

Model Prediksi Otomatis Jenis Penyakit Hipertensi dengan Pemanfaatan Algoritma Machine Learning Artificial Neural Network

Purwono Purwono^{*1}, Pramesti Dewi², Sony Kartika Wibisono³, Bala Putra Dewa⁴

^{1,3,4}Prodi Informatika, Universitas Harapan Bangsa, Purwokerto

²Prodi Keperawatan, Universitas Harapan Bangsa, Purwokerto

e-mail: ^{*1}purwono@uhb.ac.id, ²pramestidewi@uhb.ac.id, ³wibisono.sony@gmail.com,

⁴bdewa@uhb.ac.id,

Abstrak

Hipertensi merupakan faktor utama dalam perkembangan penyakit seperti stroke, gagal jantung, infark miokard, fibrilasi atrium, penyakit arteri perifer, dan diseksi aorta. Prediksi dini jenis hipertensi dari riwayat kesehatan merupakan hal yang penting agar kita dapat mengetahui penyakit yang disebabkan olehnya. Prediksi ini dapat diperoleh dengan memanfaatkan teknologi machine learning untuk menemukan pengetahuan baru dari data dasar sehingga menemukan pola yang valid, berguna, dan mudah dipelajari. Model klasifikasi neural network diusulkan dalam penelitian ini. Kontribusi kami dalam penelitian ini adalah membuat model klasifikasi neural network. Kami melihat peneliti sebelumnya hanya mengejar nilai akurasi yang tinggi semata. Berbeda dengan penelitian sebelumnya, kami menggunakan teknik optimasi hyperparameter gridsearch cv pada model klasifikasi artificial neural network. Parameter yang digunakan dalam model ini yaitu solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=1. Nilai akurasi ketepatan prediksi dalam menentukan jenis hipertensi ini sebesar 85% yang menunjukkan bahwa model yang dibangun ternyata sudah cukup baik dalam proses klasifikasi.

Kata kunci—klasifikasi, hipertensi, neural network, penyakit, kesehatan

1. PENDAHULUAN

Kesehatan adalah hal yang paling dibutuhkan masyarakat baik secara fisik, mental dan sosial [1]. Tekanan darah merupakan daya yang digunakan dalam proses peredaran darah menuju dinding arteri tubuh manusia [2]. Terdapat dua jenis tekanan darah manusia yaitu tekanan darah sistolik dan tekanan darah diastolik [2]. Tekanan dalam pembuluh darah saat jantung berdetak disebut dengan tekanan darah sistolik sedangkan untuk tekanan diastolik terjadi diantara ketukan denyut jantung. Penyakit tekanan darah tinggi atau bisa kita sebut hipertensi didiagnosis jika nilai tekanan sistolik sama atau lebih besar daripada 140 mmHg [3] serta untuk tekanan diastolik lebih besar atau sama dengan 90 mmHg [4]. Hipertensi menjadi salah satu faktor penting penyebab beberapa penyakit seperti stroke, gagal jantung, infark miokard, atrial fibrilasi, penyakit arteri perifer hingga diseksi aorta [5].

Merujuk pada data yang diambil dari National High Blood Pressure Education Program, terdapat beberapa jenis tekanan darah yang terjadi pada manusia dewasa [2]. Tekanan darah normal berkisar pada angka tekanan sistolik kurang dari 120 mmHg dan diastolik kurang dari 80mmHg. Tekanan disebut dengan *pre-hypertension* jika angka sistolik antara 120 hingga 139 mmHg atau diastolik antara 80 hingga 99 mmHg. *Stage-1 hypertension* jika angka sistolik

berkisar antara 140-159 mmHg atau diastolik antara 90-99 mmHg. *Stage-2 hypertension* jika angka sistolik lebih dari 160 mmHg atau diastolik sama dengan atau lebih dari 100 mmHg. Tahapan *pre-hypertension* biasanya tidak diikuti dengan gejala-gejala yang dirasakan oleh individu, namun berpotensi pada resiko menjadi penyakit hipertensi hingga penyakit kardiovaskule [6].

Kejadian hipertensi secara global mencapai angka 1,13 milyar orang, dimana 31% diantaranya beresiko terhadap orang dewasa dan terus semakin meningkat sebanyak 5,1% dibandingkan prevalensi dunia dalam kurun waktu tahun 2000-2010 [7]. Dilansir dari data Kementerian Kesehatan Indonesia, hipertensi terjadi dalam kategori usia 31-44 tahun dengan persentase 31,6%, pada umur 45-54 tahun sebanyak 45,3 dan 55-64 tahun sebanyak 55,2% [8].

