

Deteksi Hipertensi Dengan Metode Artificial Neural Network

Nurhafifah Matondang, Mayanda Mega Santoni,
Nurul Chamidah

Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jakarta
email:nurhafifahmatondang@upnvj.ac.id
Jl. Rs. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450, Indonesia

Abstrak

Hipertensi atau tekanan darah tinggi merupakan salah satu penyakit yang menjadi penyebab kematian di Indonesia. Penelitian ini menggunakan metode Artificial Neural Network (ANN) Feed Forward Backpropagation untuk mendeteksi hipertensi berdasarkan ciri-ciri dan atau kebiasaan seseorang. Data hipertensi diperoleh dengan kuesioner kemudian diproses dengan melakukan imputasi pada missing value serta transformasi data untuk mengubah data pada range yang sama. Hasil pra-proses digunakan untuk melatih model ANN dengan 25 input, 3 hidden neuron serta 1 output. Dengan learning rate 0.25 serta epoch sebanyak 1000, hasil percobaan menunjukkan akurasi hingga 96% dengan pembagian data latih sebesar 80% dan data uji sebesar 20%.

Kata kunci: hipertensi, ANN, backpropagation

1 PENDAHULUAN

Hipertensi atau tekanan darah tinggi merupakan salah satu penyakit yang menjadi penyebab kematian. Saat ini pengecekan hipertensi merupakan kondisi yang sering ditemukan pada pelayanan kesehatan primer. Di Indonesia data penderita hipertensi setiap tahun mengalami peningkatan selama 5 (lima) tahun terakhir dari tahun 2013-2018, menurut data Riskesdas meningkat sekitar 9.4%, ditentukan dalam beberapa prevalensi diantaranya berdasarkan umur, diagnosis dokter dan penggunaan obat antihipertensi. Informasi ini diperoleh dari laman web Departemen Kesehatan Badan Penelitian dan Pengembangan Kesehatan (Anwar, 2017).

Machine Learning (ML) merupakan sub bidang dari *Artificial Intelligence* (AI), salah satu pendekatan yang menuju ke arah kecerdasan buatan bertujuan untuk mempelajari dan menangani data dalam jumlah besar dengan menggunakan algoritma komputer. *Machine learning* akan mempelajari dari pola-pola yang diberikan untuk dipelajari dan dianalisa. Beberapa tipe-tipe model yang terdapat dalam *machine learning* diantaranya klasifikasi, prediksi, *clustering* dan *explanation*.

Beberapa contoh program *machine learning* yang telah digunakan dalam kehidupan sehari-hari diantaranya pendeteksi spam, pendeteksi wajah, rekomendasi produk, asisten virtual, diagnosa medis dan masih banyak lagi contoh lainnya yang juga telah dilakukan penelitian. Diagnosa medis dengan *machine learning* memudahkan tenaga medis dalam memprediksi berbagai penyakit dari kumpulan data pasien-pasien sebelumnya, sehingga diharapkan hasil analisis dapat membantu tim medis dalam mengambil keputusan.

Kehadiran Kecerdasan Buatan (*Artificial Intelligence*) di industri kesehatan dengan cepat meningkatkan kualitas pelayanan kesehatan khususnya di Indonesia. Dua model kecerdasan buatan seperti *machine learning* dan *deep learning* dengan mudah dan cepat dalam

mempelajari dan mengolah data untuk memperoleh informasi yang dibutuhkan.

Salah satu contoh hasil penelitian yang memanfaatkan *machine learning* di bidang kesehatan seperti klasifikasi untuk diagnosa diabetes menggunakan metode *Bayesian Regularization Neural Network* (RBNN) untuk mengklasifikasi gejala diabetes secara cepat dan akurat (Muhammad, et.al., 2017). Selain itu, Hidayatul, et.al (2018) juga telah melakukan desain sistem klasifikasi kelainan jantung menggunakan *Learning Vector Quantization* untuk klasifikasi kelainan jantung berdasarkan citra digital electrocardiogram.

