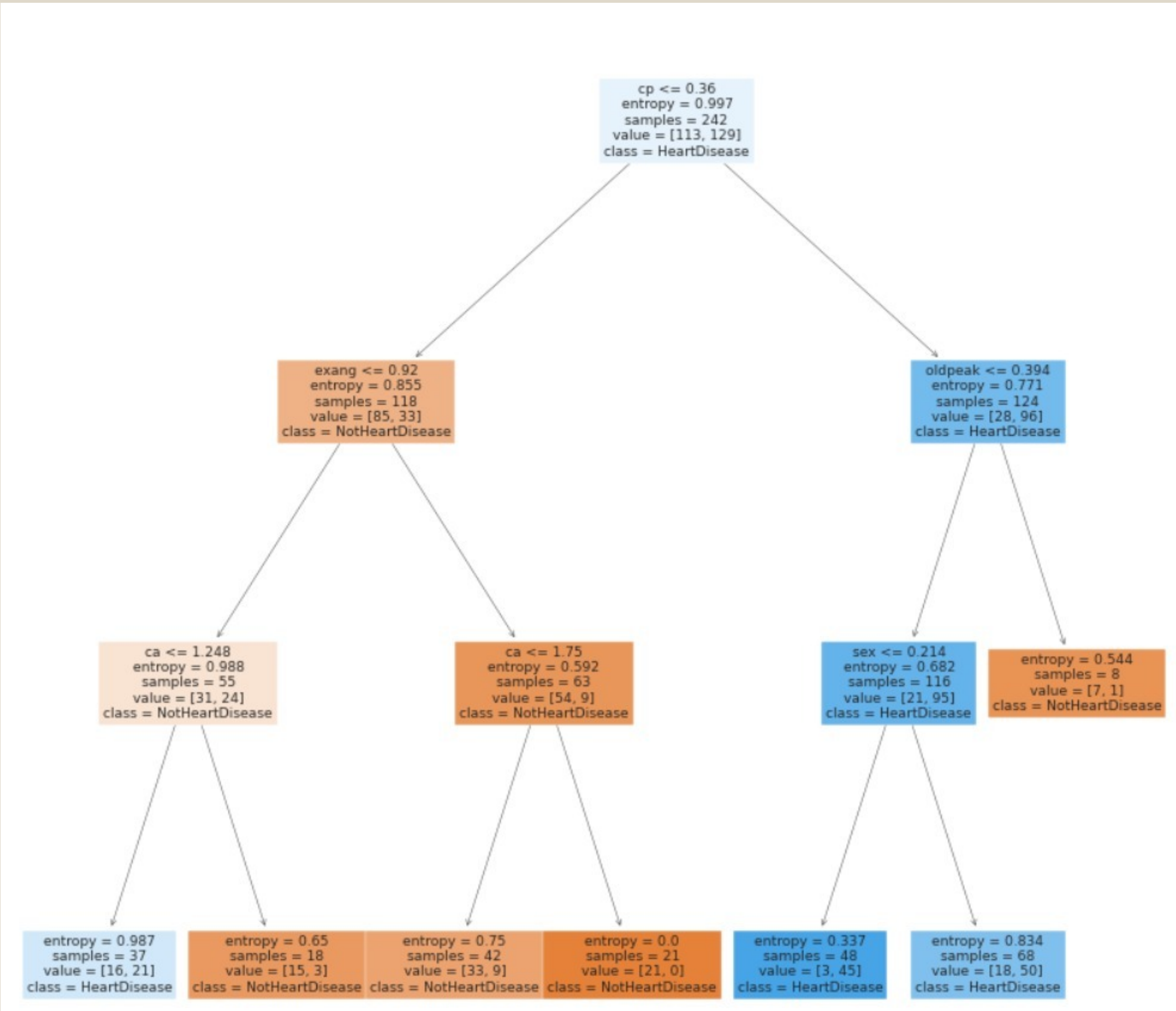
      DTCclassifier.fit(x_train, y_train)
      dt_clf = DTCclassifier.fit(x_train, y_train)
      y_pred_DTC = DTCclassifier.predict(x_test)