Perkembangan teknologi pada bidang kesehatan terus meningkat [9]. Teknologi seperti pemanfaatan *machine learning* pada aplikasi keperawatan dan medis merupakan salah satu contoh nyata dari perkembangan teknologi pada bidang kesehatan [10]. *Machine learning* ialah teknologi berbasis komputer dan matematika yang menggunakan data sebagai media pembelajaran oleh mesin komputer dan menghasilkan prediksi di masa depan [11]. *Machine learning* hadir sebagai solusi dalam upaya prediksi dini sebuah penyakit [12]. Prediksi diagnostik pada *machine learning* telah terbukti menjadi metode yang menjanjikan dalam berbagai permasalahan kehidupan [13]. Metode *machine learning* dibuat dengan berbagai rangkaian alat dan teknik yang dapat mengeksplorasi kumpulan data dalam jumlah besar hingga menghasilkan pola-pola tertentu yang bermanfaat bagi ilmu pengetahuan [14].

Berbagai penelitian terkait dengan upaya prediksi jenis penyakit hipertensi telah dilakukan oleh para peneliti. Penelitian yang dilakukan oleh [15] memanfaatkan model *machine learning* tanpa teknik *oversampling*. Penelitian tersebut menghasilkan nilai ketepatan prediksi jenis hipertensi sebesar 91%. Penelitian yang dilakukan oleh [16] telah menggunakan model *machine learning* LVQ pada klasifikasi jenis hipertensi dengan pedoman ESH dan berhasil mendapatkan akurasi prediksi pada model tersebut sebesar 94,6%. Penelitian yang telah dilakukan oleh [17] menggunakan *machine learning* dengan optimalisasi fitur yang diekstraksi dari sinyal PPG melalui transformasi *wavelet* dan klinis seperti usia, indeks masa tubuh dan detak jantung dalam memprediksi normotensi dan prehipertensi. Hasil penelitian tersebut mendapatkan nilai akurasi prediksi sebesar 71,42%.

Berdasarkan beberapa penelitian terdahulu, kami secara khusus melakukan penelitian untuk membuat model *machine learning* dengan memanfaatkan algoritma *artificial neural network*. Model selanjutnya diukur nilai akurasi dalam memprediksi jenis hipertensi. Model ini digunakan sebagai sistem prediksi ketika ada data baru yang akan diuji jenis hipertensinya.