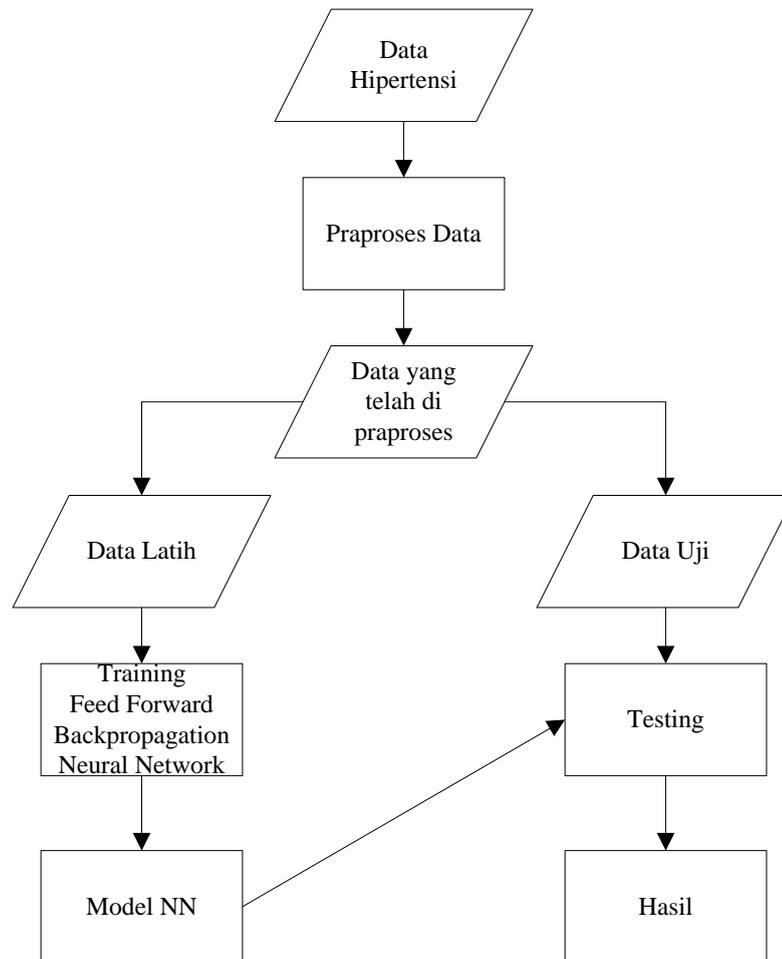
Pada penelitian Lafreniere, et.al. (2016), faktor risiko penting yang digunakan untuk memprediksi hipertensi dengan *artificial neural network* terdiri dari 11 variable yang meliputi tahun kelahiran, jenis kelamin, indeks masa tubuh, tekanan darah sistolik, tekanan darah diastolic, lipoprotein kepadatan tinggi, lipoprotein kepadatan rendah, triglycerides, kolesterol, micro-albumin dan urine albumin. Tingkat dengan akurasi yang dihasilkan adalah 82% dengan menggunakan dataset berjumlah 185.371 untuk memprediksi pasien hipertensi dan tidak hipertensi.

Penelitian dari Wang, et.al. (2015) bertujuan untuk membangun model prediksi penyakit hipertensi berdasarkan penggunaan *hybrid* dari *logistic regression* dan *artificial neural network*. Adapun variable yang digunakan terdapat 14 variabel terdiri dari usia, jenis kelamin, tinggi badan, berat badan, status pernikahan, pendidikan, penghasilan, kebiasaan olahraga, diabetes, *hyperlipemia*, kebiasaan merokok, minum alkohol dan riwayat hipertensi. Jumlah dataset yang digunakan untuk pengujian ini yaitu 308.711 pasien. Hasil pengujian dengan model *artificial neural network* menunjukkan akurasi 72%.

Penelitian Assaghir, et.al. (2017) mendeteksi hipertensi dengan metode *neural network*. Terdapat 10 variable yang digunakan pada penelitian ini diantaranya jenis kelamin, detak jantung, indeks masa tubuh, berat badan, pinggang, panggul, aktivitas fisik, kebiasaan merokok, penggunaan garam dan tingkat stress. Jumlah data set untuk memperoleh hasil akurasi yaitu sejumlah 3000 partisipan. Hasil prediksi terhadap tekanan darah sistolik dengan akurasi 7.54% dan tekanan darah diastolic dengan akurasi 9.84%.

2 METODOLOGI PENELITIAN

Metodologi yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1. Penelitian dilakukan dengan mengumpulkan data, dan kemudian dilakukan suatu praproses data dengan imputasi *missing value* dan transformasi data. Data yang telah dipraproses dibagi menjadi data latih dan data uji. Data latih digunakan pada proses pelatihan untuk membangun model *neural network* dengan proses training dan data uji digunakan untuk menguji model yang telah dibangun pada saat pelatihan.



Gambar 1:Metode Penelitian

2.1 Data

Data diperoleh dari kuesioner yang disebarakan kepada responden dan diperoleh 80 *record* data. Data hipertensi terdiri dari 80 *record* dengan 25 fitur dan kelas hipertensi yang berisi 40 *record* dan tidak hipertensi sebanyak 40 *record*.

