

Perbandingan Algoritma Decision Tree dan Deep Learning dalam Prediksi Masalah Kesehatan berdasarkan Kebiasaan Gaya Hidup

Fadhillah Rashidatul A'la, Zaehol Fatah

Universitas Ibrahimy, Indonesia

E-mail: fadhhillah5@gmail.com, zaeholfatah@gmail.com

KEYWORD

*lifestyle habits;
health prediction;
decision tree; deep
learning*

ABSTRACT

The development of modern lifestyles that tend to be unhealthy has caused a significant increase in various health problems, such as obesity, hypertension, and diabetes. To overcome this problem, a health prediction system is needed that is able to identify risks based on lifestyle habits. This study aims to compare the performance of the Decision Tree and Deep Learning algorithms in predicting health problems based on lifestyle habits. An unhealthy lifestyle is often the main cause of various health problems, so an accurate prediction system is needed to prevent them. The method used is quantitative analysis with a lifestyle habit dataset, using both algorithms to predict health status. The results showed that Deep Learning was superior with an accuracy of 92.31%, compared to Decision Tree which only reached 89.74%. This finding strengthens the effectiveness of Deep Learning in handling complex data and non-linear relationships between lifestyle variables. In conclusion, the Deep Learning algorithm is more appropriate for preventive health prediction systems. This study contributes to the development of data-based health prediction technology and provides recommendations for expanding the scope of future research, such as increasing sample size and using other algorithms. In conclusion, Deep Learning is proven to be more effective in handling complex data with non-linear patterns, making it more suitable for health prediction systems based on lifestyle habits. This study contributes to the development of health prediction technology and recommends the application of more adaptive algorithms and testing with wider samples for further research.

KATA KUNCI

kebiasaan gaya hidup; prediksi kesehatan; pohon keputusan; pembelajaran mendalam

ABSTRAK

Perkembangan gaya hidup modern yang cenderung kurang sehat telah menyebabkan peningkatan signifikan pada berbagai masalah kesehatan, seperti obesitas, hipertensi, dan diabetes. Untuk mengatasi masalah ini, diperlukan sistem prediksi kesehatan yang mampu mengidentifikasi risiko berdasarkan kebiasaan gaya hidup. Penelitian ini bertujuan untuk membandingkan kinerja algoritma Decision Tree dan Deep Learning dalam memprediksi masalah kesehatan berdasarkan kebiasaan gaya hidup. Gaya hidup yang tidak sehat sering kali menjadi penyebab utama berbagai

masalah kesehatan, sehingga diperlukan sistem prediksi yang akurat untuk mencegahnya. Metode yang digunakan adalah analisis kuantitatif dengan dataset kebiasaan gaya hidup, menggunakan kedua algoritma tersebut untuk memprediksi status kesehatan. Hasil penelitian menunjukkan bahwa Deep Learning lebih unggul dengan akurasi 92.31%, dibandingkan Decision Tree yang hanya mencapai 89.74%. Temuan ini memperkuat efektivitas Deep Learning dalam menangani data yang kompleks dan hubungan non-linear antara variabel gaya hidup. Kesimpulannya, algoritma Deep Learning lebih tepat digunakan untuk sistem prediksi kesehatan preventif. Penelitian ini memberikan kontribusi pada pengembangan teknologi prediksi kesehatan berbasis data dan memberikan rekomendasi untuk memperluas cakupan penelitian di masa depan, seperti peningkatan ukuran sampel dan penggunaan algoritma lain. Kesimpulannya, Deep Learning terbukti lebih efektif dalam menangani data kompleks dengan pola non-linear, sehingga lebih sesuai untuk sistem prediksi kesehatan berbasis kebiasaan gaya hidup. Penelitian ini berkontribusi pada pengembangan teknologi prediksi kesehatan dan merekomendasikan penerapan algoritma yang lebih adaptif serta pengujian dengan sampel yang lebih luas untuk penelitian lanjutan.

PENDAHULUAN

Dalam beberapa dekade terakhir, prevalensi masalah kesehatan yang terkait dengan gaya hidup seperti obesitas, diabetes, hipertensi, dan penyakit jantung terus meningkat secara global. Menurut Organisasi Kesehatan Dunia (World Health Organization), gaya hidup yang tidak sehat, seperti kurangnya aktivitas fisik, pola makan yang buruk, dan kebiasaan merokok, menjadi faktor utama penyebab munculnya berbagai penyakit kronis ((Ataey et al., 2020). Fenomena ini memerlukan perhatian serius mengingat dampaknya yang besar terhadap kualitas hidup dan beban ekonomi pada sistem kesehatan masyarakat (Utiarahman & Pratama, 2024). Oleh karena itu, identifikasi dini faktor-faktor risiko kesehatan yang terkait dengan gaya hidup menjadi tantangan penting dalam dunia kesehatan modern (Gilmore, 2024).

