



## Perbandingan Performansi Algoritma Pengklasifikasian Terpandu Untuk Kasus Penyakit Kardiovaskular

Adi Nugroho<sup>1</sup>, Agustinus Bimo Gumelar<sup>2</sup>, Adri Gabriel Sooi<sup>3</sup>, Dyana Sarvasti<sup>4</sup>, Paul L Tahalele<sup>5</sup>

<sup>1,2</sup>Sistem Informasi, Fakultas Ilmu Komputer, Universitas Narotama, Surabaya, Indonesia

<sup>3</sup>Ilmu Komputer, Fakultas Teknik, Universitas Katolik Widya Mandira, Kupang, Indonesia

<sup>4</sup>Ilmu Penyakit Jantung, Fakultas Kedokteran, Universitas Katolik Widya Mandala, Surabaya, Indonesia

<sup>5</sup>Ilmu Bedah Toraks, Kardiak dan Vaskular, Fakultas Kedokteran, Universitas Katolik Widya Mandala, Surabaya, Indonesia  
<sup>1</sup>adinugroho696@gmail.com, <sup>2</sup>bimogumelar@ieee.org, <sup>3</sup>adrigabriel@ieee.org, <sup>4</sup>dyana@ukwms.ac.id, <sup>5</sup>pltahalele@yahoo.com

### Abstract

One of the health problems that occur in Indonesia is the increasing number of NCD (Non-Communicable Disease) such as heart attack and cardiovascular disease. There are two factors that cause cardiovascular disease, i.e. factor that can be changed and cannot be changed. This study aim to analyze the best performance of several classification algorithms such as *k*-nearest neighbors algorithm (*k*-NN), stochastic gradient descent (SGD), random forest (RF), neural network (NN) and logistic regression (LR) in classifying cardiovascular based on factors that caused those diseases. There are two aspects that need to be examined, the performance of each algorithm which is evaluated using the Confusion matrix method with the parameters of accuracy, precision, recall and AUC (Area Under the Curve). The dataset uses 425.195 samples from result data of cardiovascular disease diagnosed. The testing mode uses percentage split and cross-validation technique. The experimental results show that the performance of NN algorithms produces the best prediction accuracy compared to other algorithms, which is accuracy of 89.60%, AUC of 0.873, precision of 0.877, and recall of 0.896 using percentage split and cross-validation testing mode using Orange. For the accuracy of 89.46%, AUC of 0.865, precision of 0.875, and recall of 0.895 using cross-validation testing mode using Weka. By KNIME, the result of accuracy value is 88.55%, AUC value is 0.768, precision value is 0.854, and recall value is 0.886 using cross-validation testing mode.

**Keywords:** data mining, classification, cardiovascular disease, supervised learning algorithm, confusion matrix.

### Abstrak

Salah satu permasalahan kesehatan yang terjadi di Indonesia adalah meningkatnya angka kejadian Penyakit Tidak Menular (PTM) seperti penyakit jantung dan pembuluh darah (kardiovaskular). Terdapat dua faktor risiko yang menyebabkan terjadinya penyakit kardiovaskular yaitu faktor risiko yang bisa diubah dan faktor risiko yang tidak bisa diubah. Penelitian ini mengkaji untuk menganalisa kinerja terbaik dari beberapa algoritma pengklasifikasian terpandu yaitu *k*-nearest neighbors (*k*-NN), stochastic gradient descent (SGD), random forest (RF), neural network (NN) dan logistic regression (LR) dalam mengklasifikasikan penyakit kardiovaskular berdasarkan faktor-faktor risiko penyebab terjadinya penyakit tersebut. Aspek yang akan dikaji adalah hasil kinerja dari masing-masing algoritma yang dievaluasi menggunakan metode *Confusion matrix* dengan parameter akurasi, presisi, *recall* dan AUC (*Area Under the Curve*). Basis data yang digunakan adalah data hasil pemeriksaan penyakit kardiovaskular sejumlah 425.195 data sampel. Mode pengujian yang digunakan adalah pembagian prosentase dan validasi silang. Hasil eksperimen menunjukkan bahwa kinerja algoritma NN menghasilkan kinerja terbaik dibandingkan algoritma yang lain yaitu nilai akurasi sebesar 89.60%, nilai AUC sebesar 0.873, nilai presisi sebesar 0.877 dan nilai *recall* sebesar 0.896 menggunakan mode pengujian pembagian prosentase dan mode pengujian validasi silang pada *tools* Orange. Pada *tools* Weka, didapat nilai akurasi sebesar 89.46%, nilai AUC sebesar 0.865, nilai presisi sebesar 0.875 dan nilai *recall* sebesar 0.895 menggunakan mode pengujian validasi silang. Dengan *tools* KNIME, didapatkan nilai akurasi sebesar 88.55%, nilai AUC sebesar 0.768, nilai presisi sebesar 0.854 dan nilai *recall* sebesar 0.886 menggunakan mode pengujian validasi silang.

**Kata kunci:** data mining, klasifikasi, penyakit kardiovaskular, algoritma pengklasifikasian terpandu, *confusion matrix*.

