Predicting Heart Disease Using the XgBoost Algorithm and RandomizedSearch Optimizer

Prediksi Penyakit Jantung dengan Menggunakan Algoritma XgBoost dan RandomizedSearch Optimizer

Reo Sahobby1, Dessyanto Boedi P2, Mangaras Yanu F3

1,2,3 Informatika, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

1\*123170067@student.upnyk.ac.id, 2dess@upnyk.ac.id, 3mangaras.yanu@upnyk.ac.id

\*: Penulis korenspondensi (corresponding author)

|  |  |
| --- | --- |
| **Informasi Artikel** | **Abstract** (menggunakan style abstract) |
| Received: December 2020  Revised: -  Accepted: -  Published: - | Purpose: The purpose of this study was to identify heart disease based on tabular data or table data containing parameters from cardiac record data, and also to implements the XgBoost algorithm to predicting heart disease while reducing overfitting  Design/methodology/approach: Implements the XgBoost algorithm to build machine learning models and the RandomizedSearch Optimizer, then calculate the model's accuracy performance in prediciting heart disease  Findings/result: Machine learning models created with the XgBoost algorithm get 91% accuracy on training data and 83% on testing data. Tests carried out using other datasets get 78% accuracy, and the general model gets 90% accuracy.  Originality/value/state of the art: This research was conducted using the XgBoost algorithm combined with the RandomizedSearch Optimizer as a hyper parameter tuning for machine learning model making. |
|  | **Abstrak** (menggunakan style abstrak) |
| *Keywords:one; two; three*  Kata kunci: satu; dua; tiga | Tujuan: Tujuan penelitian ini adalah untuk mengidentifikasi penyakit jantung berdasarkan data tabular atau data tabel yang berisi parameter dari data rekam jantung, dan juga dilakukan untuk menerapkan algoritma XgBoost untuk mengidentifikasi penyakit jantung sekaligus mengurangi *overfitting*.  Perancangan/metode/pendekatan: Menerapkan algoritma XgBoost untuk pembuatan model *machine learning* dan RandomizedSearch Optimizer, kemudian menghitung performa akurasi model dalam mengidentifikasi penyakit jantung.  Hasil: Model *machine learning* yang dibuat dengan algoritma XgBoost mendapatkan akuasi 91% pada *data training* dan 83% pada *data testing*. Pengujian yang dilakukan menggunakan *dataset* lain mendapatkan akurasi 78%, dan model yang dibuat secara umum mendapatkan akuasi sebesar 90%.  Keaslian/ *state of the art*: Penelitian ini dilakukan menggunakan algoritma XgBoost yang dikombinasikan dengan RandomizedSearch Optimizer sebagai tuning *hyper parameter* untuk pembuatan model *machine learning*. |

# Pendahuluan (Heading 1)

Penulisan setiap paragraf pada jurnal Telematika menggunakan *style* **Normal** untuk memudahkan penulisan. Setiap judul bab menggunakan *style* **Heading 1**, dan seterusnya. Sub-bab terakhir merupakan sub-bab level 3 dengan menggunakan *style* **Heading 3.**

# Metode/Perancangan (Heading 1)

Pada metode dibahas tentang metode, tahapan, maupun model yang digunakan dalam penelitian yang dilakukan.

## Pengumpulan Data (Heading 2)

Pada penulisan metode terdapat tabel, gambar, maupun persamaan yang digunakan dalam penelitian tersebut.

Gambar ditampilkan dengan melakukan rujukan dengan menebalkan huruf terlebih dahulu, seperti diperlihatkan pada **Gambar 1**.



**Gambar 1**. Telematika: Jurnal Informatika dan Teknologi Informasi (menggunakan style caption)

Gambar tidak perlu menggunakan bingkai, berikan alasan penggunaan atau penjelasan terkait gambar setelah gambar ditampilkan. Selain gambar, ada juga tabel yang disajikan seperti diperlihatkan pada **Tabel 1**.

**Tabel 1**. Daftar terbitan jurnal Telematika 2020-2021

|  |  |  |
| --- | --- | --- |
| No. | Volume dan Nomor | Bulan Terbit |
| 1. | Volume 17, No. 1 | April 2020 |
| 2. | Volume 17, No. 2 | Oktober 2020 |
| 3. | Volume 18, No. 1 | Februari 2021 |
| 4. | Volume 18, No. 2 | Juni 2021 |
| 5. | Volume 18, No. 3 | Oktober 2021 |

Judul tabel menggunakan style **Caption**, kepala/kop tabel menggunakan style **HeadingTabel**, dan isi tabel menggunakan style **IsiTabel**.

### Tata Cara Penulisan Persamaan (Heading 3)

Persamaan yang digunakan pada penelitian dituliskan dengan menggunakan *equation* pada perangkat lunak Microsoft Word dan style **persamaan**.

(1)

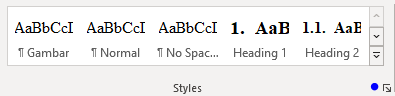
# Hasil dan Pembahasan (Heading 1)

Hasil dan pembahasan memuat hasil penelitian dan pembahasana terkait hasil penelitian tersebut. Setiap gambar tabel yang ditampilkan harus disertai penjelasan agar pembaca bisa memahami isi dari gambar maupun tabel tersebut. Penjelasan terkait data yang disajikan harus disampaikan pada bagian ini dengan tujuan untuk memperjelas kegunaan data pada penelitian.