Di Indonesia, perubahan pola gaya hidup masyarakat yang semakin sedentari dan meningkatnya konsumsi makanan cepat saji berkontribusi pada peningkatan angka obesitas dan penyakit terkait lainnya (Carney et al., 2023). Menurut data dari Kementerian Kesehatan Indonesia, prevalensi obesitas pada orang dewasa meningkat dari tahun ke tahun, dan masalah ini tidak hanya terjadi di perkotaan tetapi juga di pedesaan (Kementrian kesehatan RI, 2018). Di tengah kondisi ini, ada kebutuhan yang mendesak untuk mengembangkan sistem prediksi kesehatan yang dapat membantu mengidentifikasi individu-individu yang berisiko mengalami masalah kesehatan akibat gaya hidup mereka. Namun, metode manual yang digunakan oleh tenaga medis saat ini sering kali terbatas oleh waktu dan tenaga (Hermawan & Hananto, 2020).

Penelitian sebelumnya telah mengusulkan berbagai pendekatan untuk memprediksi risiko kesehatan, termasuk algoritma machine learning. Sebuah penelitian yang dilakukan oleh Siti Andini Utiarahman menunjukkan bahwa algoritma Decision Tree lebih unggul diantara algoritma lainnya untuk hasil akurasi, presisi, recall, dan f1-score (Telaumbanua et al., 2019). Di sisi lain, penelitian oleh Fitriana Sari membuktikan bahwa penggunaan Deep Learning dalam kesehatan digital mampu meningkatkan akurasi diagnosis penyakit, seringkali melebihi kemampuan manusia dalam beberapa kasus. Deep Learning menunjukkan efektivitas tinggi dalam menganalisis data

medis kompleks, seperti gambar medis dan data sekuensial, menghasilkan prediksi diagnosis dengan akurasi yang sangat baik (Sari, 2023). Meskipun demikian, perbandingan menyeluruh antara kedua algoritma ini dalam konteks prediksi masalah kesehatan berbasis gaya hidup masih minim (Collins et al., 2009).

Urgensi dari penelitian ini terletak pada perlunya metode prediksi yang lebih efektif dan efisien dalam mendeteksi masalah kesehatan berdasarkan data kebiasaan gaya hidup. Pendekatan yang menggabungkan machine learning dan artificial intelligence dapat menjadi solusi inovatif dalam dunia medis, membantu pencegahan penyakit sebelum gejala klinis muncul (Tamalumu et al., 2021). Kebaharuan (novelty) dari penelitian ini adalah pada perbandingan langsung antara algoritma Decision Tree dan Deep Learning, yang belum banyak diteliti secara mendalam dalam konteks gaya hidup dan masalah kesehatan di Indonesia.

Tujuan penelitian ini adalah untuk menganalisis dan membandingkan kinerja algoritma Decision Tree dan Deep Learning dalam memprediksi masalah kesehatan berdasarkan data kebiasaan gaya hidup. Hasil dari penelitian ini diharapkan dapat memberikan wawasan baru bagi praktisi kesehatan dalam memilih metode prediksi yang lebih tepat dan efisien, serta menjadi referensi dalam pengembangan teknologi kesehatan berbasis machine learning di masa depan (Badidi, 2023). Implikasi dari penelitian ini juga dapat berkontribusi pada upaya pencegahan dini masalah kesehatan melalui intervensi yang lebih tepat sasaran (Wicaksono et al., 2023).

Manfaat penelitian ini meliputi beberapa aspek penting, baik dalam ranah praktis maupun akademis. Dalam dunia kesehatan, penelitian ini berpotensi membantu penyedia layanan kesehatan dan praktisi medis dalam mengembangkan sistem prediksi kesehatan berbasis data yang dapat memantau risiko penyakit akibat kebiasaan gaya hidup. Hal ini memungkinkan deteksi dini risiko kesehatan yang lebih efektif, sehingga intervensi dapat dilakukan dengan cepat dan tepat (Adistie et al., 2018). Dari segi peningkatan akurasi, penelitian ini menyediakan perbandingan antara algoritma Decision Tree dan Deep Learning, memberikan wawasan yang berharga tentang algoritma yang lebih efektif untuk memprediksi risiko kesehatan terkait gaya hidup (Yudistira & Kom, 2024). Pengetahuan ini penting dalam pengembangan sistem rekomendasi kesehatan yang lebih akurat, yang dapat diterapkan dalam aplikasi kesehatan digital. Kontribusi akademis penelitian ini juga signifikan dalam bidang kesehatan dan data science, menambah literatur terkait penggunaan machine learning untuk prediksi kesehatan berbasis kebiasaan gaya hidup. Hasil penelitian ini diharapkan menjadi referensi bagi studi-studi lanjutan di bidang kesehatan preventif dan data science, khususnya yang memerlukan pemilihan algoritma yang sesuai untuk data kesehatan yang kompleks. Selain itu, penelitian ini juga mendukung pengembangan teknologi prediksi berbasis data gaya hidup, memungkinkan sistem untuk memberikan rekomendasi kesehatan personal yang dapat meningkatkan kesadaran masyarakat tentang dampak gaya hidup mereka. Terakhir, penelitian ini memberikan rekomendasi yang dapat diimplementasikan dalam aplikasi teknologi kesehatan, terutama bagi pengembang aplikasi kesehatan yang ingin memanfaatkan data gaya hidup untuk tujuan preventif dalam perawatan kesehatan pada skala yang lebih luas.