## 1. Pendahuluan

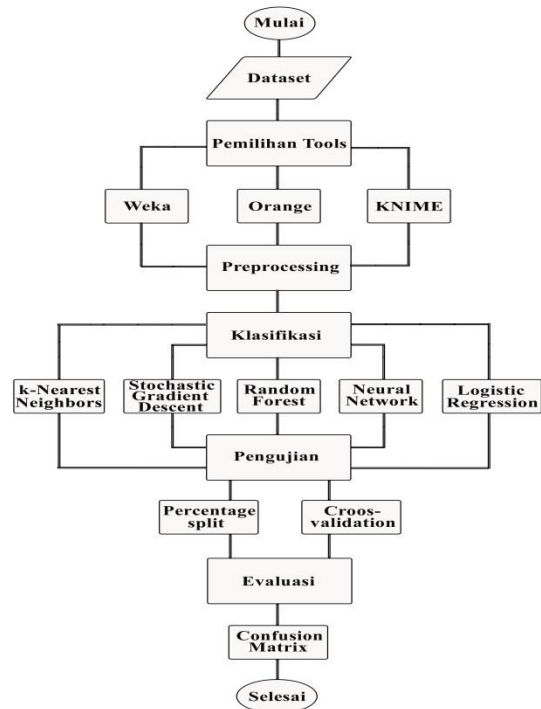
Penyakit kardiovaskular merupakan suatu penyakit yang disebabkan karena terjadinya gangguan pada fungsi jantung dan pembuluh darah seperti hipertensi, jantung koroner, stroke dan gagal jantung [1]. Penyakit ini menjadi penyebab utama terjadinya kasus kematian diseluruh dunia [2]. Menurut data dari RISKESDAS (Riset Kesehatan Dasar) tahun 2018, terjadinya kasus penyakit kardiovaskular di Indonesia mengalami peningkatan setiap tahunnya. Setidaknya sebanyak 2.784.064 penduduk Indonesia menderita penyakit jantung [3]. Terdapat dua faktor risiko terjadinya penyakit kardiovaskular yaitu faktor risiko yang tidak bisa diubah seperti riwayat penyakit keluarga, usia, jenis kelamin, dan obesitas. Sedangkan terdapat faktor risiko yang bisa diubah yaitu : *dislipidemia* atau kondisi kolestrol yang tidak normal, *diabetes melitus* atau peningkatan kadar glukosa, hipertensi atau peningkatan tekanan darah, stres, diet yang tidak sehat dan kurangnya aktifitas fisik [2].

Tujuan penelitian ini adalah membandingkan kinerja dari beberapa algoritma pengklasifikasian terpandu yaitu *k-nearest neighbors*, *stochastic gradient descent*, *random forest*, *neural network* dan *logistic regression* dalam mengklasifikasikan penyakit kardiovaskular berdasarkan faktor risiko penyebab terjadinya penyakit tersebut. 1. mengukur kinerja dari lima algoritma tersebut menggunakan metode *Confusion matrix* dengan parameter akurasi, presisi, *recall* dan *AUC (Area Under the Curve)*. 2. mengevaluasi lima algoritma tersebut menggunakan tiga *tools data mining* yaitu Weka versi 3.83, Orange versi 3.24.1, dan KNIME versi 3. Mode pengujian yang digunakan pada ketiga *tools* adalah pembagian prosentase dan validasi silang .

Sistematika penulisan artikel ini tersusun menjadi beberapa bagian. Pada bagian 1 merupakan introduksi yang membahas tentang penyakit kardiovaskular. Pada bagian 2 dijelaskan mengenai rincian metode penelitian yang digunakan dalam penelitian ini, mulai dari pengumpulan data, informasi atribut yang digunakan, pemilihan perangkat lunak yang akan dibandingkan dalam pekerjaan ini, tahap pra-proses (*preprocessing*), metode klasifikasi, metode pengujian dan metode evaluasi yang digunakan dalam penelitian ini. Selanjutnya pada bagian 3, dijelaskan tentang hasil dan pembahasan yang meliputi evaluasi prediksi dan *metric performance*. Untuk bagian akhir yaitu pada bagian 4, penyampaian kesimpulan terhadap keseluruhan proses penelitian yang dilakukan.

## 2. Metode Penelitian

Metode penelitian yang digunakan didalam penelitian dapat dilihat pada Gambar 1.



Gambar 1. Flowchart Metode Penelitian

### 2.1. Dataset

*Dataset* yang digunakan merupakan hasil dari penelitian [4] yaitu data hasil pemeriksaan pasien terhadap penyakit kardiovaskular sebesar 425.196 data sampel. *Dataset* tersebut terdiri dari dua kelas yaitu positif kardiovaskular dan negatif kardiovaskular.

Tabel 1. Informasi Detail *Dataset*

Informasi	Detail
Dataset	Data Penyakit Kardiovaskular
Tipe File	Comma-separated values (CSV)
Banyak Data	425.196
Banyak Atribut	11
Karakteristik Atribut	Real
Kelas	Positif Kardiovaskular dan Negatif Kardiovaskular
Distribusi Kelas	Tidak Merata
Karakteristik Dataset	Single Variate

Atribut data yang terdapat didalam *dataset* merupakan faktor risiko penyebab terjadinya penyakit kardiovaskular diantaranya :

- Age* : kategori umur pasien (0-20 = 0, 20-40 = 1, 40-55 = 2, 55-70 = 3, 70-90 = 4)
- NumAge* : usia pasien (0-90)
- Race* : kategori ras/suku : (American Indian=0, Asian / Pasific Islander=1, Black/African American=2, White=3)
- Hypertension* : pasien penderita hipertensi (0=Tidak, 1=Ya)
- Treatment* : pasien menjalani pengobatan hipertensi (0=Tidak, 1=Ya)

f. <i>Smoking</i> : pasien seorang perokok (0=Tidak, 1=Ya)	<i>Treatment</i>	0
g. <i>Gender</i> : jenis kelamin pasien (0=Wanita, 1=Pria)	<i>Smoking</i>	0
h. <i>Diabetes</i> : pasien penderita diabetes (0=Tidak, 1=Ya)	<i>Gender</i>	0
i. <i>Body Mass Index</i> : indeks massa tubuh (15 – 36)	<i>Diabetes</i>	0
j. <i>Cholesterol</i> : jumlah kolesterol pasien (155 – 245)	<i>Body Mass Index</i>	0
k. <i>Systolic</i> : tekanan darah pasien (74 – 266).	<i>Cholesterol</i>	0
	<i>Systolic</i>	0

## 2.2. Pemilihan Tools

Proses pengujian kinerja pada algoritma klasifikasi dilakukan menggunakan tiga *tools data mining* yaitu Weka, Orange, dan KNIME.

### a. Weka

Weka merupakan perangkat lunak *open source* yang didalamnya terdapat berbagai macam algoritma yang dapat melakukan tugas pada *data mining* seperti *preprocessing* data, klasifikasi, klasterisasi, visualisasi data dan lain sebagainya [5]. Pada penelitian ini digunakan tools Weka versi 3.8.3.