2. METODE PENELITIAN

2.1 Dataset

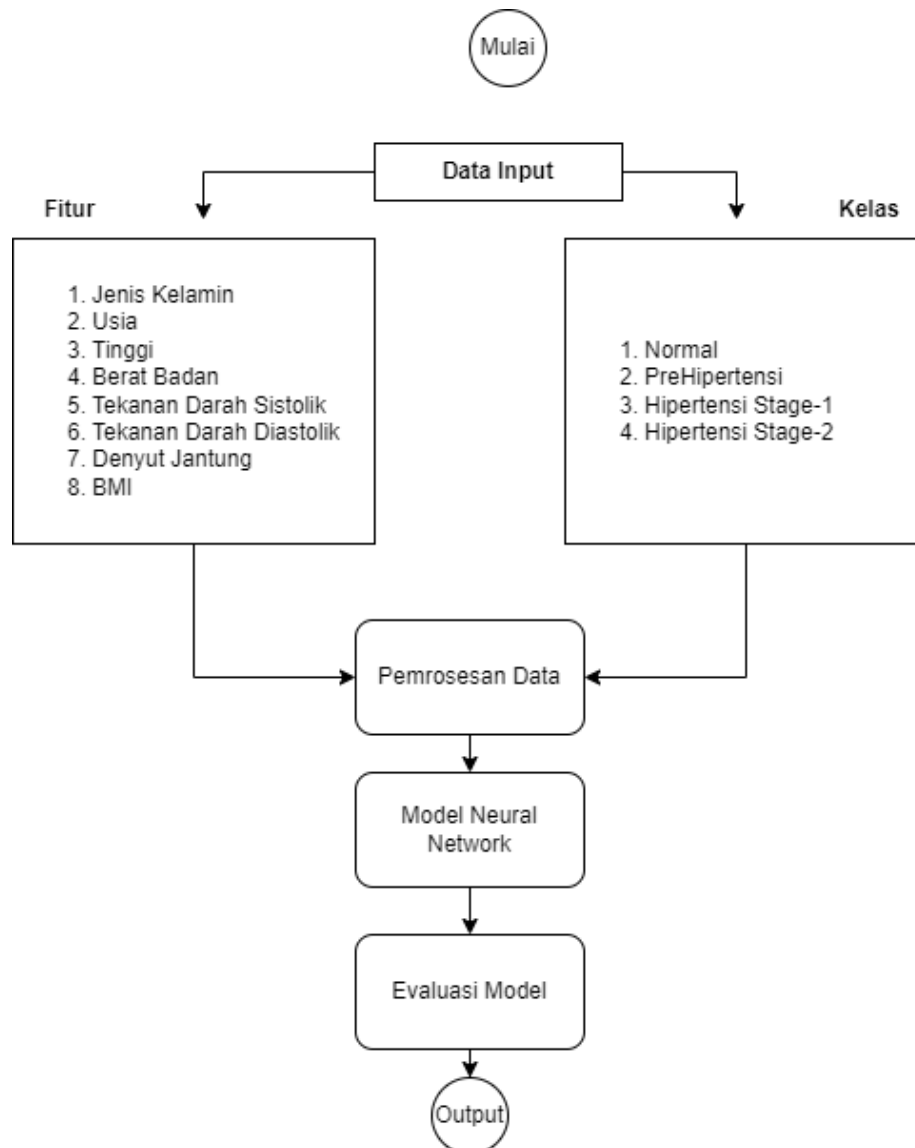
Penelitian ini memanfaatkan dataset yang diperoleh dari PPG-BP (*photoplethysmography-blood pressure*) yang dapat digunakan untuk mengetahui kinerja dari model klasifikasi *machine learning* [18]. Dataset ini memiliki beberapa fitur yang digunakan dalam prediksi hipertensi antara lain adalah jenis kelamin, usia, tinggi (cm), berat badan (kg), tekanan sistolik (mmHg), tekanan diastolik (mmHg), denyut jantung (bpm), dan BMI (kg/m). Kelas utama hipertensi yang terdapat pada dataset tersebut ialah normal (sehat), prehipertensi, hipertensi stadium-1, dan hipertensi stadium-2. Dataset PPG-BP dikumpulkan dari 219 orang dewasa dengan rentang umur 21-86 tahun. Jumlah masing-masing fitur pada dataset ini dapat dilihat pada Tabel 1.

Tabel 1. Jumlah Data Fitur

| Fitur | Jumlah |
|----------------------|--------|
| Prehypertension | 85 |
| Normal | 80 |
| Stage 1 hypertension | 34 |
| Stage 2 hypertension | 20 |

2. 2 Metode

Pemanfaatan salah satu algoritma *machine learning* digunakan dalam mengklasifikasi jenis hipertensi berdasarkan data pribadi. Pembelajaran mesin dengan teknik diawasi (*supervised learning*) digunakan dalam penelitian ini. Pembelajaran mesin diawasi yaitu melatih mesin dengan memberikan label, artinya beberapa data sudah diberikan data yang benar [19]. Sebagai data inputan pada pembelajaran mesin ini, kami menggunakan delapan fitur dari data pribadi yang telah dikumpulkan dalam bentuk *dataset*. Algoritma yang digunakan ialah *neural network*. Diagram blok yang diusulkan pada penelitian ini bisa dilihat pada Gambar 1.