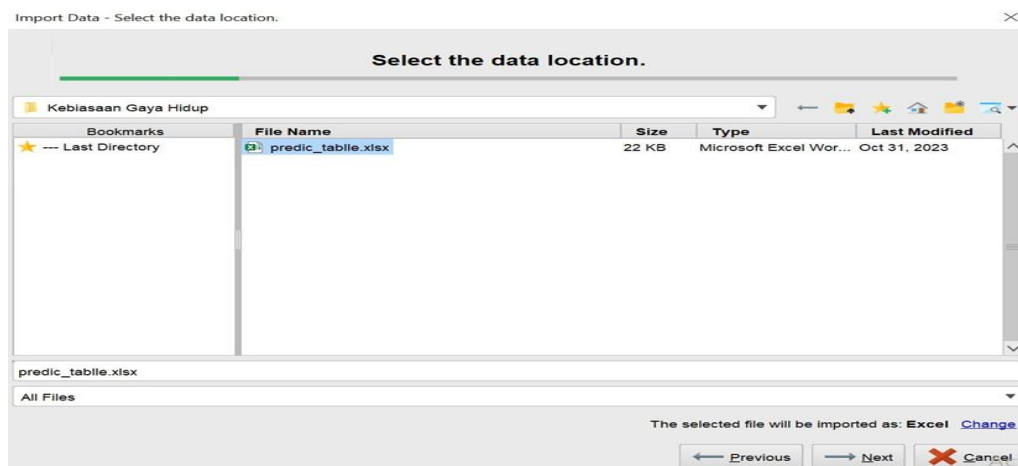
METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan eksperimen yang memfokuskan pada perbandingan dua algoritma machine learning, yaitu Decision Tree dan Deep Learning, untuk memprediksi apakah seseorang memiliki masalah kesehatan berdasarkan kebiasaan gaya hidupnya. Data gaya hidup responden dianalisis menggunakan kedua algoritma ini untuk kemudian dibandingkan performanya berdasarkan akurasi dan metrik lainnya.

Data yang digunakan dalam penelitian ini adalah data publik yang dapat diakses melalui <https://www.kaggle.com/datasets/rustaas/kebiasaan-buruk-berdampak-ke-kesehatan>. Tabel dalam dataset yang digunakan mencakup berbagai variabel yang merepresentasikan kebiasaan gaya hidup, seperti usia, jenis kelamin, kebiasaan merokok, bekerja atau tidak, kebiasaan begadang, aktivitas olahraga, pola makan, dan penyakit bawaan. Data ini diklasifikasikan menjadi dua kategori, yaitu memiliki masalah kesehatan atau tidak, yang dinyatakan dalam variabel "Hasil". Metode pengumpulan data dalam penelitian ini melibatkan beberapa langkah utama untuk pengolahan data menggunakan RapidMiner, yaitu:

1. Import Data

Langkah pertama adalah mengimpor dataset publik ke RapidMiner. Dataset ini berisi variabel-variabel terkait kebiasaan gaya hidup individu seperti pola makan, aktivitas fisik, merokok, dan aktivitas begadang.



Gambar 1. Import Data



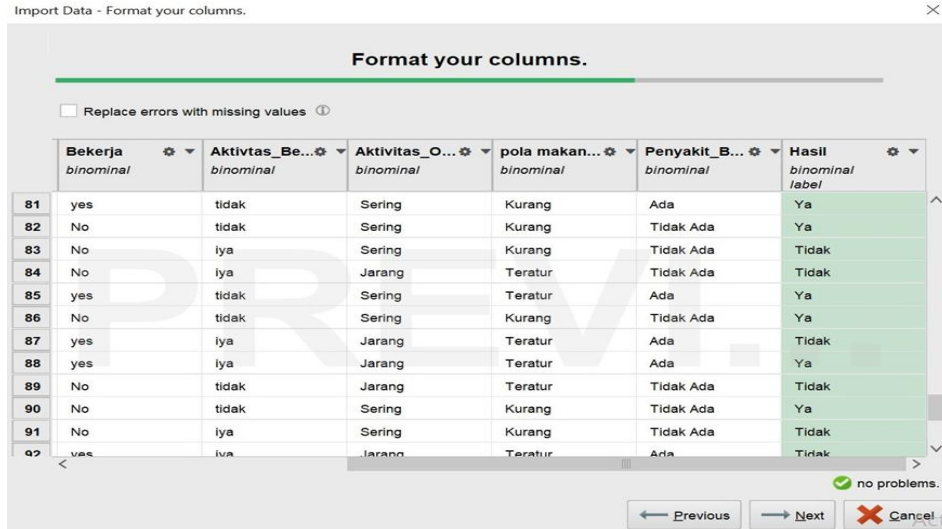
Gambar 2. Meletakkan Dataset Pada Panel Operator

2. Pemberian Label Data

Selanjutnya, data diberi label sesuai dengan kategori hasil kesehatan. Dalam penelitian ini, label diberikan dalam dua kategori:

- “Ya” untuk responden dengan kebiasaan gaya hidup yang tidak baik.
- “Tidak” untuk responden yang memiliki gaya hidup yang baik.