### b. Orange

Orange adalah *platform* analisis data yang berbasis *python*. Perangkat lunak ini menggunakan pemrograman visual dengan fungsi utama yaitu analisis dan visualisasi data [5]. Perangkat lunak Orange yang digunakan didalam penelitian ini adalah versi 3.24.1.

### c. KNIME

KNIME adalah perangkat lunak analisis data yang mampu melakukan tiga tahap pengolahan data sekaligus, yaitu ekstraksi, transformasi dan pemuatan data [5]. Berbasis bahasa pemrograman JAVA dan menyediakan berbagai macam modul integrasi data. Dalam penelitian ini digunakan versi 3.

## 2.3 Preprocessing

Tahapan *preprocessing* data didalam penelitian ini terdiri dari dua tahap yaitu pemilihan atribut data dan *filtering dataset*.

a. **Pemilihan Atribut** : terdapat 11 atribut didalam dataset saling berhubungan satu sama lain. Sehingga pada penelitian ini digunakan semua atribut dari *dataset* tersebut

b. **Filtering Dataset** : Pemeriksaan data kosong pada setiap atribut *dataset*. Filter data pada ketiga tools yang digunakan yaitu :

1. Weka : menggunakan filter *replace missing values*.
2. Orange : menggunakan filter *impute*.
3. KNIME : menggunakan filter *missing value*.

Tabel 2. Hasil Pengecekan *Missing Value*

Nama Atribut	<i>Missing Value Rate (%)</i>
<i>Age</i>	0
<i>NumAge</i>	0
<i>Race</i>	0
<i>Hypertension</i>	0

## 2.4. Metode Klasifikasi

Proses klasifikasi *dataset* menggunakan lima algoritma pengklasifikasian terpandu yaitu *k-nearest neighbors*, *stochastic gradient descent*, *random forest*, *logistic regression*, dan *neural network*.

### a. *k-Nearest Neighbors (k-NN)*

Algoritma k-NN adalah suatu algoritma nonparametrik yang bisa memprediksi sebuah kelas dari suatu objek berdasarkan kelas terdekatnya [6]. Tahapan algoritma k-NN dimulai dari menghitung jarak ( $N_0$ ) dari pengamatan  $y_i$  hingga ke pengamatan  $y_j$ . Selanjutnya mencari probabilitas bersyarat untuk kelas  $j$  seperti pada persamaan 1.

$$Pr(Y = j | X = x_0) = \frac{1}{k} \sum_{i \in N_0} I(y_i = j) \quad (1)$$

### b. *Stochastic Gradient Descent (SGD)*

Algoritma SGD secara kontras melakukan pembaharuan terhadap parameter setiap pelatihan. SGD mengoptimalkan setiap fungsi berdasarkan gradien yang bergerak secara menurun. SGD melakukan penghapusan redundansi dan pembaharuan parameter secara bersamaan sehingga memiliki kinerja yang lebih cepat [7].

### c. *Random Forest (RF)*

Algoritma RF merupakan sebuah pendekatan pembelajaran mesin berbasis pohon keputusan yang memerlukan dua parameter yaitu jumlah pohon ( $n_{tree}$ ) dan jumlah fitur pemisah ( $mtry$ ) [8]. Nilai standar/bawaan  $mtry$  dijelaskan pada persamaan 2.

$$mtry = \sqrt{p} \quad (2)$$

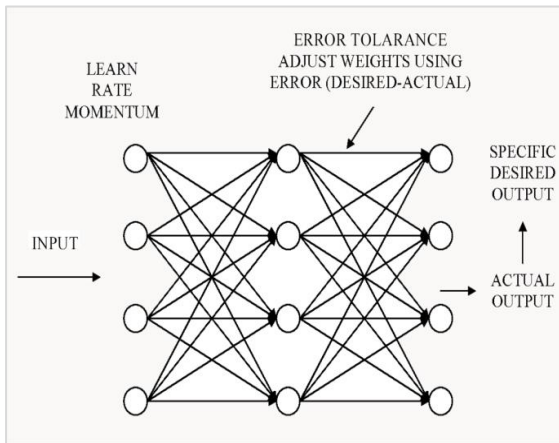
dengan  $p$  adalah jumlah variabel prediktor yang digunakan didalam proses klasifikasi.

### d. *Logistic Regression (LR)*

LR merupakan bagian dari analisa regresi yang terdiri dari variabel respon dan variabel dikotomi. LR digunakan ketika nilai variabel respon sama dengan nilai variabel dikotomi. Variabel dikotomi memiliki dua nilai yang mewakili ada tidaknya suatu kejadian [9].

### e. *Neural Network (NN)*

NN merupakan algoritma yang mengadopsi sistem kerja dari otak manusia, digunakan untuk membuat suatu model yang menggambarkan kompleksitas relasi antara *input* dan *output* dengan tujuan menemukan pola hubungan antar *neuron* [9].



Gambar 2 Arsitektur Neural Network [10]

Algoritma NN memiliki susunan yang berlapis yaitu *input layer*, lapisan tersembunyi (*hidden layer*), dan *output layer* [11].

### 2.5. Model Pengujian

Model pengujian proses klasifikasi terdiri dari dua mode pengujian yaitu pembagian prosentase dan validasi silang.

#### a. Pembagian prosentase

Mode pengujian dengan cara membagi *dataset* menjadi data *training* dan data *testing*. Komposisi data yang digunakan adalah 80% untuk data *training* dan 20% untuk data *testing*.

#### b. Validasi silang

Mode pengujian dengan cara membagi data *training* menjadi *n* subhimpunan. Dengan *n* adalah nilai dari *fold*. Banyaknya nilai *fold* yang digunakan didalam penelitian ini sebesar 10 *fold*. Artinya data akan diuji sepuluh kali, *dataset* akan menjadi data *testing* sebanyak satu kali dan menjadi data *training* sebanyak *n - 1*.

