

# ALGORITME *STACKING* UNTUK KLASIFIKASI PENYAKIT JANTUNG PADA DATASET *IMBALANCED CLASS*

Atik Nurmasani<sup>1</sup>, Yoga Pristyanto<sup>2</sup>

<sup>1,2</sup> Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta  
Jl. Ring Road Utara, Condong Catur Sleman, Yogyakarta (telp: 0274-884201-207; 0274-884208)

<sup>1</sup>nurmasani@pertama.ac.id

<sup>2</sup>yoga.pristyanto@amikom.ac.id

**Abstrak:** Berdasarkan data Riset Kesehatan Dasar (Riskesdas) tahun 2018, angka kejadian penyakit jantung dan pembuluh darah semakin meningkat dari tahun ke tahun. Setidaknya, 15 dari 1000 orang, atau sekitar 2.784.064 individu di Indonesia menderita penyakit jantung. Data mining merupakan bidang yang dapat menjadi solusi untuk digunakan sebagai alat deteksi dini penyakit jantung. Pada penelitian yang dilakukan sebelumnya mayoritas menggunakan *single classifier*, hal ini akan menimbulkan sebuah permasalahan baru ketika dalam dataset penyakit terdapat ketidakseimbangan kelas. Keberadaan ketidakseimbangan tersebut dapat menyebabkan kinerja *single classifier* menjadi tidak maksimal. Oleh karena itu pada penelitian ini akan digunakan metode *ensemble* atau *meta learning*. Berdasarkan pengujian yang dilakukan menunjukkan bahwa algoritme *stacking* mampu menghasilkan kinerja dari sisi akurasi TPR, TNR, G-Mean dan AUC yang lebih baik dibandingkan *single classifier* lainnya. Dengan adanya peningkatan nilai tersebut diharapkan penelitian ini mampu menjadi referensi untuk pengembangan berbagai sistem yang mendukung dan memaksimalkan tingkat keberhasilan proses deteksi dini penyakit jantung menggunakan data mining.

**Kata Kunci:** Data Mining, *Stacking*, Penyakit Jantung, *Imbalanced Class*, *Ensemble*, Klasifikasi.

**Abstract:** Based on data from Basic Health Research (Riskesdas) in 2018, the incidence of heart and blood vessel disease has increased from year to year. At least, 15 out of 1000 people or about 2,784,064 individuals in Indonesia suffer from heart disease. Data mining is a field that can be a solution to be used as an early detection tool for heart disease. In previous research, the majority used a *single classifier*, this would create a new problem when the disease dataset contained class imbalances. The existence of this imbalance can cause the performance of the *single classifier* to be not optimal. Therefore, this study will use the *ensemble* or *meta-learning* method. Based on the tests conducted, it shows that the *stacking* algorithm can produce better performance in terms of TPR, TNR, G-Mean, and AUC accuracy than other *single classifiers*. With this increase in value, it is hoped that this research can become a reference for the development of various systems that support and maximize the success rate of the early detection process of heart disease using data mining.

**Keywords:** Data Mining, *Stacking*, Heart Disease, *Imbalanced Class*, *Ensemble*, Classification

## I. PENDAHULUAN

*Cardiovascular Diseases* (CVD) merupakan salah satu penyakit mematikan nomor satu didunia. Berdasarkan laporan *World Health Organization* (WHO) angka kematian yang disebabkan oleh penyakit *Cardiovascular Diseases* (CVD) mencapai 17,7 juta orang setiap tahunnya dan 31% merupakan penyebab dari seluruh kematian global. Angka kematian akibat penyakit *Cardiovascular Diseases* diprediksi akan terus meningkat dari tahun ke tahun dan diperkirakan pada tahun 2030 akan mencapai 23,3 juta kematian [1] [2]. Salah satu penyakit *Cardiovascular Diseases* ialah

penyakit jantung. Berdasarkan data Riset Kesehatan Dasar (Riskesdas) tahun 2018, angka kejadian penyakit jantung dan pembuluh darah semakin meningkat dari tahun ke tahun. Setidaknya, 15 dari 1000 orang, atau sekitar 2.784.064 individu di Indonesia menderita penyakit jantung [3].