Perbandingan Algoritma Decision Tree dan Deep Learning dalam Prediksi Masalah Kesehatan berdasarkan Kebiasaan Gaya Hidup



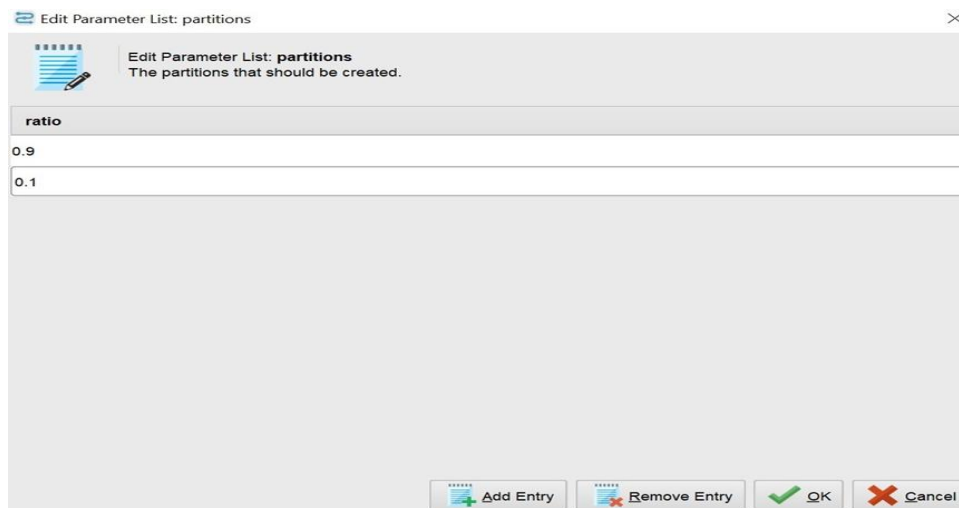
Gambar 3. Pemberian Label

3. Pembagian Data (Data Splitting) dan Pembagian Parameter

Dataset kemudian dibagi menjadi dua bagian, yaitu training set dan testing set. Pembagian ini dilakukan untuk melatih dan menguji model, dengan perbandingan umum adalah 90% data untuk pelatihan dan 10% untuk pengujian. Operator yang digunakan dalam RapidMiner adalah Split Data atau Cross Validation untuk membagi data secara acak.



Gambar 4. Split Data

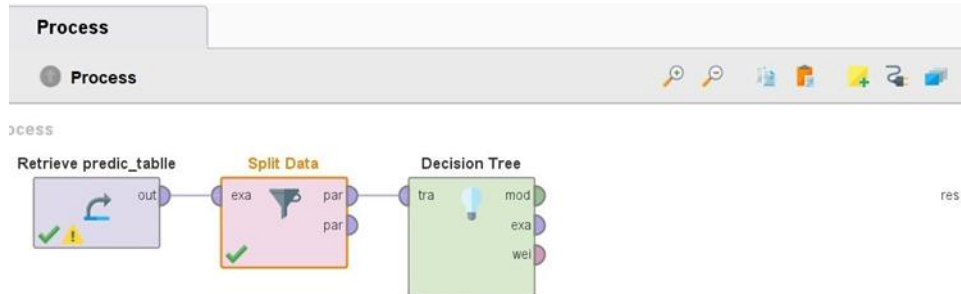


Gambar 5. Pembagian Parameter

4. Pelatihan Model (Model Training)

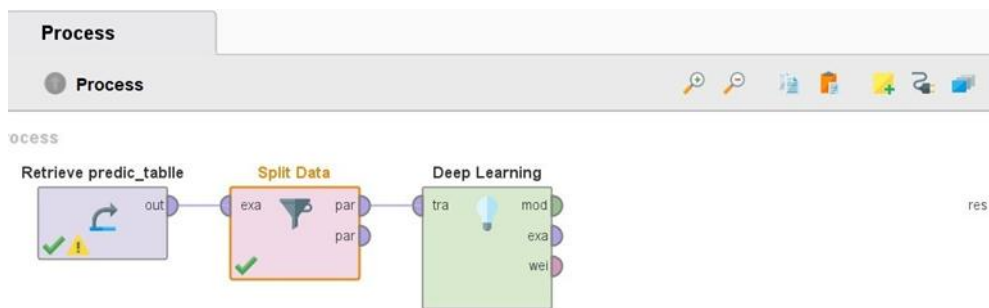
Pada tahap ini, dilakukan pelatihan model menggunakan dua algoritma:

- a. Decision Tree: Operator yang digunakan adalah Decision Tree, yang membangun model berbasis pohon keputusan untuk prediksi.