### 2.6. Model Evaluasi

Model evaluasi yang digunakan untuk mengukur performa dari algoritma klasifikasi adalah *Confusion matrix*. Model ini menunjukkan perbandingan hasil klasifikasi yang dilakukan oleh model dengan data yang sebenarnya dalam bentuk sebuah tabel *matrix* seperti pada Tabel 2.

Tabel 2. *Confusion Matrix*

	Predicted		Total Actual
	Positif	Negatif	
Actual	Positif	FPC	aPC
	Negatif	FNC	aNC
Total Predicted	pPC	pNC	nData

Dengan *PC* adalah data positif kardiovaskular yang diprediksi dengan benar. *NC* adalah data negatif kardiovaskular yang diprediksi dengan benar. *FPC*

adalah data positif kardiovaskular yang diprediksi sebagai data negatif kardiovaskular. *FNC* adalah data negatif kardiovaskular yang diprediksi sebagai data positif kardiovaskular. *aPC* adalah total data positif kardiovaskular yang sebenarnya. *aNC* adalah total data negatif kardiovaskular yang sebenarnya. *pPC* adalah total hasil prediksi positif kardiovaskular. *pNC* adalah total hasil prediksi negatif kardiovaskular. Sedangkan *nData* merupakan total *instance dataset* yang digunakan didalam proses klasifikasi. Parameter yang digunakan untuk mengukur *performance metrics* sebagai acuan dalam menilai kinerja algoritma adalah akurasi, presisi, dan *recall* dan AUC [12]. Nilai akurasi dapat dihitung menggunakan persamaan 3.

$$\text{Akurasi} = \frac{PC+NC}{pPC+pNC} \quad (3)$$

Akurasi merupakan rasio prediksi benar dari keseluruhan data, dan nilai presisi merupakan perbandingan rasio prediksi benar dengan keseluruhan hasil yang diprediksi dengan benar [13]. Untuk mencari nilai presisi dari setiap kelas digunakan persamaan 4 dan 5.

$$\text{Presisi (PC)} = \frac{PC}{pPC} \quad (4)$$

$$\text{Presisi (NC)} = \frac{NC}{pNC} \quad (5)$$

Selanjutnya mencari nilai presisi dari keseluruhan kelas menggunakan persamaan 6.

$$\text{Presisi} = \left( \frac{aPC}{nData} * \text{Presisi PC} + \frac{aNC}{nData} * \text{Presisi NC} \right) \quad (6)$$

Sedangkan *recall* merupakan perbandingan rasio prediksi benar dengan keseluruhan data yang benar [13]. Nilai *recall* dari setiap kelas dapat dicari menggunakan persamaan 7 dan 8.

$$\text{Recall (PC)} = \frac{PC}{aPC} \quad (7)$$

$$\text{Recall (NC)} = \frac{NC}{aNC} \quad (8)$$

Selanjutnya menentukan nilai *recall* pada keseluruhan kelas menggunakan persamaan 9.

$$\text{Recall} = \left( \frac{aPC}{nData} * \text{recall PC} + \frac{aNC}{nData} * \text{recall NC} \right) \quad (9)$$

Kurva ROC (*Receiver Operator Characteristic*) digunakan untuk membandingkan hasil kinerja dari suatu algoritma dalam bentuk visual. Visualisasi pada kurva ROC didasarkan pada nilai dari *Confusion matrix*. Kurva ROC berbentuk grafik dua dimensi dengan sumbu x adalah rasio dari *false-positive* dan sumbu y adalah rasio dari *true-positive*. Sedangkan untuk menghitung perbedaan performa dari algoritma digunakan nilai AUC (*Area Under the Curve*) [13]. Kurva ROC memiliki lima pengkategorian terhadap hasil kinerja suatu algoritma diantaranya :

- Excellent Classification* : AUC bernilai 0.90-1.00
- Good Classification* : AUC bernilai 0.80-0.90

- c. *Fair Classification* : AUC bernilai 0.70-0.80
- d. *Poor Classification* : AUC bernilai 0.60-0.70
- e. *Failure Classification* : AUC bernilai 0.50-0.60

	Validasi silang (10-folds)	89.60	10.40
LR	Pembagian prosentase (80%)	89.40	10.60
	Validasi silang (10-folds)	89.50	10.50

### 3. Hasil dan Pembahasan

Evaluasi hasil akurasi prediksi dari algoritma k-NN, SGD, RF, NN dan LR diukur menggunakan parameter *Correctly Classified* dan *Incorrectly Classified*.

Tabel 3. Hasil Evaluasi Akurasi Prediksi pada *Tools Weka*

Algoritma	Mode Pengujian	Hasil Akurasi Klasifikasi	
		Akurasi benar (%)	Akurasi salah (%)
k-NN	Pembagian prosentase (80%)	84.75	15.25
	Validasi silang (10-folds)	84.96	15.04
SGD	Pembagian prosentase (80%)	89.37	10.63
	Validasi silang (10-folds)	89.37	10.63
RF	Pembagian prosentase (80%)	88.65	11.35
	Validasi silang (10-folds)	88.70	11.30
NN	Pembagian prosentase (80%)	89.32	10.68
	Validasi silang (10-folds)	89.46	10.54
LR	Pembagian prosentase (80%)	89.37	10.63
	Validasi silang (10-folds)	89.48	10.52

Hasil evaluasi pada Tabel 3 menunjukkan bahwa pada mode pengujian pembagian prosentase, hasil akurasi prediksi terbaik diraih oleh algoritma SGD dan LR dengan nilai akurasi sebesar 89.37%. Sedangkan pada mode pengujian validasi silang, hasil akurasi prediksi terbaik diraih oleh algoritma LR dengan nilai akurasi sebesar 89.48%.