Pada dasarnya penyakit jantung dapat dicegah dengan berbagai faktor, diantaranya pola hidup sehat. Selain itu deteksi dini penyakit jantung juga diperlukan untuk mencegah terjadinya kematian pada penderitanya. Salah satu cara untuk melakukan deteksi dini ialah menggunakan teknologi informasi [4]. Data mining merupakan bidang yang dapat menjadi solusi untuk digunakan sebagai alat deteksi dini penyakit jantung. Data mining dapat diartikan sebagai berbagai macam cabang ilmu pengetahuan yang menjadi satu, terdiri atas sistem basis data, statistika, machine learning, *visualization*, dan informasi pengetahuan. Data mining telah berhasil diterapkan diberbagai bidang ilmu seperti ekonomi, bioinformatika, genetika, kedokteran, pendidikan dan lain sebagainya [5].

Pada data mining terdapat berbagai teknik antara lain klustering, asosiasi, prediksi dan klasifikasi [6]. Pada kasus deteksi dini penyakit jantung mayoritas teknik yang digunakan ialah klasifikasi. Teknik klasifikasi bekerja dengan mengenali pola atau model dari sebuah dataset khususnya dataset penyakit jantung yang sudah banyak tersedia di *repository*. Tujuannya agar model tersebut dapat digunakan untuk memprediksi ataupun klasifikasi apakah seseorang tersebut menderita penyakit jantung atau tidak. Model tersebut didasarkan pada analisis data latih. Model dari hasil klasifikasi dapat dimanfaatkan untuk mengklasifikasikan serta memprediksi tren

data masa depan [7]. Beberapa algoritme klasifikasi yang umum digunakan pada deteksi dini penyakit jantung menggunakan data mining antara lain *K-Nearest Neighbour* (kNN), *Tree C4.5*, *Support Vector Machine* (SVM) dan *Neural Network* (NN) [8]. Pada penelitian yang dilakukan sebelumnya mayoritas menggunakan *single classifier*, hal ini akan menimbulkan sebuah permasalahan baru ketika dalam dataset penyakit terdapat ketidakseimbangan kelas. Keberadaan ketidakseimbangan tersebut dapat menyebabkan kinerja *single classifier* menjadi tidak maksimal. Oleh karena itu pada penelitian ini akan digunakan metode *ensemble* atau *meta learning*. Metode *ensemble* merupakan gabungan dari beberapa algoritme yang dijadikan satu sehingga kinerjanya lebih baik dibandingkan *single classifier*. Algoritme ensemble yang digunakan pada penelitian ini ialah *Stacking*. Tujuannya ialah untuk meningkatkan kinerja *single classifier*. Dengan adanya peningkatan kinerja klasifikasi diharapkan dapat digunakan sebagai referensi, terutama dalam rangka pengembangan berbagai sistem yang dapat meningkatkan keberhasilan deteksi dini penyakit jantung menggunakan teknologi informasi.

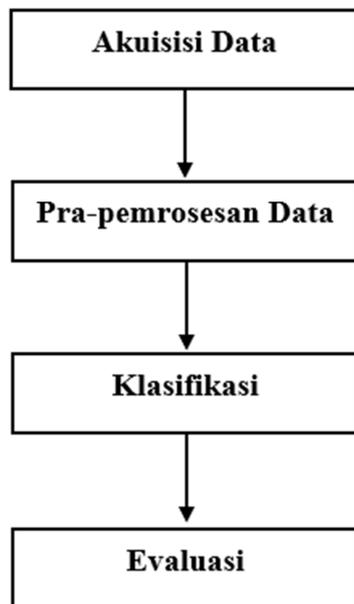
## II. TINJAUAN PUSTAKA

Berbagai penelitian telah dilakukan terkait deteksi dini penyakit jantung menggunakan data mining. Penelitian yang dilakukan oleh [9] menggunakan algoritme *Support Vector Machine* (SVM) dengan kernel linier menghasilkan akurasi sebesar 77% pada kasus klasifikasi penyakit jantung. Penelitian serupa juga dilakukan oleh [10] menggunakan algoritme SVM, *Decision Tree*, dan *Neural Network*. Masing-masing algoritme menghasilkan akurasi sebesar SVM 69%, *Decision*

*Tree* 70% dan *Neural Network* sebesar 69%. Kedua penelitian tersebut menggunakan dataset sama yaitu *heart failure clinical records data set* yang bersumber di *UCI Machine Learning Repository* [11]. Namun kondisi kelas yang tidak seimbang yang terdapat dalam dataset tersebut berpengaruh terhadap akurasi pada dua penelitian tersebut masih dibawah 80%. Oleh karena itu pada penelitian ini akan menggunakan metode *ensemble stacking* untuk mengatasi permasalahan tersebut dan meningkatkan kinerja model klasifikasi penyakit jantung.