## Cara menampilkan style

Untuk menampilkan style apa saja yang digunakan pada template ini (bisa dilakukan pada MS Office Word 2010 dan versi setelahnya), bisa perhatikan langkah-langkah berikut ini:

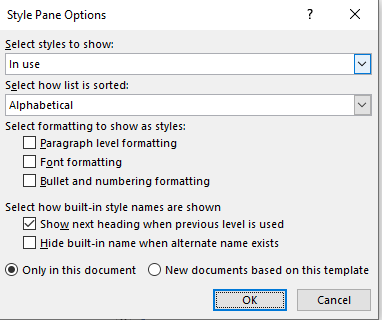
1. Pilih menu **Home** pada toolbar
2. Pada ribbon **Styles**, klik icon arah panah ke pojok kanan bawah seperti diperlihatkan pada **Gambar 2.**



**Gambar 2.** *Ribbon Styles*

Kemudian akan muncul jendela styles pada bagian kanan layar kerja.

1. Pada jendela styles, pilih menu **options**. Kemudian akan muncul jendela menu options.
2. Pada jendela menu options, buatlah pengaturan seperti diperlihatkan pada **Gambar 3**.



**Gambar 3.** *Style Pane*

1. Maka pada jendela styles akan muncul semua jenis style yang digunakan pada template ini.
2. Petunjuk ini berlaku bagi pengguna MS Office Word 2010 – Word 365.
3. Petunjuk nomor 1-7 dibuat menggunakan style **List Paragraph**

## Sitasi dan Referensi

Sitasi dan refrensi yang digunakan mengacu pada **IEEE Style.** Untuk mempermudahpenulisan, disarankan untuk menggunakan *software* pendukung untuk sitasi dan referensi seperti Mendeley, Zotero, EndNote, dll.

Contoh penggunaan sitasi dengan menggunakan IEEE style adalah sebagai berikut.

Contoh sitasi terhadap satu artikel: The feature extraction method applied in the pre-processed image is the Gray Level Co-occurrence Matrix (GLCM) with six parameters (Entropy, Angular Second Moment, Contrast, Inverse Different Moment, Correlation, and Variance) [1].

Contoh sitasi terhadap dua artikel: Pengumpulan data dilakukan dengan menggunakan API untuk kemudian diterapkan pada sebuah perusahaan [2], [3].

Contoh sitasi terhadap lebih dari dua artikel berurutan: Dalam proses pembersihan data dilakukan penghapusan data yang tidak lengkap atau tidak sesuai dengan atribut yang akan digunakan dalam penelitian. Sementara pada tahap transformasi, data diubah menjadi adjacency matrix yang kemudian dianalisis menggunakan SNA [3]–[5].

Contoh sitasi terhadap lebih dari dua artikel yang tidak semuanya berurutan: Sistem informasi skripsi yang saat ini digunakan oleh program studi Teknik Informatika UPN “Veteran” Yogyakarta masih menggunakan cara konvensional, sehingga ditemui kesulitan dalam proses pemeriksaan kemiripan dokumen yang ada dalam proposal tugas akhir mahasiswa. Pemeriksaan kemiripan dokumen dapat dilakukan dengan menggunakan algoritme Winnowing [1], [3]–[5].

## Syarat referensi

Pustaka atau referensi utama harus berasal dari artikel yang diterbitkan pada Jurnal Ilmiah dengan jumlah **minimal 15 referensi utama**. Referensi utama **harus** terbit dalam rentang waktu **10 tahun terakhir** dari pengajuan artikel.

# Kesimpulan dan Saran

Kesimpulan yang ditulis harus bisa menjawab tujuan penelitian. Pada bagian ini juga harus disertai saran untuk penelitian selanjutnya. Kesimpulan dan saran ditulis dalam bentuk paragraf dan bukan dalam bentuk penomoran.

# Daftar Pustaka

[1] S. Saifullah and V. A. Permadi, “Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms,” in *Proceeding - 2019 5th International Conference on Science in Information Technology: Embracing Industry 4.0: Towards Innovation in Cyber Physical System, ICSITech 2019*, 2019, pp. 140–145.

[2] S. P. Tahalea and A. SN, “Central Actor Identification of Crime Group using Semantic Social Network Analysis,” *Indones. J. Inf. Syst.*, vol. 2, no. 1, p. 24, 2019.

[3] Hermantoro, A. P. Suryotomo, A. I. Uktoro, and R. A. Renjani, “Unmanned Aerial Vehicle Application for Plantation Mapping and Automatic Oil Palm Trees Counting on Oil Palm Plantation Management,” in *International Conference on the Role of Agricultural Engineering for Sustainable Agriculture Production*, 2016, no. December, pp. 47–50.

[4] Awang Hendrianto Pratomo, W. Kaswidjanti, and S. Mu’arifah, “Implementasi Algoritma Region of Interest ( ROI ) Untuk Meningkatkan Performa Algoritma Deteksi Dan Klasifikasi Kendaraan,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 1, pp. 155–162, 2020.

[5] S. P. Tahalea, “Identifikasi Peran Hero DOTA2 Menggunakan Social Network Analysis,” *TEKNOMATIKA*, vol. 12, no. 2, pp. 81–86, 2020.