[266] feature_names = ["age", "sex", "cp", "trestbps", "chol", "fbs", "restecg", "thalach", "exang", "oldpeak", "slope", "ca", "thal"]
      class_names = ["NotHeartDisease", "HeartDisease"]

      fig = plt.figure(figsize=(10,10))
      plot = plot_tree(dt_clf,
                      feature_names=feature_names,
                      class_names=class_names,
                      filled=True)
```

**MACHINE LEARNING MENGGUNAKAN  
METODE DECISION TREE**





# TAHAP-TAHAP MACHINE LEARNING

↳ ... Decision Tree Accuracy: 83.61% ...

...: Classification Report

\*\*\*\*\*

	precision	recall	f1-score	support
0	0.86	0.72	0.78	25
1	0.82	0.92	0.87	36
accuracy			0.84	61
macro avg	0.84	0.82	0.83	61
weighted avg	0.84	0.84	0.83	61

AKURASI DECISION TREE

# TAHAP-TAHAP MACHINE LEARNING

```
] # --- Import Prediction ---  
prediction = pd.read_csv("/content/drive/MyDrive/Heart disease kaggle/Prediksi Heart Disease.csv")  
  
# --- Reading Prediction Dataset ---  
prediction.head().style.background_gradient(cmap='Reds').set_properties(**{'font-family': 'Segoe UI'}).hide_index()
```

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
32	0	3	145	233	1	0	136	0	2.300000	0	0	1
41	1	2	130	250	0	1	146	0	3.500000	0	0	2
59	0	1	130	204	0	0	167	0	1.400000	2	0	2
58	1	1	120	236	0	1	141	0	0.800000	2	0	2
43	0	0	120	354	0	1	153	1	0.600000	2	0	2

IMPORT DATA BARU UNTUK DI PREDIKSI  
OLEH MACHINE LEARNING

# TAHAP-TAHAP MACHINE LEARNING

```
# Ini scaling lengkap dari atas sampe bawah
cols = ['age', 'trestbps', 'chol', 'thalach', 'oldpeak']
for col in cols:
    minimum = min(prediction[col])
    maximum = max(prediction[col])
    prediction[col] = (df[col] - minimum) / (maximum - minimum)
# Data values now range from 0 to 1 after min-max normalization.
prediction.head()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
0	-0.682292	0	3	-0.882253	-0.287113	1	0	-2.156616	0	0.059834	0	0	1
1	-0.695833	1	2	-0.883588	-0.287025	0	1	-2.151908	0	0.091051	0	0	2
2	-0.693750	0	1	-0.883588	-0.287265	0	0	-2.153817	0	0.036420	2	0	2
3	-0.685937	1	1	-0.884478	-0.287098	0	1	-2.153053	0	0.020812	2	0	2
4	-0.685417	0	0	-0.884478	-0.286483	0	1	-2.154962	1	0.015609	2	0	2

**DATA PREDIKSI DI SCALING AGAR  
MUDAH DIPROSES OLEH MACHINE  
LEARNING UNTUK DAPAT MENGIKUTI  
ALGORITMA PADA DATA TRAINING YANG  
DI SCALING**

# TAHAP-TAHAP MACHINE LEARNING

```
# --- Prediction using Random Forest ---  
result = DTClassifier.predict(prediction)  
prediction['Hasil'] = result  
print(prediction)  
prediction.to_csv('Hasil Prediksi.csv', index=False)  
  
# --- Print Heart Disease Status ---  
# if result[0] == 1:  
#     print('\033[1m' + '... Heart Disease Detected!..' + '\033[0m')  
# else:  
#     print('\033[1m' + '... Heart Disease Not Detected!..' + '\033[0m')
```