Gambar 1. Diagram Blok

Berdasarkan Gambar 1, ada beberapa tahapan-tahapan penelitian yang dilakukan. Penelitian dimulai dengan memberikan data input yaitu fitur-fitur dan kelas pada *dataset* hipertensi. *Dataset* tersebut tentunya sudah melewati tahap pemrosesan data (*pre-processing*). Tahap ini memberikan efek perbaikan jika terdapat data yang bersifat *nullable* atau kosong yang kemudian dapat kita hapus atau ganti dengan data rata-rata (*mean*) jika data berupa angka [20]. Data-data fitur yang bersifat string selanjutnya diubah ke dalam bentuk numerik agar algoritma dapat memprosesnya dengan baik [21]. Sebagai contoh kelas “normal” diubah menjadi nilai 1, prehipertensi menjadi 2, hipertensi stadium-1 menjadi 3 dan hipertensi stadium-2 menjadi 4. Hal tersebut dibuat dengan teknik *label encoding* yang mampu merubah data string menjadi data numerik. Data yang sudah bersifat numerik kemudian dilakukan penskalaan fitur yang akan menyatukan tiap variabel mandiri dalam format data baru [22]. Teknik normalisasi juga digunakan untuk menormalkan kolom fitur yang dimiliki pada kisaran [0,1] dengan metode *min-max scaling*. Selanjutnya pemilihan fitur memanfaatkan *principal component analysis* (PCA) untuk menangani masalah *overfitting* akibat redundansi data. PCA berfungsi untuk mengurangi jumlah fitur dari kumpulan data dengan mempertahankan variabel sebanyak mungkin [21]. Data-data yang sudah melewati pemrosesan data selanjutnya dibagi menjadi data latih dan data uji dengan perbandingan 70% berbanding 30% [23]. Data yang sudah matang selanjutnya diterapkan pada model klasifikasi *machine learning* yaitu *neural network*. Hasil pembelajaran ini kemudian di optimasi parameternya hingga menghasilkan parameter yang terbaik dalam pembelajarannya. Hasil optimasi di uji coba kembali dengan teknik evaluasi yang umum digunakan pada model klasifikasi *machine learning*

2. 3 Artificial Neural Networks

Jaringan saraf atau *artificial neural networks* dapat digunakan untuk berbagai tujuan termasuk klasifikasi dan prediksi dalam penambangan data. Jaringan saraf pada awalnya dibangun dalam pembelajaran mesin untuk mencoba meniru *neurofisiologi* otak manusia melalui kombinasi elemen komputasi sederhana (*neuron*) ke dalam sistem yang saling berhubungan [24]. Menurut [25], aktivitas algoritma *neural network* dan otak manusia pada dasarnya terkonsentrasi dalam dua aspek yaitu pengetahuan didapatkan dari proses pembelajaran dan intensitas koneksi antar neuron dikenal sebagai bobot sebagai media yang dapat menyimpan pengetahuan yang diperoleh. *Neural network* dikenal sebagai algoritma yang baik untuk menangani berbagai data yang memiliki *noise* [26].

2. 4 Algoritma Backpropagation

Algoritma ini merupakan salah satu teknik dari *artificial neural network*. *Backpropagation* dapat meminimalisir kesalahan dalam hasil luaran yang dihasilkan oleh jaringan. Algoritma ini dapat melatih jaringan agar menghasilkan keseimbangan antara kemampuan jaringan dalam membaca pola yang digunakan pada saat proses pelatihan dan kemampuan dalam menghasilkan respon yang tepat pada pola yang serupa dengan pola yang digunakan pada proses pelatihan [27]. Terdapat tiga fase dalam pelatihan *backpropagation* antara lain yaitu: (a) *fase forward* dimana pada fase ini pola data latih dihitung mulai dari lapisan input hingga lapisan output; (b) *back forward* yaitu saat masing-masing neuron output menerima target pola input agar terhitung nilai kesalahan yang dipropagasikan mundur; (c) fase modifikasi bobot yang bermanfaat untuk menurunkan nilai kesalahan muncul [28].

Artificial neural network dengan *backpropagation* merupakan jenis topologi multi lapis dengan komposisi satu lapis masukan dengan $X_1 - X_{12}$, satu atau lebih lapisan *hidden layer* $Z_1 - Z_{12}$ dan satu lapis output yaitu Y . Masing-masing lapisan mempunyai *neuron-neuron* yang dimodelkan dengan lingkaran. Antara *neuron* pada satu lapis dengan satu neuron lapis berikutnya terhubung dengan koneksi yang memiliki bobot ke simpul output W_{ij} dan bobot ke simpul *hidden* V_{ij} . Pada *hidden layer* memiliki nilai bias (b) yang memiliki bobot sama dengan 1 [29].

2. 5 Confusion Matrix

Evaluasi model *artificial neural networks* yang digunakan adalah *confusion matrix*. Kriteria klasifikasi dengan metode *confusion matrix* dapat dilihat pada Tabel 2.

