

Sistem Prediksi Penyakit Jantung Berbasis Web Menggunakan Metode SVM dan Framework Streamlit

Oleh:

Ary Putranto,

Nuril Lutvi Azizah

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Mei, 2023



Pendahuluan

- Berdasarkan laporan World Health Organization (WHO) tahun 2005, dari 58 juta kematian di dunia, 17,5 juta (30%) diantaranya disebabkan Oleh penyakit jantung dan pembuluh darah, terutama Oleh serangan Jantung (716 juta) dan stroke (57 juta).
- Di Indonesia, Angka kematian penyakit tidak menular terutama penyakit jantung dan pembuluh darah meningkat dari 41,7% pada tahun 1995 menjadi 59,5% pada tahun 2007 (Riskesdas 2007).
- Penderita penyakit jantung dan pembuluh darah tidak lagi mengenal kelompok status sosial ekonomi masyarakat. Tidak sedikit penderita penyakit jantung dan pembuluh darah yang justru datang dari kalangan ekonomi menengah kebawah yang tergolong masyarakat miskin, tidak mampu dan kurang mampu. Hal tersebut kemungkinan diakibatkan oleh perubahan gaya hidup yang tidak sehat serta meningkatnya faktor risiko penyakit jantung dan pembuluh darah.

Pertanyaan Penelitian (Rumusan Masalah)

Permasalahan yang menjadi topik utama penelitian ini adalah sebagai berikut.

1. Bagaimana membangun sebuah model machine learning untuk melakukan prediksi penyakit Jantung?
2. Bagaimana melakukan deployment model ke dalam aplikasi berbasis web?

Metode

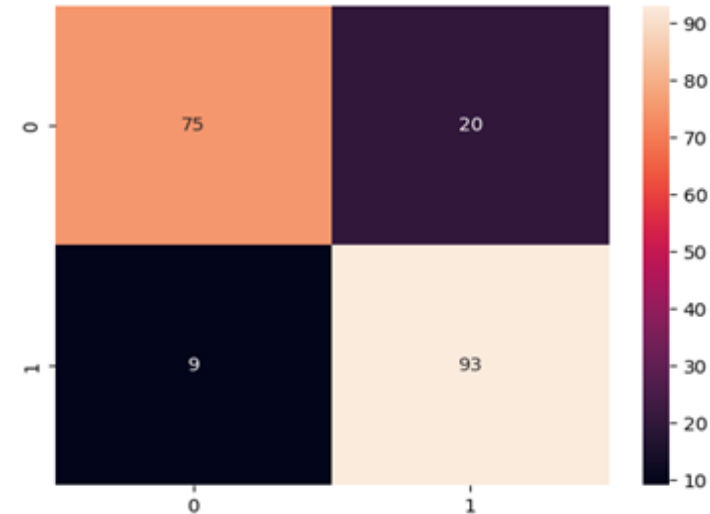
Algoritma yang digunakan adalah SVM

Algoritma Support Vector Machine merupakan salah satu algoritma yang termasuk dalam kategori Supervised Learning, yang artinya data yang digunakan untuk belajar oleh mesin merupakan data yang telah memiliki label sebelumnya. Sehingga dalam proses penentuan keputusan, mesin akan mengkategorikan data testing ke dalam label yang sesuai dengan karakteristik yang dimilikinya.

Cara kerja dari metode Support Vector Machine khususnya pada masalah non-linear adalah dengan memasukkan konsep kernel ke dalam ruang berdimensi tinggi. Tujuannya adalah untuk mencari hyperplane atau pemisah yang dapat memaksimalkan jarak (margin) antar kelas data. Untuk menemukan hyperplane terbaik, kita dapat mengukur margin kemudian mencari titik maksimalnya. Proses pencarian hyperplane yang terbaik ini adalah inti dari metode Support Vector Machine ini.

Hasil

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) Type I Error
	0 (Negative)	FN (False Negative) Type II Error	TN (True Negative)



Precision= $TP/(TP+FP)$

Precision= $75/(75+20) \times 100\%$

Precision= $75/95 \times 100\%$

Precision= $0,78 \times 100\%$

Precision= 78%

Recall= $TP/(TP+FN)$

Recall= $75/(75+9) \times 100\%$

Precision= $75/84 \times 100\%$

Precision= $0,89 \times 100\%$

Precision= 89%

- Akurasi= $(TP+TN)/(TP+FN+FP+TN) \times 100\%$
- Akurasi= $(75+93)/(75+9+20+93) \times 100\%$
- Akurasi= $168/197 \times 100\%$
- Akurasi= $0,853 \times 100\%$
- Akurasi= 85%

Pembahasan

Terdapat beberapa Langkah untuk membangun sebuah model machine learning yang dapat mendeteksi penyakit jantung. Diantaranya adalah:

- Pengumpulan data
- Preprocessing
- Processing
- deployment

Temuan Penting Penelitian

Data yang digunakan pada penelitian ini merupakan data indicator penyakit jantung yang diambil UCI Machine Learning

Manfaat Penelitian

Dari penelitian ini ada beberapa manfaat yang bisa di ambil adalah sebagai berikut:

- a. Sesorang dapat melakukan pencegahan penyakit jantung lebih dini
- b. Memberikan kontribusi bagi upaya pengembangan aplikasi yang dapat membantu masyarakat medis terutama dokter dan rumah sakit dalam mendiagnosis penyakit jantung dengan tingkat akurasi yang baik.