MENERAPKAN DECISION TREE UNTUK  
MEMPREDIKSI DATA YANG BARU



# TAHAP-TAHAP MACHINE LEARNING

Prediksi Heart Disease.csv X Data Train Scaled.csv Hasil Prediksi.csv

1 to 10 of 303 entries Filter

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
32	0	3	145	233	1	0	136	0	2.3	0	0	1
41	1	2	130	250	0	1	146	0	3.5	0	0	2
59	0	1	130	204	0	0	167	0	1.4	2	0	2
58	1	1	120	236	0	1	141	0	0.8	2	0	2
43	0	0	120	354	0	1	153	1	0.6	2	0	2
45	1	0	140	192	0	1	180	0	0.4	1	0	1
44	0	1	140	294	0	0	144	0	1.3	1	0	2
38	1	1	120	263	0	1	159	0	0	2	0	3
67	1	2	172	199	1	1	130	0	0.5	2	0	3
51	1	2	150	168	0	1	163	0	1.6	2	0	2

Show 10 per page

1 2 10 30 31

**DATA PREDIKSI SEBELUM DI SCALING  
DAN TANPA ADA HASIL PREDIKSI  
APAKAH HEART DISEASE ATAU TIDAK**

# TAHAP-TAHAP MACHINE LEARNING

Prediksi Heart Disease.csv Hasil Prediksi.csv X

1 to 10 of 303 entries Filter

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	Hasil
-0.6822916666666667	0	3	-0.8822534709861161	-0.28711348804236775	1	0	-2.156615776081425	0	0.05983350676378771	0	0	1	1
-0.6958333333333333	1	2	-0.8835884656461374	-0.28702487437709806	0	1	-2.151908396946565	0	0.09105098855358999	0	0	2	1
-0.69375	0	1	-0.8835884656461374	-0.2872646525301808	0	0	-2.15381679389313	0	0.036420395421436	2	0	2	1
-0.6859375	1	1	-0.8844784620861517	-0.2870978503367319	0	1	-2.1530534351145034	0	0.02081165452653486	2	0	2	1
-0.6854166666666667	0	0	-0.8844784620861517	-0.28648276724838934	0	1	-2.1549618320610686	1	0.015608740894901144	2	0	2	0
-0.6854166666666667	1	0	-0.8826984692061232	-0.2873272033527241	0	1	-2.1568702290076334	0	0.01040582726326743	1	0	1	1
-0.6859375	0	1	-0.8826984692061232	-0.28679552136110587	0	0	-2.1562340966921116	0	0.033818938605619145	1	0	2	1
-0.6921875	1	1	-0.8844784620861517	-0.28695711098600946	0	1	-2.1536895674300256	0	0.0	2	0	3	1
-0.6880208333333333	1	2	-0.8798504805980777	-0.2872907153729071	1	1	-2.1550890585241733	0	0.013007284079084287	2	0	3	1
-0.6854166666666667	1	2	-0.881808472766109	-0.2874523049978107	0	1	-2.1535623409669213	0	0.04162330905306972	2	0	2	1

Show 10 per page

1 2 10 30 31

SETELAH DI SCALING DAN DIPREDIKSI  
OLEH MACHINE DAN DIDAPATKAN HASIL  
PADA KOLOM SEBELAH KANAN YAITU 0  
DAN 1. JIKA HASIL 0 MAKA NOT HEART  
DISEASE DAN 1 HEART DISEASE

# PENUTUPAN

PROYEK INI BERFOKUS UNTUK MENYELIDIKI SUATU ALGORITMA, APAKAH MEMILIKI TINGKAT AKURASI YANG TINGGI GUNA MENDETEKSI PENYAKIT JANTUNG MELALUI OBJEK DATASET (HEART DISEASE). DECISION TREE MENCAPAI SKOR AKURASI TINGGI UNTUK MEMPREDIKSI KEMUNGKINAN SERANGAN JANTUNG. SETELAH MENGEKSEKUSI DATASHEET DENGAN ALGORITMA YANG DIPILIH DIDAPATKANLAH ALGORITMA DECISION TREE DENGAN AKURASI 83.61%. DENGAN ADANYA PROYEK INI KITA MAMPU MENINGKATKAN KEWASPADAAN TERHADAP RESIKO DETAK JANTUNG SEHINGGA DAPAT DITINDAKLANJUTI SEJAK DINI.



---

*sepian*  
**TERIMA KASIH**

---