Tabel 4. Hasil Evaluasi Akurasi Prediksi pada *Tools Orange*

Algoritma	Mode Pengujian	Hasil Akurasi Klasifikasi	
		Akurasi benar (%)	Akurasi salah (%)
k-NN	Pembagian prosentase (80%)	87.90	12.10
	Validasi silang (10-folds)	87.90	12.10
SGD	Pembagian prosentase (80%)	88.50	11.50
	Validasi silang (10-folds)	88.60	11.40
RF	Pembagian prosentase (80%)	88.60	11.40
	Validasi silang (10-folds)	88.60	11.40
NN	Pembagian prosentase (80%)	89.60	10.40

Hasil evaluasi pada Tabel 4 menunjukkan bahwa algoritma RF dan NN meraih hasil akurasi prediksi terbaik dengan nilai akurasi sebesar 89.60% baik pada mode pengujian pembagian prosentase dan mode pengujian validasi silang.

Tabel 5. Hasil Evaluasi Akurasi Prediksi pada *Tools KNIME*

Algoritma	Mode Pengujian	Hasil Akurasi Klasifikasi	
		Akurasi benar (%)	Akurasi salah (%)
k-NN	Pembagian prosentase (80%)	87.82	12.18
	Validasi silang (10-folds)	87.71	12.29
SGD	Pembagian prosentase (80%)	89.38	10.62
	Validasi silang (10-folds)	89.46	10.54
RF	Pembagian prosentase (80%)	89.56	10.44
	Validasi silang (10-folds)	89.56	10.44
NN	Pembagian prosentase (80%)	88.27	11.73
	Validasi silang (10-folds)	88.55	11.45
LR	Pembagian prosentase (80%)	89.49	10.51
	Validasi silang (10-folds)	89.48	10.52

Hasil evaluasi pada Tabel 5 menunjukkan bahwa algoritma RF meraih hasil akurasi prediksi terbaik dengan nilai akurasi sebesar 89.56% baik pada mode pengujian pembagian prosentase dan mode pengujian validasi silang. Dengan perhitungan klasifikasi menggunakan algoritma k-NN, SGD, RF, NN, dan LR, dapat disimpulkan bahwa hasil akurasi prediksi terbaik menggunakan *tools Weka* diraih oleh algoritma LR pada mode pengujian validasi silang. Sedangkan pada *tools Orange*, hasil akurasi prediksi terbaik diraih oleh algoritma RF dan NN pada mode pengujian pembagian prosentase dan mode pengujian validasi silang. Serta pada *tools KNIME*, hasil akurasi prediksi terbaik diraih oleh algoritma SGD dan LR pada mode pengujian validasi silang. Setelah melakukan evaluasi terhadap hasil akurasi prediksi dari masing-masing algoritma, tahap selanjutnya adalah melakukan evaluasi terhadap kinerja algoritma menggunakan parameter *performance metrics* yaitu AUC, presisi, dan *recall*. Hasil evaluasi menggunakan *tools Weka* dapat dilihat pada Tabel 6.

Tabel 6. Hasil Evaluasi *Performance Metrics* pada *Tools* Weka

Algoritma	Mode Pengujian	<i>Performance Metrics</i>		
		AUC	Presisi	<i>Recall</i>
k-NN	Pembagian prosentase (80%)	0.639	0.848	0.847
	Validasi silang (10-folds)	0.640	0.849	0.850
SGD	Pembagian prosentase (80%)	0.868	0.874	0.894
	Validasi silang (10-folds)	0.868	0.875	0.894
RF	Pembagian prosentase (80%)	0.840	0.866	0.886
	Validasi silang (10-folds)	0.841	0.867	0.887
NN	Pembagian prosentase (80%)	0.870	0.876	0.893
	Validasi silang (10-folds)	0.865	0.875	0.895
LR	Pembagian prosentase (80%)	0.868	0.874	0.894
	Validasi silang (10-folds)	0.869	0.876	0.895

Hasil evaluasi *performance metrics* pada Tabel 6 menunjukkan bahwa algoritma LR memiliki nilai *performance metrics* terbaik dibandingkan algoritma yang lain dengan nilai AUC sebesar 0.869, nilai presisi sebesar 0.876, dan nilai *recall* sebesar 0.895 pada mode pengujian validasi silang. Berdasarkan nilai AUC tersebut, maka kinerja algoritma LR termasuk dalam kategori *Good Classification*.

Tabel 7. Hasil Evaluasi *Performance Metrics* pada *Tools* Orange

Algoritma	Mode Pengujian	<i>Performance Metrics</i>		
		AUC	Presisi	<i>Recall</i>
k-NN	Pembagian prosentase (80%)	0.755	0.854	0.879
	validasi silang (10-folds)	0.755	0.853	0.879
SGD	Pembagian prosentase (80%)	0.611	0.863	0.886
	validasi silang (10-folds)	0.589	0.856	0.883
RF	Pembagian prosentase (80%)	0.825	0.866	0.886
	validasi silang (10-folds)	0.825	0.866	0.866
NN	Pembagian prosentase (80%)	0.873	0.877	0.896
	validasi silang (10-folds)	0.873	0.877	0.896
LR	Pembagian prosentase (80%)	0.868	0.875	0.894
	validasi silang (10-folds)	0.868	0.875	0.895

Hasil evaluasi *performance metrics* pada Tabel 7 menunjukkan bahwa algoritma NN memiliki nilai *performance metrics* terbaik dibandingkan algoritma

yang lain dengan nilai AUC sebesar 0.873, nilai presisi sebesar 0.877, dan nilai *recall* sebesar 0.896 pada mode pengujian pembagian prosentase dan mode pengujian validasi silang. Berdasarkan nilai AUC tersebut, maka kinerja algoritma NN termasuk dalam kategori *Good Classification*.

Tabel 8. Hasil Evaluasi *Performance Metrics* pada *Tools* KNIME

Algoritma	Mode Pengujian	<i>Performance Metrics</i>		
		AUC	Presisi	<i>Recall</i>
K-NN	Pembagian prosentase (80%)	0.692	0.845	0.870
	Validasi silang (10-folds)	0.691	0.844	0.887
SGD	Pembagian prosentase (80%)	0.869	0.876	0.895
	Validasi silang (10-folds)	0.869	0.875	0.895
RF	Pembagian prosentase (80%)	0.825	0.877	0.896
	Validasi silang (10-folds)	0.824	0.877	0.896
NN	Pembagian prosentase (80%)	0.847	0.883	0.883
	Validasi silang (10-folds)	0.768	0.854	0.886
LR	Pembagian prosentase (80%)	0.869	0.876	0.895
	Validasi silang (10-folds)	0.869	0.876	0.895

Hasil evaluasi *performance metrics* pada Tabel 8 menunjukkan bahwa algoritma SGD dan LR memiliki nilai *performance metrics* terbaik dibandingkan algoritma yang lain dengan nilai AUC sebesar 0.869, nilai presisi sebesar 0.876, dan nilai *recall* sebesar 0.895 pada mode pengujian pembagian prosentase dan mode pengujian validasi silang. Berdasarkan nilai AUC tersebut, maka kinerja algoritma SGD dan LR termasuk dalam kategori *Good Classification*.