2.2 Praproses

Praproses dilakukan dengan mengimputasi nilai yang hilang dengan metode median, setelah diimputasi, data ditransformasi min-max agar nilai pada setiap variable memiliki range yang sama. Imputasi median dilakukan dengan mengurutkan data pada variabel yang sama dari terkecil hingga terbesar. Kemudian mengambil nilai tengah untuk mengisi nilai yang kosong (*missing value*) pada variabel tersebut.

Setelah dilakukan imputasi pada *missing value*, praproses dilakukan dengan melakukan transformasi min-max. Transformasi ini dilakukan pada setiap variabel secara terpisah. Berikut formula normalisasi min-max (Han et. al., 2011).

$$X_{baru} = \frac{X_{lama} - \min}{maks - \min} (maks_baru - \min_baru) + \min_baru \quad (1)$$

Xbaru : nilai baru setelah transformasi
Xlama : nilai lama sebelum transformasi
min : nilai minimum pada variabel
maks : nilai maksimum pada variabel
maks_baru : nilai maksimum baru yang diinginkan

min_baru : nilai minimum baru yang diinginkan

Dalam hal ini, transformasi dilakukan untuk mengubah nilai dari setiap variabel ke dalam range yang sama yakni 0-1, sehingga nilai maks_baru adalah 1 dan nilai min_baru adalah 0.

2.3 Data Latih dan Data Uji

Data yang telah dipraproses dibagi menjadi data latih dan data uji. Data latih digunakan untuk membangun model *Neural Network* melalui proses pelatihan, sedangkan data uji digunakan untuk menguji model yang telah dibangun pada proses pelatihan sebelumnya. Pengujian ini dilakukan dalam proses pengujian. Data yang digunakan untuk pelatihan adalah 75% dan 25% sisanya digunakan untuk pengujian, kemudian 80% dan 20%, serta 70% dan 30%. Pemilihan data untuk pelatihan dan pengujian dilakukan secara random.

2.4 Feed Forward Backpropagation Neural Network

Model yang digunakan untuk klasifikasi hipertensi ini adalah model *Feed Forward Neural Network* dengan *input layer* sebanyak 25 *neuron* yakni sejumlah variabel dari hipertensi, 1 *hidden layer* terdiri dari 3 *neuron* dan *output layer* sebanyak 1 *neuron* untuk menentukan hipertensi atau tidak. Untuk aktivasinya, digunakan fungsi sigmoid biner.

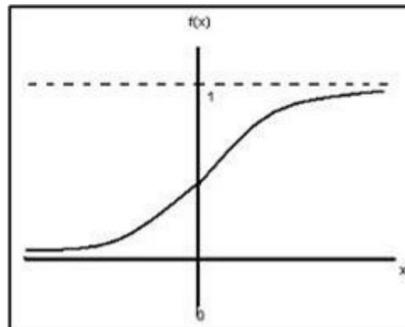
Fungsi Sigmoid biner digunakan untuk *Neural Network* yang menggunakan metode Backpropagation. Fungsi Sigmoid unipolar ini memiliki fase antara nol sampai dengan satu untuk model yang membutuhkan nilai output benar atau salah saja ($true=1$ atau $false=0$). Persamaan dari fungsi Sigmoid unipolar adalah sebagai berikut:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

dengan turunan

$$f'(x) = f(x)(1 - f(x)) \quad (3)$$

Sementara itu, ilustrasi dari fungsi Sigmoid unipolar dapat dilihat pada Gambar 2.



Gambar 2: Sigmoid Biner

3 HASIL DAN PEMBAHASAN

Tabel 1 merupakan potongan data yang diperoleh dari kuesioner. Terlebih dahulu, data tersebut dilakukan praproses. Pada baris 2 di kolom fitur 1, terdapat data *missing value*. Data ini tidak diisi oleh responden.

Tabel 1. Potongan Data Penelitian

Fitur 1	Fitur 2	Fitur 3	Fitur 4	Fitur 5	...	Fitur 25	Kelas
32	1	1	167	64	...	1	1

	2	1	158	82	...	1	1
30	1	1	178	68	...	2	2
25	2	2	150	95	...	1	2
30	2	2	156	45	...	2	2
23	2	3	163	88	...	2	2

Praproses data dimulai dari melakukan imputasi nilai yang hilang (*missing value*) dengan metode median. Data hasil imputasi dapat dilihat pada Tabel 2. Pada fitur 1 baris ke dua menunjukkan data setelah dilakukan imputasi.