Gambar 6. Training Decision Tree

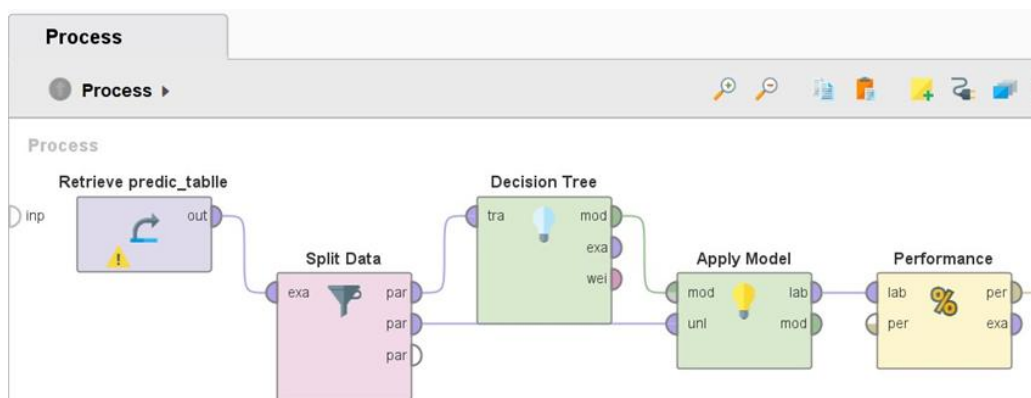
- b. Deep Learning: Menggunakan Operator Deep Learning untuk membangun model dengan jaringan saraf tiruan, yang mampu menangkap hubungan kompleks antar variabel.



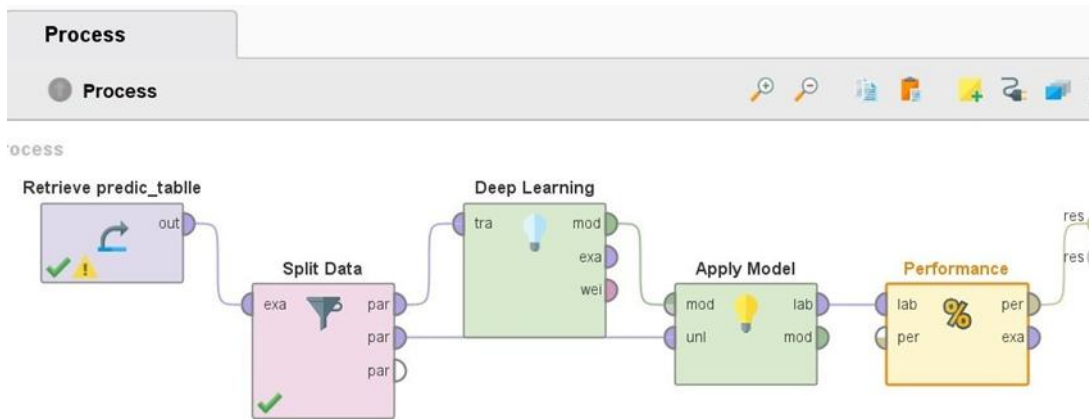
Gambar 7. Training Deep Learning

5. Evaluasi Model (Model Evaluation)

Setelah model dilatih, tahap evaluasi dilakukan untuk menguji kinerja model pada testing set. Operator Performance (Classification) digunakan untuk mengevaluasi model berdasarkan metrik seperti akurasi, presisi, recall, dan AUC. Operator ini juga memungkinkan perbandingan langsung antara dua model, yaitu Decision Tree dan Deep Learning.



Gambar 8. Evaluasi Decision Tree



Gambar 9. Evaluasi Deep Learning

6. Interpretasi Hasil

Hasil dari model-model yang dilatih kemudian dianalisis dan dibandingkan. Perbandingan performa antara algoritma Decision Tree dan Deep Learning dilakukan berdasarkan metrik yang telah diperoleh. Dalam penelitian ini, Deep Learning terbukti lebih unggul dengan akurasi yang lebih tinggi.

Populasi dan Sampel

Dataset ini berisi data publik yang mencakup individu dengan karakteristik gaya hidup berbeda. Sampel penelitian terdiri dari data individu dengan variasi usia yang dikategorikan sebagai ‘Muda’ atau ‘Tua’, serta variasi jenis kelamin dan kebiasaan gaya hidup lainnya. Sampel dipilih berdasarkan ketersediaan data mengenai keragaman variable gaya hidup lainnya dan berkaitan dengan hasil kesehatan. Data sampel yang menyajikan contoh dari beberapa item variabel terdapat pada tabel 1 berikut.

Tabel 1. Variabel Data Kebiasaan Hidup

Usia	Jenis Kelamin	Merokok	Bekerja	Aktivitas Begadang	Aktivitas Olahraga	Pola Makan	Penyakit Bawaan	Hasil
Muda	Pria	Aktif	No	Iya	Jarang	Teratur	Tidak Ada	Tidak
Muda	Wanita	Pasif	Yes	Tidak	Sering	Kurang	Ada	Ya
Muda	Wanita	Pasif	Yes	Tidak	Sering	Kurang	Ada	Ya
Muda	Pria	Aktif	No	Iya	Sering	Kurang	Tidak Ada	Ya
Muda	Pria	Aktif	No	Iya	Jarang	Teratur	Tidak Ada	Tidak
Tua	Wanita	Aktif	Yes	Iya	Jarang	Teratur	Ada	Tidak
Muda	Wanita	Pasif	Yes	Iya	Sering	Kurang	Ada	Ya
Tua	Wanita	Aktif	Yes	Iya	Jarang	Teratur	Ada	Tidak
Tua	Pria	Pasif	No	Iya	Sering	Teratur	Tidak Ada	Tidak
Tua	Pria	Aktif	Yes	Tidak	Jarang	Teratur	Ada	Tidak

Berdasarkan tabel 1 diatas, dapat dilihat bahwa beberapa poin utama terakit kebiasaan hidup dan hasil kesehatan adalah :

1. Usia dan Jenis Kelamin: Tidak menunjukkan pola yang jelas terhadap hasil kesehatan.
2. Merokok: Hasil kesehatan bervariasi antara perokok aktif dan perokok pasif, tanpa pola yang konsisten.
3. Aktivitas Begadang dan Aktivitas Olahraga: Tidak selalu berkorelasi langsung dengan hasil kesehatan.
4. Pola Makan dan Penyakit Bawaan: Pola makan kurang teratur dan adanya penyakit bawaan cenderung terkait dengan masalah kesehatan, tapi tidak selalu.