Berdasarkan hasil perbandingan *performance metrics* pada tabel 6-8, dapat disimpulkan bahwa hasil kinerja klasifikasi terbaik menggunakan *tools* Weka diraih oleh algoritma LR pada mode pengujian validasi silang. Sedangkan pada *tools* Orange, hasil kinerja terbaik diraih oleh algoritma NN pada mode pengujian pembagian prosentase dan mode pengujian validasi silang. Serta pada *tools* KNIME, hasil akurasi prediksi terbaik diraih oleh algoritma SGD dan LR pada mode pengujian validasi silang.

Dari hasil evaluasi kinerja dari masing-masing algoritma, maka dapat dibentuk sebuah tabel *Confusion Matrix* berdasarkan kinerja terbaik menggunakan *tools* Weka, Orange, dan KNIME. Tabel *Confusion Matrix* dari algoritma k-NN dapat dilihat pada Tabel 9.

Tabel 9. *Confusion Matrix* Kinerja Terbaik Algoritma k-NN

Akurasi = 87.90%	Prediksi		Kelas <i>Recall</i>
	Positif	Negatif	
Positif	11.457	38.413	0.230
Aktual Negatif	13.194	362.131	0.965
Kelas Presisi	0.465	0.904	

Tabel 9 menunjukkan hasil evaluasi *Confusion matrix* dari kinerja terbaik algoritma k-NN menggunakan *tools* Orange pada mode pengujian validasi silang. Dari 425.195 data, 11.457 data diprediksi dengan benar sebagai kelas positif dan 38.413 data kelas positif diprediksi sebagai kelas negatif. Serta 362.131 data diprediksi benar sebagai kelas negatif dan 13.194 data kelas negatif diprediksi sebagai kelas positif.

Tabel 10. *Confusion Matrix* Kinerja Terbaik Algoritma SGD

Akurasi = 89.46%	Prediksi		Kelas <i>Recall</i>
	Positif	Negatif	
Positif	13.121	36.749	0.263
Aktual Negatif	8.051	367.274	0.979
Kelas Presisi	0.620	0.909	

Tabel 10 menunjukkan hasil evaluasi *Confusion matrix* dari kinerja terbaik algoritma SGD menggunakan *tools* KNIME pada mode pengujian validasi silang. Dari 425.195 data, 13.121 data diprediksi dengan benar sebagai kelas positif dan 36.749 data kelas positif diprediksi sebagai kelas negatif. Serta 367.274 data diprediksi benar sebagai kelas negatif dan 8.051 data kelas negatif diprediksi sebagai kelas positif.

Tabel 11. *Confusion Matrix* Kinerja Terbaik Algoritma RF

Akurasi = 89.56%	Prediksi		Kelas <i>Recall</i>
	Positif	Negatif	
Positif	11.912	37.958	0.239
Aktual Negatif	6.429	368.896	0.983
Kelas Presisi	0.649	0.907	

Tabel 11 menunjukkan hasil evaluasi *Confusion matrix* dari kinerja terbaik algoritma RF menggunakan *tools* KNIME pada mode pengujian validasi silang. Dari 425.195 data, 11.912 data diprediksi dengan benar sebagai kelas positif dan 37.958 data kelas positif diprediksi sebagai kelas negatif. Serta 368.896 data diprediksi benar sebagai kelas negatif dan 6.429 data kelas negatif diprediksi sebagai kelas positif. Berikutnya adalah tabel *Confusion matrix* dari algoritma NN seperti pada Tabel 12.

Tabel 12. *Confusion Matrix* Kinerja Terbaik Algoritma NN

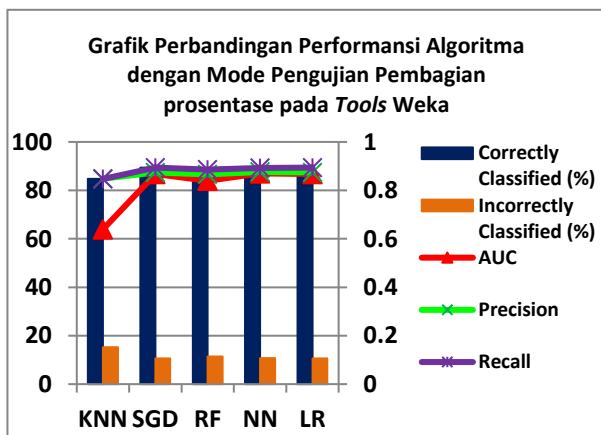
Akurasi = 89.60%	Prediksi		Kelas <i>Recall</i>
	Positif	Negatif	
Positif	13.325	36.545	0.267
Aktual Negatif	7.646	367.679	0.980
Kelas Presisi	0.635	0.910	

Tabel 12 menunjukkan hasil evaluasi *Confusion matrix* dari kinerja terbaik algoritma NN menggunakan *tools* Orange pada mode pengujian validasi silang. Dari 425.195 data, 13.325 data diprediksi dengan benar sebagai kelas positif dan 36.545 data kelas positif diprediksi sebagai kelas negatif. Serta 367.679 data diprediksi benar sebagai kelas negatif dan 7.646 data kelas negatif diprediksi sebagai kelas positif.