### III. METODE PENELITIAN

Penelitian ini dilakukan berdasarkan tahapan penelitian yang meliputi akuisisi data, pra-pemrosesan data, klasifikasi, serta evaluasi sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Tahap pertama pada penelitian ini ialah proses akuisisi data. Pada penelitian dataset yang digunakan merupakan dataset publik yang termasuk kategori *benchmarking*. Dataset

bersumber dari *repository UCI Machine Learning*. Dataset yang digunakan ialah *heart failure clinical records data set* [11]. *Heart failure clinical* merupakan dataset utama yang digunakan dalam penelitian ini. Selain itu dataset ini terdiri dari 13 *features*, 299 *instances* dan 2 *class/label*. Jumlah instance minoritas sebesar 96 *instances* sedangkan jumlah kelas mayoritas sebesar 203 *instances*. Berikut Tabel 1 merupakan detail fitur pada dataset *heart failure clinical*.

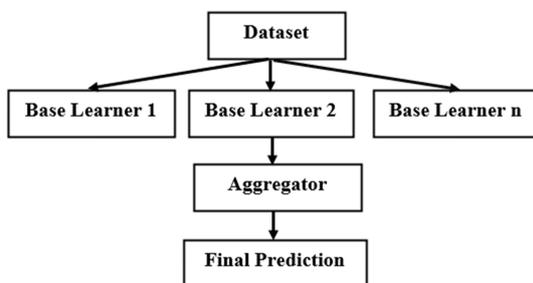
Tabel 1. Detail Fitur Dataset Heart Failure Clinical

Variabel	Tipe
Age	Numeric
Anemia	Categorical
High blood pressure	Categorical
Creatine phosphokinase (CPK)	Numeric
Diabetes	Categorical
Ejection Fraction	Numeric
Platelets	Numeric
Sex	Categorical
Serum creatine	Numeric
Serum sodium	Numeric
Smoking	Categorical
Time	Numeric
Death event (target/class/label)	Categorical

Tahap selanjutnya setelah akuisisi data ialah melakukan pra-pemrosesan terhadap dataset yang digunakan. Tahap pra-pemrosesan data pada umumnya meliputi beberapa hal anantara lain pengisian data yang kosong, menghilangkan duplikasi data, dan memeriksa inkonsistensi data. Biasanya data yang kosong disebabkan oleh kesalahan alat pada saat pengambilan data maupun adanya data baru yang belum ada informasinya. Pada tahap pra-pemrosesan kami melakukan pengisian missing *values* pada dataset. Keberadaan missing values pada dataset tentunya juga akan berpengaruh pada hasil klasifikasi dataset itu sendiri. Oleh karena itu pada tahap ini kami melakukan pengisian missing values menggunakan suatu nilai konstan. Pada data numerik missing values diganti dengan nilai rata-rata, sedangkan

untuk data kategorik maka *missing values* diganti dengan nilai modus.

Setelah tahap pra-pemrosesan data, tahap berikutnya ialah proses klasifikasi menggunakan algoritme *ensemble* atau *meta learning*. Algoritme *ensemble* yang dipilih ialah *Stacking Algoritm*. Tujuan utama menggunakan algoritme Stacking ialah sebagai solusi untuk mengatasi keberadaan ketidakseimbangan kelas yang ada pada dataset yang digunakan. *Stacking* merupakan salah satu algoritme ensemble, pada Gambar 3 menunjukkan algoritma stacking. N himpunan bagian yang berbeda dari kumpulan data pelatihan dibuat dengan menggunakan pengambilan sampel bertingkat dengan penggantian di mana proporsi relatif dari kelas yang berbeda dipertahankan di semua himpunan bagian. Setiap subset dari set pelatihan digunakan untuk menentukan kinerja pengklasifikasi pada set pelatihan. *Meta-classifier* dalam bentuk bobot relatif untuk masing-masing *classifier* dibuat dengan menetapkan bobot ke *classifier* yang sebanding dengan kinerjanya. *Meta classifier* dapat digambarkan dalam berbagai tahap dalam skenario *meta-learning* yang disederhanakan [12]. Berikut Gambar 2 merupakan gambaran proses kerja algoritme *stacking*.