Metode Analisis Data

Data yang telah didapat akan dianalisis menggunakan dua algoritma utama, yaitu:

1. Decision Tree, adalah algoritma yang digunakan untuk klasifikasi dan regresi dengan memecah data ke dalam subset berdasarkan atribut tertentu. Pemecahan ini dilakukan secara berulang hingga mencapai hasil yang optimal. Berikut adalah rumus-rumus utama yang digunakan dalam *Decision Tree*.

a. Entropy

$$E(S) = - \sum_{i=1}^n p_i \log_2 p_i$$

Dimana:

$E(S)$ adalah *entropy* dari dataset S ,

p_i adalah probabilitas kemunculan kelas ke- i dalam dataset,

n adalah jumlah kelas

b. Information Gain

$$IG(S, A) = E(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} E(S_v)$$

Dimana:

$IG(S, A)$ adalah *Information Gain* dari atribut A ,

$E(S)$ adalah *entropy* dari dataset sebelum pemisahan,

$\frac{|S_v|}{|S|}$ adalah proporsi data subset S_v yang dipisahkan oleh nilai v dari atribut A ,

$E(S_v)$ adalah *entropy* dari subset data S_v .

c. Gini index (Gini Impurity)

$$Gini(S) = 1 - \sum_{i=1}^n p_i^2$$

Dimana:

$Gini(S)$ adalah Gini Index dari dataset S ,

p_i adalah proporsi kelas ke- i dalam dataset.

d. Reduction in Variance (Untuk Regresi)

$$\text{Variance}(S) = \frac{1}{|S|} \sum_{i=1}^n (x_i - \bar{x})^2$$

Dimana:

\bar{x} adalah rata-rata dari nilai dalam dataset S ,

x_i adalah setiap nilai dalam dataset (Maimon & Rokach, 2014).

2. Deep Learning, *Deep Learning* akan diterapkan menggunakan jaringan saraf tiruan (*Artificial Neural Networks*), yang terdiri dari beberapa lapisan tersembunyi. Algoritma ini mampu mendeteksi pola non-linear yang lebih kompleks dalam data kebiasaan gaya hidup. Pengujian akurasi model juga dilakukan dengan *cross-validation*.

Perhitungan Neuron pada Jaringan Saraf Tiruan

$$z = \sum_{i=1}^n w_i x_i + b$$

Dimana:

z adalah nilai sebelum fungsi aktivasi diterapkan,

w_i adalah bobot untuk setiap input x_i ,

x_i adalah nilai input,

b adalah bias yang ditambahkan

- a. Fungsi Aktivasi

Sigmoid

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

ReLU (Rectified Linear Unit)

$$f(z) = \max(0, z)$$

Fungsi Biaya (*Loss Function*)

$$L = -\frac{1}{n} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Dimana:

L adalah nilai fungsi biaya,

y_i adalah nilai aktual,

\hat{y}_i adalah nilai prediksi.

- b. *Backpropagation* dan Pembaruan Bobot

Gradien dari fungsi biaya terhadap bobot:

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial w}$$

Setelah gradien dihitung, bobot diperbarui sebagai berikut:

$$w = w - \eta \cdot \frac{\partial L}{\partial w}$$

Dimana:

w adalah bobot yang diperbarui,

η adalah laju pembelajaran (*learning rate*),

$\frac{\partial L}{\partial w}$ adalah gradien dari fungsi biaya terhadap bobot.

- c. *Softmax Function* (Untuk Multi-Kelas)

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

Dimana:

z_i adalah skor dari kelas ke- i ,

K adalah jumlah kelas (Yudistira, 2024).

Penelitian ini dilakukan dengan mempertimbangkan standar etika yang ketat. Data publik yang digunakan bersifat anonim, sehingga identitas individu tidak dapat diidentifikasi, dan semua data diakses serta digunakan hanya untuk tujuan penelitian. Dalam hal validitas dan reliabilitas, instrumen data publik telah diuji melalui uji validitas content validity untuk memastikan keakuratan isi. Sementara itu, reliabilitas data diuji menggunakan Cronbach's Alpha untuk memastikan konsistensi hasil. Validitas model prediksi diuji lebih lanjut melalui teknik cross-validation, dan reliabilitas model dipastikan dengan melakukan pengujian pada berbagai subset data, sehingga hasil penelitian ini memiliki dasar yang kuat dalam hal etika, validitas, dan reliabilitas.