Tabel 13. *Confusion Matrix* Kinerja Terbaik Algoritma LR

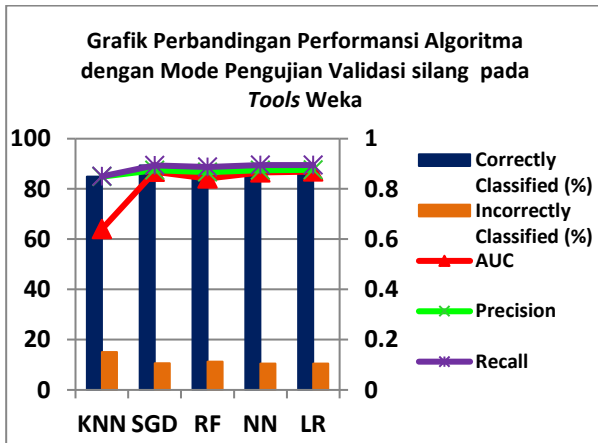
akurasi = 89.50%	Prediksi		Kelas <i>Recall</i>
	Positif	Negatif	
Positif	13.455	36.415	0.270
Aktual Negatif	8.380	366.945	0.978
Kelas Presisi	0.616	0.910	

Tabel 13 merupakan hasil evaluasi *Confusion matrix* dari kinerja terbaik algoritma LR menggunakan *tools* Orange pada mode pengujian validasi silang. Dari 425.195 data, 13.455 data diprediksi dengan benar sebagai kelas positif dan 36.415 data kelas positif diprediksi sebagai kelas negatif. Serta 366.945 data diprediksi benar sebagai kelas negatif dan 8.380 data kelas negatif diprediksi sebagai kelas positif.

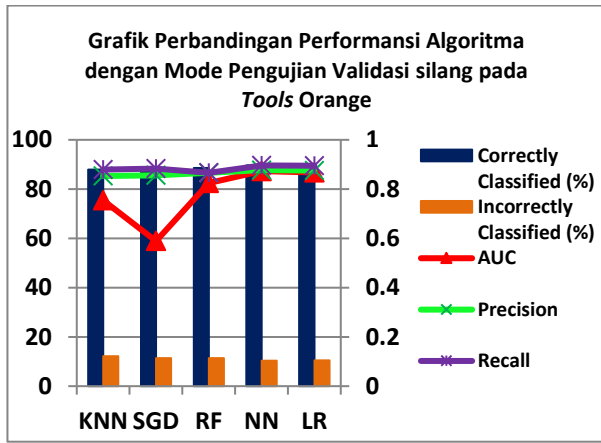


Gambar 3. Grafik Perbandingan Performansi Klasifikasi pada Mode Pengujian Pembagian prosentase menggunakan *Tools* Weka

Gambar 3 merupakan grafik perbandingan performansi dengan mode pengujian pembagian prosentase menggunakan *tools* Weka yang menunjukkan bahwa algoritma SGD dan LR memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma k-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.



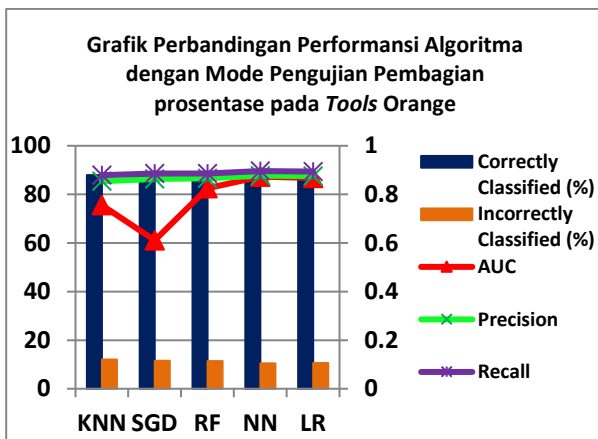
Gambar 4. Grafik Perbandingan Performansi Klasifikasi pada Mode Pengujian Validasi silang menggunakan Tools Weka



Gambar 6. Grafik Perbandingan Performansi Klasifikasi pada Mode Pengujian Validasi silang menggunakan Tools Orange

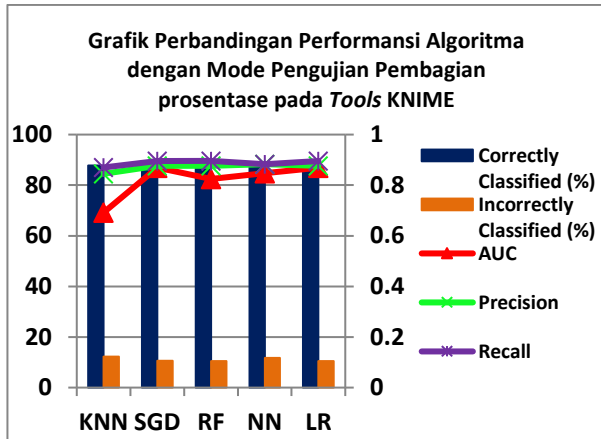
Gambar 4 merupakan grafik perbandingan performansi dengan mode pengujian validasi silang menggunakan tools Weka yang menunjukkan bahwa algoritma LR memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma k-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.

Gambar 6 merupakan grafik perbandingan performansi dengan mode pengujian validasi silang menggunakan tools Orange yang menunjukkan bahwa algoritma NN memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma K-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.



Gambar 5. Grafik Perbandingan Performansi klasifikasi pada Mode Pengujian Pembagian prosentase menggunakan Tools Orange

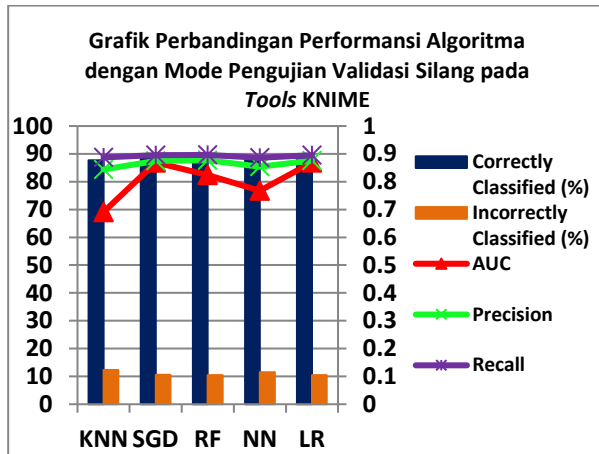
Gambar 5 merupakan grafik perbandingan performansi dengan mode pengujian pembagian prosentase menggunakan tools Orange yang menunjukkan bahwa algoritma NN memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma k-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.