Referensi

- [1] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 2, 2020, doi: 10.30865/mib.v4i2.2080.
- [2] S. Agustin, "Ciri-Ciri Sakit Jantung yang Harus Diwaspadai," <https://www.alodokter.com/kenali-ciri-ciri-sakit-jantung>, May 08, 2022.
- [3] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 2, p. 437, Apr. 2020, doi: 10.30865/mib.v4i2.2080.
- [4] Samsudiney, "Penjelasan Sederhana tentang Apa Itu SVM?," <https://medium.com/@samsudiney/penjelasan-sederhana-tentang-apa-itu-svm-149fec72bd02>, Jul. 25, 2019.
- [5] A. F. HIDAYATULLAH, "MEMBUAT APLIKASI WEB SAINS DATA DENGAN MUDAH MENGGUNAKAN STREAMLIT," <https://informatics.uui.ac.id/2021/03/15/streamlit-membuat-aplikasi-web-sains-data/>, Mar. 15, 2021.
- [6] M. A. Bianto, K. Kusri, and S. Sudarmawan, "Perancangan Sistem Klasifikasi Penyakit Jantung Menggunakan Naïve Bayes," *Creative Information Technology Journal*, vol. 6, no. 1, 2020, doi: 10.24076/citec.2019v6i1.231.
- [7] D. P. Utomo, P. Sirait, and R. Yunis, "Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi Menggunakan Algoritma C5.0," *Jurnal Media Informatika Budidarma*, vol. 4, no. 4, 2020.
- [8] A. Nurmasani and Y. Pristyanto, "ALGORITME STACKING UNTUK KLASIFIKASI PENYAKIT JANTUNG PADA DATASET IMBALANCED CLASS," 2021. [Online]. Available: www.ejournal.unib.ac.id/index.php/pseudocode
- [9] A. Novita Sari and S. Alfionita, "Klasifikasi Penyakit Jantung Menggunakan Metode Naïve Bayes," vol. 1, no. 1, pp. 22–26, 2022, doi: 10.12487/AMRI.v1i1.xxxxx.
- [10] P. R. Fitrayana and D. R. S. Saputro, "Algoritme Clustering Large Application (CLARA) untuk Menangani Data Outlier," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 5, pp. 721–725, 2022, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [11] P. R. , D. A. Sihombing, S. Suryadiningrat, and Y. P. A. C. Yuda, . "Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya," *Jurnal Ekonomi dan Statistik Indonesia*, pp. 307–316, 2022.
- [12] R. Fikri, "Handling Imbalanced Dataset," <https://medium.com/@rusnandifikri96/handling-imbalanced-dataset-260378b2a21b>, Jul. 23, 2020.
- [13] A. A. Arifiyanti and E. D. Wahyuni, "SMOTE: METODE PENYEIMBANG KELAS PADA KLASIFIKASI DATA MINING," *ISSN 2686-6099*, vol. 11, 2020.
- [14] A. Liani, "Analisis Perbandingan Kernel Algoritma Support Vector Machine dalam Mengklasifikasikan Skripsi Teknik Informatika berdasarkan Abstrak," *JOINS (Journal of Information System)*, vol. 5, no. 2, 2020, doi: 10.33633/joins.v5i2.3715.
- [15] P. M. Kouate, "Machine Learning: GridSearchCV & RandomizedSearchCV," <https://towardsdatascience.com/machine-learning-gridsearchcv-randomizedsearchcv-d36b89231b10>, Sep. 11, 2020.
- [16] M. S. ANGGREANY, "Confusion Matrix," <https://socs.binus.ac.id/2020/11/01/confusion-matrix/>, Nov. 01, 2020.
- [17] R. Maulid, "Kriteria Jenis Teknik Analisis Data dalam Forecasting," <https://www.dqlab.id/kriteria-jenis-teknik-analisis-data-dalam-forecasting>, Jan. 14, 2022.
- [18] A. A. Suryanto, A. Muqtadir, and S. Artikel, "PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI Info Artikel : ABSTRAK," no. 1, p. 11, 2019.
- [19] Y. D. Wijaya and M. W. Astuti, "PENGUJIAN BLACKBOX SISTEM INFORMASI PENILAIAN KINERJA KARYAWAN PT INKA (PERSERO) BERBASIS EQUIVALENCE PARTITIONS," *Jurnal Digital Teknologi Informasi*, vol. 4, no. 1, 2021, doi: 10.32502/digital.v4i1.3163.
- [20] E. Agustin, A. Eviyanti, and N. Lutvi Azizah, "Deteksi Penyakit Epilepsi Melalui Sinyal EEG Menggunakan Metode DWT dan Extreme Gradient Boosting," vol. 7, no. 1, pp. 117–127, 2023, doi: 10.30865/mib.v7i1.5412.