HASIL DAN PEMBAHASAN

Hasil Temuan

Setelah dilakukan analisis terhadap data kebiasaan gaya hidup menggunakan algoritma Decision Tree dan Deep Learning, beberapa temuan utama yang diperoleh dapat disajikan dalam bentuk tabel sebagai berikut:

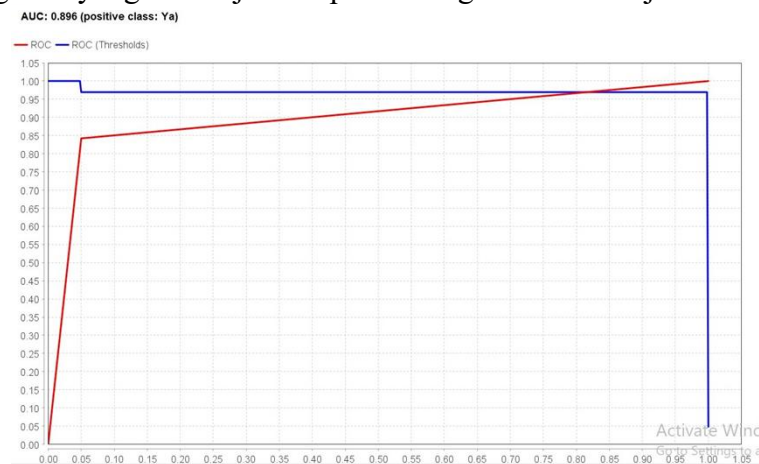
Tabel 2. Perbandingan Kinerja Algoritma Decision Tree dan Deep Learning

Formula	Decision Tree	Deep Learning
Akurasi	89.74%	92.31%
Presisi	94.12%	100.00%
Recall	84.21%	84.21%
AUC	0.896	0.933

Dari tabel di atas, terlihat bahwa algoritma Deep Learning menghasilkan kinerja yang lebih baik dibandingkan Decision Tree dalam hal akurasi, presisi, recall, dan AUC. Akurasi Deep Learning mencapai 92.31%, sementara Decision Tree hanya 89.74%. Begitu pula dengan presisi, recall, dan AUC Deep Learning unggul dengan akurasi yang lebih tinggi.

Selain itu, dari segi kemampuan memprediksi hasil kesehatan berdasarkan kebiasaan gaya hidup, algoritma Deep Learning juga lebih unggul dalam mengenali pola-pola kompleks dalam data, seperti interaksi antara kebiasaan olahraga, pola makan, dan kebiasaan merokok.

Berikut adalah grafik yang menunjukkan perbandingan hasil kinerja kedua algoritma:



Gambar 10. AUC Decision Tree

Perbandingan Algoritma Decision Tree dan Deep Learning dalam Prediksi Masalah Kesehatan berdasarkan Kebiasaan Gaya Hidup



Gambar 11. AUC Deep Learning

Pembahasan

1. Interpretasi Hasil

Hasil temuan menunjukkan bahwa algoritma Deep Learning lebih efektif dalam memprediksi masalah kesehatan berdasarkan kebiasaan gaya hidup dibandingkan dengan Decision Tree. Hal ini disebabkan oleh kemampuan Deep Learning dalam menangani kompleksitas data, terutama ketika ada hubungan non-linear antara variabel yang sulit dijelaskan oleh algoritma yang lebih sederhana seperti Decision Tree. Sebagai contoh, pola makan dan olahraga mungkin memiliki interaksi yang lebih kompleks dengan penyakit bawaan, yang berhasil diidentifikasi oleh Deep Learning tetapi tidak oleh Decision Tree.

Kinerja yang baik dari Deep Learning dibuktikan dengan penelitian tentang penyakit kanker payudara yang dianggap memiliki tingkat akurasi dan prediksi yang tinggi (Ardiansyah et al., 2023). Ini juga mengindikasikan bahwa algoritma yang lebih canggih dapat menangani variasi pola kebiasaan hidup lebih baik.

2. Implikasi Hasil

Penelitian ini memiliki implikasi penting dalam pengembangan sistem prediksi kesehatan berbasis data kebiasaan gaya hidup. Dengan menggunakan algoritma Deep Learning, sistem prediksi dapat lebih akurat dan sensitif terhadap perubahan pola gaya hidup yang berdampak pada kesehatan. Ini berguna dalam pengembangan alat bantu medis atau aplikasi kesehatan yang memantau dan memberikan rekomendasi kepada pengguna berdasarkan data gaya hidup mereka.

Selain itu, algoritma Deep Learning dapat dimanfaatkan oleh praktisi kesehatan untuk mengenali pola risiko kesehatan yang tidak terdeteksi oleh pendekatan tradisional, sehingga dapat membantu dalam penanganan preventif.

3. Keterbatasan Penelitian

Meskipun penelitian ini memberikan hasil yang signifikan, terdapat beberapa keterbatasan yang perlu diperhatikan. Pertama, sampel data yang digunakan masih terbatas pada masyarakat Indonesia dengan rentang usia tertentu. Ini dapat membatasi generalisasi hasil ke populasi yang lebih luas atau dengan karakteristik yang berbeda. Kedua, algoritma yang digunakan belum mencakup teknik lain yang mungkin lebih efektif dalam kondisi tertentu, seperti Random Forest atau Support Vector Machines.