Tabel 2. *Confusion Matrix*

| Kelas | Diklasifikasikan sebagai Positif | Diklasifikasikan sebagai Negatif |
|-------|----------------------------------|----------------------------------|
| + | True Positive (TP) | False Positive (FP) |
| - | True Negative (TN) | False Negative (FN) |

Tabel 2 adalah klasifikasi dari *confusion matrix*. *True positive* (TP) berarti model klasifikasi dengan benar memberi label jumlah tupel positif. *True Negative* (TN) berarti bahwa model klasifikasi dengan benar memberi label jumlah tupel negatif. *False positive* (FP) berarti bahwa model klasifikasi memberi label yang salah untuk jumlah tupel negatif. *False Negative* (FN) menunjukkan bahwa model klasifikasi memberi label yang salah untuk jumlah tupel positif. Akurasi adalah ukuran kinerja model klasifikasi dan merupakan persentase jumlah data yang diprediksi dengan benar dari total data [30]. Rumus untuk menghitung ketelitian dapat dilihat pada Persamaan 1.

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

Pengukuran *performance metrics* terdiri dari presisi, *recall* dan *f1-score* [31]. Presisi adalah rasio positif atau derajat keandalan, yaitu proporsi prediksi berlabel positif yang benar terhadap prediksi positif keseluruhan [32]. Rumus untuk menghitung presisi dapat dilihat pada Persamaan 2.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall juga dikenal sebagai *true positive rate* atau sensitivitas. *Recall* juga disebut sebagai derajat keandalan model dalam mendeteksi data berlabel positif dengan benar [32]. Rumus untuk menghitung *recall* dapat dilihat pada persamaan 3.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-score merangkum semua hasil perhitungan *precision* dan *recall* dengan membuat rata-rata harmonik [27]. Rumus untuk menghitung *F1-score* dapat dilihat pada persamaan 4.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Dataset telah mengalami proses transformasi data hingga layak untuk diproses dengan model klasifikasi. Dataset diubah menjadi versi numerik sehingga model klasifikasi *neural network* dapat memprosesnya. Pada field kategori kelas *prehypertension* diubah menjadi angka 0, *stage 1 hypertension* menjadi 1, *stage 2 hypertension* menjadi 2 dan normal menjadi 3. Pada field *sex* juga diubah nilainya menjadi format *number* sehingga jika *male* menjadi 0 dan *female* menjadi 1. Pada field *bmi* angka desimal juga diubah menjadi bilangan bulat sebagai contoh angka

32.66 diformat menjadi 136 dengan *label encoder*. Adapun beberapa perubahan transformasi data dapat dilihat secara rinci pada Tabel 3.

Tabel 3. Transformasi Data

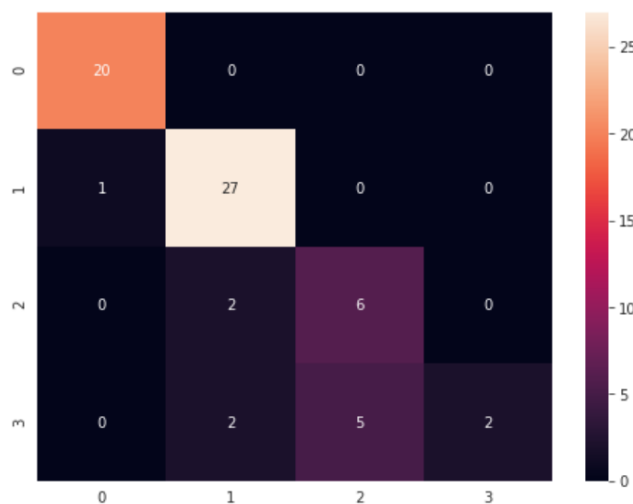
| index | subject | sex | age | height | weight | sbp | dbp | hr | bmi |
|-------|---------|-----|-----|--------|--------|-----|-----|-----|-----|
| 0 | 1 | 2 | 0 | 45 | 152 | 63 | 161 | 89 | 97 |
| 1 | 2 | 3 | 0 | 50 | 157 | 50 | 160 | 93 | 76 |
| 2 | 3 | 6 | 0 | 47 | 150 | 47 | 101 | 71 | 79 |
| 3 | 4 | 8 | 1 | 45 | 172 | 65 | 136 | 93 | 87 |
| 4 | 5 | 9 | 0 | 46 | 155 | 65 | 123 | 73 | 73 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

Modul seleksi memisahkan kelompok data menjadi data pelatihan dan pengujian. Komposisi perbandingan *dataset* data latih dan data uji yaitu 70% berbanding 30% [23]. Penerapan algoritma klasifikasi *artificial neural network* kemudian diuji pada *dataset* yang telah diolah dan dipisahkan sebelumnya. Tabel 4 merupakan hasil performa pengujian (*testing*) dari 219 data uji klasifikasi algoritma *neural network* dengan parameter *solver='lbfgs'*, *alpha=1e-5*, *hidden_layer_sizes=(5, 2)*, *random_state=1*.