Gambar 7. Grafik Perbandingan Performansi Klasifikasi pada Mode Pengujian Pembagian prosentase menggunakan Tools KNIME

Gambar 7 merupakan grafik perbandingan performansi mode pengujian pembagian prosentase menggunakan tools KNIME yang menunjukkan bahwa algoritma RF memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma k-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.





Gambar 8. Grafik Perbandingan Performansi Klasifikasi pada Mode Pengujian Validasi silang menggunakan Tools KNIME

Gambar 8 merupakan grafik perbandingan performansi dengan mode pengujian validasi silang menggunakan tools KNIME yang menunjukkan bahwa algoritma RF memiliki grafik performansi tertinggi disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih tinggi daripada algoritma yang lain. Sedangkan grafik performansi terendah ditunjukkan oleh algoritma k-NN disebabkan nilai akurasi dan *performance metrics* yang dihasilkan lebih rendah daripada algoritma yang lain.

#### 4. Kesimpulan

Berdasarkan hasil evaluasi perbandingan kinerja dari tiga algoritma yaitu *k-nearest neighbors*, *stochastic gradient descent*, *random forest*, *neural network* dan *logistic regression* dalam mengklasifikasikan dataset hasil pemeriksaan penyakit kardiovaskular menunjukkan bahwa algoritma NN menduduki posisi pertama dengan kinerja klasifikasi terbaik yaitu nilai akurasi prediksi sebesar 89.60%, nilai AUC sebesar 0.873, nilai presisi sebesar 0.877, dan nilai *recall* sebesar 0.896 pada mode pengujian pembagian prosentase dan validasi silang menggunakan tools Orange. Sedangkan posisi kedua diraih oleh algoritma RF dengan nilai akurasi prediksi sebesar 89.56%, nilai AUC sebesar 0.825, nilai presisi sebesar 0.877, dan nilai *recall* sebesar 0.896 pada mode pengujian pembagian prosentase menggunakan tools KNIME. Dan posisi ketiga diraih oleh algoritma LR dengan nilai akurasi prediksi sebesar 89.50%, nilai AUC sebesar 0.868, nilai presisi sebesar 0.875, dan nilai *recall* sebesar 0.895 pada mode pengujian validasi silang menggunakan tools Orange. Serta pada posisi keempat diraih oleh algoritma SGD dan posisi terakhir diraih oleh algoritma K-NN.

Berdasarkan hasil nilai AUC dari masing-masing algoritma menunjukkan rata-rata nilai AUC dalam rentang 0.80-0.90 sehingga termasuk didalam kategori *Good Classification*. Untuk penelitian kedepannya dapat dilakukan eksperimen menggunakan dataset yang berbeda sehingga dapat dijadikan sebagai bahan perbandingan tingkat kinerja algoritma klasifikasi.

#### Daftar Rujukan

- [1] Team, R. P., 2017. *Buku Saku Pedoman Pengkajian dan Pengelolaan Risiko Penyakit Kardiovaskuler Versi Bahasa Indonesia*. Yogyakarta: Universitas Gajah Mada.
- [2] Kemenkes RI, 2014. *Situasi kesehatan jantung*. Jakarta: INFODATIN Pusat Data dan Informasi Kementerian Kesehatan RI.
- [3] Perhimpunan Dokter Spesialis Kardiovaskular Indonesia (PERKI), 2019. *Press Release, World Heart Day PERKI 2019* [Online] (Updated 26 Sept 2019) Tersedia di: [http://www.inaheart.org/news\\_and\\_events/news/2019/9/26/press\\_release\\_world\\_heart\\_day\\_perki\\_2019](http://www.inaheart.org/news_and_events/news/2019/9/26/press_release_world_heart_day_perki_2019). [Accessed: 01-Jul-2020].
- [4] Laderas, T., Vasilevsky, N., Pederson, B., Haendel, M., McWeeney, S. and Dorr, D.A., 2018. Teaching data science fundamentals through realistic synthetic clinical cardiovascular data. *bioRxiv*, p.232611.
- [5] SangeethaLakshmi, M.G. and Jayashree, M.M., 2019. Comparative Analysis of Various Tools for Data Mining and Big Data Mining. *International Research Journal of Engineering and Technology (IRJET)*, 6(4), pp.704–708.
- [6] Saberioon, M., Cisař, P., Labbė, L., Souček, P., Pelissier, P. and Kerneis, T., 2018. Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (*oncorhynchus mykiss*) classification using image-based features. *Sensors*, 18(4), p.1027.
- [7] Ruder, S., 2016. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- [8] Thanh Noi, P. and Kappas, M., 2018. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, 18(1), p.18.
- [9] Dewi, S., 2016. Komparasi 5 metode algoritma klasifikasi data mining pada prediksi keberhasilan pemasaran produk layanan perbankan. *Jurnal Techno Nusa Mandiri*, 13(1), pp.60-65.
- [10] Gupta, C., Sharma, S.G. and Bansal, M.G., 2007. *Implementation of Back Propagation Algorithm (of neural networks) in VHDL* Doctoral dissertation: THAPAR INSTITUTE.
- [11] Saritas, M.M. and Yasar, A., 2019. Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*, 7(2), pp.88-91.
- [12] Nugroho, K.S., 2019. *Confusion Matrix untuk Evaluasi Model pada Supervised Learning* [Online] (Updated 13 Nov 2019) Tersedia di: <https://medium.com/@ksnugroho/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f>. [Accessed: 01-Jul-2020]
- [13] Rosandy, T., 2017. Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4. 5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus: Kspps/Bmt Al-fadhila. *Jurnal Teknologi Informasi Magister*, 2(01), pp.52-62.