Gambar 2. Gambaran Proses Stacking

Pada penelitian ini kami menggunakan algoritme stacking dengan *base learner* C4.5 dan *Support Vector Machine*, dengan asumsi bahwa

masing-masing *single classifier* tersebut tidak cukup baik dalam kasus *imbalanced class classification*. Sedangkan untuk aggregatornya menggunakan *default* dari *stacking* yaitu *Logistic Regression*. Pada penelitian ini proses pengujian menggunakan alat bantu python 3 dengan library *sklearn* dan *jupyter notebook*. Berikut Gambar 3 merupakan *source code* metode *stacking* yang diujikan.

```

def get_stacking():
    # basis model
    level0 = list()
    level0.append(('tree', DecisionTreeClassifier()))
    level0.append(('svm', SVC()))
    # Meta learning
    level1 = LogisticRegression()
    # Stacking
    model = StackingClassifier(estimators=level0, final_estimator=level1)
    return model
    
```

Gambar 3. Source Code Metode Stacking

Tahap terakhir pada penelitian ini ialah melakukan evaluasi kinerja dari algoritme stacking kemudian hasilnya akan dibandingkan dengan *single classifier* lainnya. Hasil dari sebuah proses klasifikasi data dapat dikategorikan menjadi empat jenis yaitu:

- *True Positive* (TP) adalah jumlah *record* positif yang diklasifikasikan sebagai positif.
- *False positif* (FP) adalah jumlah *record* negatif yang diklasifikasikan sebagai positif.
- *False negatif* (FN) adalah jumlah *record* positif yang diklasifikasikan sebagai negatif,
- *True negatif* (TN) adalah jumlah *record* negatif yang diklasifikasikan sebagai *negatif*.

Hasil tersebut didapatkan dari *confusion matrix* berikut Tabel 2 merupakan Tabel *confusion matrix* [13].

Tabel 2. Confusion Matrix

Aktual	Prediksi	
	True	False
True	TP	FN
False	FP	TN

Indikator evaluasi yang digunakan antara lain akurasi, *true positive rate* (TPR), *true negative rate* (TNR), *geometric mean* (G-Mean) dan *area under curve* (AUC) [14]. Berikut persamaan untuk menghitung indikator tersebut.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$TPR = \frac{TP}{TP+FN} \quad (2)$$

$$TNR = \frac{TN}{TN+FP} \quad (3)$$

$$geometric\ mean = \sqrt{TNR \times TPR} \quad (4)$$

#### IV. HASIL

Pada penelitian ini kami membagi masing-masing dataset ke dalam dua bagian yaitu sebagai data training dan data testing. Skema pembagian data menggunakan skema, 80% untuk data latih model dan 20% digunakan untuk validasi model atau *testing* model. Pada penelitian ini algoritme yang diusulkan yaitu algoritme *stacking* akan dibandingkan dengan *single classifier* antara lain K-NN, SVM, Tree C4.5 dan *Neural Network*. Berikut tabel 3 merupakan hasil klasifikasi yang berupa tabel *confusion matrix* masing-masing algoritme.

Tabel 3. *Confusion Matrix* Masing-Masing Algoritme

	TP	TN	FP	FN
K-NN	17	169	34	79
C4.5	61	172	31	35
SVM	52	185	18	44
Neural Network	49	189	14	47
Stacking	60	182	21	36

Kemudian nilai-nilai pada tabel *confusion matrix* tersebut akan digunakan untuk menghitung nilai evaluasi. Berikut Tabel 4 merupakan nilai hasil evaluasi algoritme menggunakan indikator yang sudah ditentukan.

Tabel 4. *Confusion Matrix* Masing-Masing Algoritme

	K-NN	C4.5	SVM	NN	Stacking
Akurasi	62%	78%	79%	80%	81%
TPR	18%	64%	54%	51%	63%
TNR	83%	85%	91%	93%	90%
G-Mean	38%	73%	70%	69%	75%
AUC	45%	73%	84%	84%	87%

Tabel 4 menunjukkan bahwa metode atau algoritme yang diusulkan yaitu *Stacking* antara *Tree* C4.5 dan SVM mampu menghasilkan nilai akurasi, TPR, TNR, G-Mean dan AUC yang lebih baik dibandingkan *single classifier* lainnya. Hal ini dikarenakan oleh keberadaan *imbalanced class* pada dataset. Selain itu dapat juga dilihat pada Tabel 3 menunjukkan bahwa nilai total *record false* atau kesalahan klasifikasi yang dihasilkan oleh algoritme *stacking* juga paling sedikit dibandingkan *single classifier*. Oleh karena itu metode atau algoritme yang diusulkan dapat menjadi solusi untuk dataset yang memiliki kondisi *imbalanced class*.