Tabel 4. Hasil Klasifikasi *Artificial Neural Networks*

| Kelas | Precision | Recall | F1-Score | Support |
|--------------------|-----------|------------|----------|---------|
| Normal | 0.95 | 1.00 | 0.98 | 20 |
| Pre Hipertensi | 0.87 | 0.96 | 0.92 | 28 |
| Hipertensi Tahap 1 | 0.55 | 0.75 | 0.63 | 8 |
| Hipertensi Tahap 2 | 1.00 | 0.22 | 0.36 | 9 |
| Accuracy : | | 85% | | 65 |

Performa model *neural network* ini juga menghasilkan *confusion matrix*. Blok pada posisi diagonal merupakan nilai TP dan TN. Matrik ini dapat dilihat pada Gambar 2. Pada posisi diagonal kita bisa melihat angka 20,27,0 dan 0. Matrik ini bisa kita gunakan sebagai perhitungan nilai akurasi sesuai dengan persamaan 1.



Gambar 2. Confusion Matrix

Berdasarkan data yang diambil dari Tabel 4, kita dapat menghitung nilai akurasi pengujian dengan rumus (1) dimana nilai secara diagonal yang dijumlahkan dibagi total data uji

sehingga $(20 + 27 + 6 + 2) / 65 = 85\%$. Tabel *confusion matrix* model *neural network* ini dapat dijelaskan secara jelas sebagai berikut:

1. 20 data *test* untuk kelas *Normal*, sistem tersebut memprediksi 20 *Normal*, 0 *Prehypertension*, 0 *Stage 1 hypertension*, dan 0 *Stage 2 hypertension*.
2. 28 data *test* untuk kelas *Prehypertension*, sistem tersebut memprediksi 1 *Normal*, 27 *Prehypertension*, 0 *Stage 1 hypertension* dan 0 *Stage 2 hypertension*.
3. 8 data *test* untuk kelas *Stage 1 hypertension*, sistem tersebut memprediksi 0 *Normal*, 2 *Prehypertension*, 6 *Stage 1 hypertension* dan 0 *Stage 2 hypertension*.
4. 9 data *test* untuk kelas *Stage 2 hypertension*, sistem tersebut memprediksi 0 *Normal*, 2 *Prehypertension*, 5 *Stage 1 hypertension* dan 0 *Stage 2 hypertension*.

4. KESIMPULAN

Berdasarkan hasil penelitian, ditemukan bahwa model klasifikasi *neural network* dapat digunakan sebagai alat klasifikasi untuk prediksi jenis penyakit hipertensi. Pengolahan data diperlukan dalam penelitian ini karena model *neural network* membutuhkan transformasi *dataset* dalam format numerik. *Dataset* yang sudah diproses selanjutnya dibagi dengan komposisi data latih sebesar 70% dan data uji sebesar 30%. Implementasi *neural network* menggunakan parameter *solver='lbfgs'*, *alpha=1e-5*, *hidden_layer_sizes=(5, 2)*, *random_state=1*. Parameter ini kemudian diterapkan kembali pada tahap klasifikasi jenis penyakit hipertensi. Hasilnya adalah akurasi prediksi jenis hipertensi diperoleh nilai sebesar 85%.

5. SARAN

Pada penelitian selanjutnya lebih baik dilakukan optimasi pada parameter yang digunakan pada model artificial neural network. Penelitian selanjutnya juga disarankan menggunakan *dataset* jenis penyakit hipertensi yang lebih besar agar proses training dan testing berjalan lebih maksimal.