Tabel 2. Data setelah dilakukan imputasi

Fitur 1	Fitur 2	Fitur 3	Fitur 4	Fitur 5	...	Fitur 25	Kelas
32	1	1	167	64	...	1	1
32	2	1	158	82	...	1	1
30	1	1	178	68	...	2	2
25	2	2	150	95	...	1	2
30	2	2	156	45	...	2	2
23	2	3	163	88	...	2	2

Selanjutnya, data ditransformasi dengan menggunakan normalisasi min-max agar nilai pada setiap variable memiliki range yang sama. Tabel 3 menunjukkan data hasil tranformasi.

Tabel 3. Data setelah dilakukan transformasi (normalisasi min-max)

Fitur 1	Fitur 2	Fitur 3	Fitur 4	Fitur 5	...	Fitur 25	Kelas
0.28	0	0	0.49	0.17	...	0	0
0.28	1	0	0.44	0.3	...	0	0
0.23	0	0	0.55	0.2	...	1	1
0.12	1	0.5	0.4	0.4	...	0	1
0.23	1	0.5	0.43	0.02	...	1	1
0.07	1	1	0.47	0.35	...	1	1

Setelah data selesai diproses, selanjutnya data siap dilakukan pelatihan dengan metode *Feed Forward Backpropagation*. Model *Feed Forward Backpropagation* yang terdiri dari 25 input variabel, 3 hidden neuron serta 1 output diujicobakan pada data hipertensi dengan fungsi aktivasi sigmoid, dan learning rate: 0.25. Pelatihan berhenti pada epoch ke-1000. Tabel 4 berikut menunjukkan hasil pengujian dengan tiga proporsi pembagian data uji dan data latih yang berbeda. Dari hasil tersebut diperoleh hasil terbaik dengan akurasi rata-rata 96% dengan membagi data dengan 80% data latih, dan 20% untuk data uji.

Tabel 4. Hasil percobaan

Pembagian Data Latih (%)	Pembagian Data Uji (%)	Akurasi Uji (%)
80	20	96
75	25	90
70	30	91

4 KESIMPULAN

Pada penelitian ini dilakukan evaluasi klasifikasi data hipertensi dengan 25 variabel menggunakan algoritma Artificial Neural Network (ANN) yakni Feed Forward

Backpropagation. Sebelum dimodelkan dengan ANN, data dipraproses dan dibagi menjadi data latih dan data uji di mana data latih digunakan untuk membangun model ANN, sedangkan data uji digunakan untuk menguji model ANN.

Model ANN yang digunakan terdiri atas 25 input variabel, 3 hidden neuron serta 1 output dengan fungsi aktivasi sigmoid, dan learning rate 0.25 serta epoch sebanyak 1000. Hasil pengujian menunjukkan pembagian data menjadi 80% data latih dan 20% data uji menghasilkan akurasi mencapai 96%.

Referensi

- Han, J., Kamber, M. and Pei, J. (2011), *Data mining: concepts and techniques*. Elsevier.
- Anwar F. (2017). *Kemenkes Sebut Kasus Hipertensi di Indonesia Terus Meningkat*. Retrieved October 1, 2018, from <https://health.detik.com/berita-detikhealth/d-3503396/kemenkes-sebut-kasus-hipertensi-di-indonesia-terus-meningkat>
- Assaghir, Z., Janbain, A., Makki, S., Kurdi, M., & Karam, R. (2017). Using Neural Network to predict the Hypertension 1, 2(2), 35–38.
- Hidayatul, S., Aini, A., Sari, Y. A., & Arwan, A. (2018). Seleksi Fitur Information Gain untuk Klasifikasi Penyakit Jantung Menggunakan Kombinasi Metode K-Nearest Neighbor dan Naïve Bayes, 2(9), 2546–2554.
- Muhammad F. R., Muhammad I D., Dion A. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN), 11(1), 36-45.
- Lafreniere, D., Zulkernine, F., Barber, D., & Martin, K. (2016). Using Machine Learning to Predict Hypertension from a Clinical Dataset. *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–7. <https://doi.org/10.1109/SSCI.2016.7849886>
- Wang, A., An, N., Chen, G., Li, L., & Alterovitz, G. (2015). Expert Systems with Applications Predicting hypertension without measurement: A non-invasive, questionnaire-based approach. *EXPERT SYSTEMS WITH APPLICATIONS*, 42(21), 7601–7609. <https://doi.org/10.1016/j.eswa.2015.06.012>