#### V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan menunjukkan bahwa data mining dapat digunakan sebagai solusi untuk deteksi dini penyakit jantung. Akan tetapi keberadaan *imbalanced class* pada dataset penyakit jantung menyebabkan kinerja *single classifier* menjadi tidak maksimal. Oleh karena algoritme yang diusulkan yaitu *stacking* dapat menjadi solusi untuk permasalahan *imbalanced class*. Berdasarkan pengujian yang dilakukan menunjukkan bahwa algoritme *stacking* mampu menghasilkan kinerja dari sisi akurasi TPR, TNR, G-Mean dan AUC yang lebih baik dibandingkan *single classifier* lainnya. Dengan adanya peningkatan nilai tersebut diharapkan penelitian ini mampu menjadi referensi untuk

pengembangan berbagai sistem yang mendukung dan memaksimalkan tingkat keberhasilan proses deteksi dini penyakit jantung menggunakan data mining.

#### UCAPAN TERIMA KASIH

Terima kasih kami ucapkan kepada program studi S1 Sistem Informasi dan Lembaga Penelitian Universitas Amikom Yogyakarta yang telah memberikan dukungan penuh dalam penelitian ini.

#### REFERENSI

- [1] M. N, R. A, R. S., and S. Stang, "Karakteristik Dan Prevalensi Risiko Penyakit Kardiovaskular Pada Tukang Masak Warung Makan Di Wilayah Kerja Puskesmas Tamalanrea," *J. Kesehat.*, vol. 11, no. 1, pp. 30–38, 2018, doi: 10.24252/kesehatan.v11i1.5029.
- [2] Kemenkes RI, *KMK No. 854 ttg Cardiovascular Diseases Guideline.pdf*. 2009, p. 32.
- [3] I. Ayu, E. Widiastuti, R. Cholidah, G. W. Buanayuda, and I. B. Alit, "Deteksi Dini Faktor Risiko Penyakit Kardiovaskuler pada Pegawai Rektorat Universitas Mataram," *J. Pengabd. Magister Pendidik. IPA*, vol. 4, pp. 137–142, 2021.
- [4] A. B. Wibisono and A. Fahrurrozi, "Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner," *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 3, pp. 161–170, 2019, doi: 10.35760/tr.2019.v24i3.2393.
- [5] Fadhilah Ahmad, N. H. Ismail, and Azwa Abdul Aziz, "The prediction of students' academic performance using classification data mining techniques," *Appl. Math. Sci.*, vol. 9, no. 129, pp. 6415–6426, 2015, doi: 10.12988/ams.2015.53289.
- [6] Y. Pristyanto, "Penerapan Metode Ensemble Untuk Meningkatkan Kinerja Algoritme Klasifikasi Pada Imbalanced Dataset," *J. TEKNOINFO*, vol. 13, no. 1, pp. 11–16, 2019, doi: 10.33365/jti.
- [7] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
- [8] V. V. Ramalingam, A. Dandapath, and M. Karthik Raja, "Heart disease prediction using machine learning techniques: A survey," *Int. J. Eng. Technol.*, vol. 7, no. 2.8 Special Issue 8, pp. 684–687, 2018, doi: 10.14419/ijet.v7i2.8.10557.
- [9] P. S. Kohli and A. L. Regression, "Application of Machine Learning in Disease Prediction," in *2020 IEEE 5th International Conference on Computing Communication and Automation, ICCCA 2020*, 2020, pp. 1–4.
- [10] C. S. Wu, M. Badshah, and V. Bhagwat, "Heart disease prediction using data mining techniques," in *ACM International Conference Proceeding Series*, 2019, pp. 7–11, doi: 10.1145/3352411.3352413.
- [11] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. 1, pp. 1–16, 2020, doi: 10.1186/s12911-020-1023-5.
- [12] N. Chanamarn, K. Tamee, and P. Sittidech, "Stacking technique for academic achievement prediction," *Int. Work. Smart Info-Media Syst. Asia (SISA 2016)*, no. Sisa 2016, pp. 14–17, 2016.
- [13] J. Han, M. Kamber, and J. Pei, *Data Mining : Concept and Techniques*, Third Edit. Massachusetts: Morgan Kauffman, 2011.
- [14] Q. Wang, "A hybrid sampling SVM approach to imbalanced data classification," *Abstr. Appl. Anal.*, vol. 2014, 2014, doi: 10.1155/2014/972786.