DAFTAR PUSTAKA

- [1] D. Masruroh, E. M. M. Has, and R. Fauziningtyas, "Pengaruh Terapi Humor dengan Media Film Komedi terhadap Penurunan Tekanan Darah Pada Lansia Dengan Hipertensi," *Indones. J. Community Heal. Nurs.*, vol. 4, no. 1, p. 29, 2019, doi: 10.20473/ijchn.v4i1.12496.
- [2] M. Nour and K. Polat, "Automatic Classification of Hypertension Types Based on Personal Features by Machine Learning Algorithms," *Math. Probl. Eng.*, vol. 2020, pp. 1–14, 2020, doi: 10.1155/2020/2742781.
- [3] A. C. Telaumbanua and Y. Rahayu, "Penyuluhan Dan Edukasi Tentang Penyakit Hipertensi," *J. Abdimas Saintika*, vol. 3, no. 1, p. 119, 2017, doi: 10.30633/jas.v3i1.1069.
- [4] Y. Nursakinah and A. Handayani, "Faktor-Faktor Risiko Hipertensi Diastolik Pada Usia Dewasa Muda," *J. Pandu Husada*, vol. 2, no. 1, p. 21, 2021, doi: 10.30596/jph.v2i1.5426.
- [5] G. Mancina *et al.*, "2018 ESC/ESH Guidelines for the management of arterial hypertension," *Eur. Heart J.*, vol. 39, pp. 3021–3104, 2018, doi: doi:10.1093/eurheartj/ehy339.
- [6] D. Tryastuti, "Determinan Pre-Hipertensi Di Kelurahan Curug Kecamatan Cimanggis Kota Depok," *Indones. J. Heal. Sci.*, vol. 11, no. 1, p. 71, 2019, doi: 10.32528/ijhs.v11i1.2240.
- [7] Y. T. G. Arum, "Hipertensi pada Penduduk Usia Produktif (15-64 Tahun)," *Higeia J. Public Heal. Res. Dev.*, vol. 3, no. 3, pp. 84–94, 2019, doi: https://doi.org/10.15294/higeia.v3i3.30235.
- [8] A. Syntya, "Hypertension and heart disease: literature review," *J. Ilm. Permas J. Ilm. STIKES Kendal*, vol. 11, no. 4, pp. 541–550, 2021, doi: doi.org/10.32583/pskm.v11i4.1621.
- [9] F. D. Telaumbanua, P. Hulu, T. Z. Nadeak, R. R. Lumbantong, and A. Dharma, "Penggunaan

- Machine Learning Di Bidang Kesehatan,” *J. Teknol. dan Ilmu Komput.*, vol. 3, no. 1, pp. 57–64, 2019, doi: <https://doi.org/10.34012/jutikomp.v2i2.657>.
- [10] A. Mustafa and M. Rahimi Azghadi, “Automated machine learning for healthcare and clinical notes analysis,” *Computers*, vol. 10, no. 2, pp. 1–31, 2021, doi: 10.3390/computers10020024.
- [11] A. Roihan, P. Abas Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: <https://doi.org/10.31294/ijcit.v5i1.7951>.
- [12] M. M. Santoni, N. Chamidah, and N. Matondang, “Prediksi Hipertensi menggunakan Decision Tree, Naïve Bayes dan Artificial Neural Network pada software KNIME,” *Techno.Com*, vol. 19, no. 4, pp. 353–363, 2020, doi: 10.33633/tc.v19i4.3872.
- [13] W. Apriliah, I. Kurniawan, M. Baydhowi, and T. Haryati, “Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest,” *Sistemasi*, vol. 10, no. 1, p. 163, 2021, doi: 10.32520/stmsi.v10i1.1129.
- [14] L. B. Moreira and A. A. Namen, “A hybrid data mining model for diagnosis of patients with clinical suspicion of dementia,” *Comput. Methods Programs Biomed.*, vol. 165, pp. 139–149, 2018, doi: 10.1016/j.cmpb.2018.08.016.
- [15] N. Chamidah, M. Mega Santoni, and N. Matondang, “Pengaruh Oversampling pada Klasifikasi Hipertensi dengan Algoritma Naïve Bayes, Decision Tree, dan Artificial Neural Network (ANN),” *J. RESTI (Rekayasa Sist. dan Teknol. Inf.)*, vol. 4, no. 4, pp. 635–641, 2020, doi: <https://doi.org/10.29207/resti.v4i4.2015>.
- [16] A. Akbar Ritonga, Ibnu Rasyid Munthe, Masrizal, “LVQ Algorithm for The Classification of Hypertension Based on ESH Guideline,” *J. Mantik*, vol. 4, no. 3, pp. 1772–1778, 2020, doi: <https://doi.org/10.35335/mantik.Vo4.2020.1006.pp1772-1778>.
- [17] E. Martinez-Ríos, L. Montesinos, and M. Alfaro-Ponce, “A machine learning approach for hypertension detection based on photoplethysmography and clinical data,” *Comput. Biol. Med.*, vol. 145, no. March, p. 105479, 2022, doi: 10.1016/j.compbimed.2022.105479.
- [18] “PPG Blood Pressure,” *IEEE Dataport*, 2019. <https://iee-dataport.org/open-access/ppg-blood-pressure>.
- [19] A. Yakimovich, A. Beaugnon, Y. Huang, and E. Ozkirimli, “Labels in a haystack: Approaches beyond supervised learning in biomedical applications,” *Patterns*, vol. 2, no. 12, pp. 1–11, 2021, doi: 10.1016/j.patter.2021.100383.
- [20] A. Toha, P. Purwono, W. Gata, and A. Toha, “Model Prediksi Kualitas Udara dengan Support Vector Machines dengan Optimasi Hyperparameter GridSearch CV,” *Bul. Ilm. Sarj. Tek. Elektro*, vol. 4, no. 1, pp. 12–21, 2022, doi: 10.12928/biste.v4i1.6079.
- [21] P. Purwono, A. Wirasto, and K. Nisa, “Comparison of Machine Learning Algorithms for Classification of Drug Groups,” *Sisfotenika*, vol. 11, no. 2, p. 196, 2021, doi: 10.30700/jst.v11i2.1134.
- [22] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, “A survey on missing data in machine learning,” *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00516-9.
- [23] P. Purwono, A. Ma’arif, I. S. Mangku Negara, W. Rahmaniar, and J. Rahmawan, “Linkage Detection of Features that Cause Stroke using Feyn Qlattice Machine Learning Model,” *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 7, no. 3, p. 423, 2021, doi: 10.26555/jiteki.v7i3.22237.
- [24] A. M. Siregar and H. H. H, “Implementasi Algoritma Neural Network untuk Mendukung Keputusan di Desa Tamanmekar,” *Petir*, vol. 13, no. 1, pp. 21–32, 2020, doi: 10.33322/petir.v13i1.768.
- [25] F. Gorunescu, *Data Mining: Concepts, models and techniques*. Berlin: Springer-Verlag, 2011.
- [26] D. T. Larose, *Discovering knowledge in data : an introduction to data mining*. Jhon Wiley & Sons Inc, 2005.
- [27] A. Wanto, “Penerapan Jaringan Saraf Tiruan Dalam Memprediksi Jumlah Kemiskinan,” *Klik - Kumpul. J. Ilmu Komput.*, vol. 5, no. 1, p. 61, 2018.
- [28] D. Monika, A. Ahmad, S. Wardani, and Solikhun, “Model Jaringan Syaraf Tiruan Dalam Memprediksi Ketersediaan Cabai Berdasarkan Provinsi,” *Teknika*, vol. 8, no. 1, pp. 17–24, 2019,

- doi: 10.34148/teknika.v8i1.140.
- [29] A. Ahmad, P. M. Putri, W. Alifah, I. Gunawan, and . S., “Analisis Jaringan Syaraf Tiruan Metode Backpropogation Dalam Memprediksi Ketersediaan Komoditas Beras Berdasarkan Provinsi Di Indonesia,” *J. Resist. (Rekayasa Sist. Komputer)*, vol. 2, no. 1, pp. 48–60, 2019, doi: 10.31598/jurnalresistor.v2i1.348.
- [30] S. Katoch, V. Singh, and U. S. Tiwary, “Indian Sign Language Recognition System using SURF with SVM and CNN,” *Array*, p. 100141, 2022, doi: <https://doi.org/10.1016/j.array.2022.100141>.
- [31] X. Xiong, S. Hu, D. Sun, S. Hao, H. Li, and G. Lin, “Detection of false data injection attack in power information physical system based on SVM–GAB algorithm,” *Energy Reports*, vol. 8, pp. 1156–1164, 2022, doi: 10.1016/j.egy.2022.02.290.
- [32] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, “The impact of class imbalance in classification performance metrics based on the binary confusion matrix,” *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.
-