Untuk penelitian mendatang, disarankan untuk meningkatkan variasi sampel dan mencoba algoritma lain untuk membandingkan kinerja yang lebih komprehensif. Selain itu, penggunaan

lebih banyak fitur dalam data, seperti informasi genetik atau riwayat medis, dapat meningkatkan akurasi model prediksi.

Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma Deep Learning memiliki performa yang lebih baik dalam memprediksi masalah kesehatan berdasarkan kebiasaan gaya hidup dibandingkan Decision Tree. Temuan ini penting untuk pengembangan alat prediksi kesehatan berbasis data dan memberikan kontribusi dalam upaya pencegahan penyakit melalui pengenalan pola risiko dari kebiasaan gaya hidup.

KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma Deep Learning lebih unggul dibandingkan Decision Tree dalam memprediksi risiko masalah kesehatan berdasarkan kebiasaan gaya hidup, dengan akurasi yang lebih tinggi dan kemampuan yang lebih baik dalam menangani data kompleks. Temuan ini memperkuat peran machine learning dalam sistem prediksi kesehatan preventif, yang dapat diaplikasikan dalam pengembangan aplikasi atau alat bantu medis untuk memberikan rekomendasi gaya hidup sehat. Penelitian ini memberikan kontribusi penting bagi bidang data science dan kesehatan, namun terbatas pada ukuran sampel yang tidak terlalu besar dan variabel yang terbatas. Untuk penelitian selanjutnya, disarankan memperluas sampel dan variabel data, serta mengeksplorasi algoritma lain untuk meningkatkan akurasi prediksi. Rekomendasi praktis meliputi penerapan hasil dalam sistem kesehatan berbasis data gaya hidup, yang dapat membantu masyarakat meningkatkan kualitas kesehatan secara preventif.

DAFTAR PUSTAKA

- Adistie, F., Lumbantobing, V. B. M., & Maryam, N. N. A. (2018). Pemberdayaan kader kesehatan dalam deteksi dini stunting dan stimulasi tumbuh kembang pada balita. *Media Karya Kesehatan, 1*(2), 173–184.
- Ardiansyah, G. R., Musayyanah, M., Aqvirandy, W., Farady, M. D., Cahya, M. N., & Hadiono, T. (2023). Deteksi kanker payudara menggunakan artificial neural network pada deep learning. *INFOTECH: Jurnal Informatika & Teknologi, 4*(2), 259–269.
- Ataey, A., Jafarvand, E., Adham, D., & Moradi-Asl, E. (2020). The relationship between obesity, overweight, and the human development index in world health organization eastern mediterranean region countries. *Journal of Preventive Medicine and Public Health, 53*(2), 98.
- Badidi, E. (2023). Edge AI for early detection of chronic diseases and the spread of infectious diseases: opportunities, challenges, and future directions. *Future Internet, 15*(11), 370.
- Carney, T. J., Wiltz, J. L., Davis, K., Briss, P. A., & Hacker, K. (2023). Peer Reviewed: Advancing Chronic Disease Practice Through the CDC Data Modernization Initiative. *Preventing Chronic Disease, 20*.
- Collins, J. L., Koplan, J. P., & Marks, J. S. (2009). Chronic disease prevention and control: coming of age at the Centers for Disease Control and Prevention. *Preventing Chronic Disease, 6*(3).
- Gilmore, G. D. (2024). Chronic Disease Epidemiology, Prevention, and Control. *American Journal of Preventive Medicine, 67*(4), 637–639.
- Hermawan, A., & Hananto, M. (2020). Faktor sosiodemografi dan perilaku pencegahan gigitan nyamuk terhadap perilaku pemberantasan sarang nyamuk di Indonesia: analisis lanjut data riskesdas 2018. *Jurnal Ekologi Kesehatan, 19*(2), 101–111.
- Maimon, O. Z., & Rokach, L. (2014). *Data mining with decision trees: theory and applications* (Vol. 81). World scientific.
- Sari, F. (2023). Penerapan Deep Learning Dalam Kesehatan Digital: Memprediksi Diagnosis

- Penyakit Dengan Akurasi Tinggi. *Jurnal Teknologi Pintar*, 3(12).
- Tamalumu, A., Kapantow, N. H., & Sanggelorang, Y. (2021). Gambaran Penerapan Pedoman Umum Gizi Seimbang Mahasiswa Semester II Fakultas Kesehatan Masyarakat Universitas Sam Ratulangi di Masa Pandemi Covid-19. *KESMAS*, 10(2).
- Telaumbanua, F. D., Hulu, P., Nadeak, T. Z., Lumbantong, R. R., & Dharma, A. (2019). Penggunaan Machine Learning Di Bidang Kesehatan. *Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP)*, 2(2), 391–399.
- Utiahman, S. A., & Pratama, A. M. M. (2024). Analisis Perbandingan KNN, SVM, Decision Tree dan Regresi Logistik Untuk Klasifikasi Obesitas Multi Kelas. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(6), 3137–3146.
- Wicaksono, E. R. T., Mulya, A. P., & Purnama, D. (2023). Pendidikan Kesehatan Gaya Hidup Sehat dan Gizi Seimbang Dalam Mengoptimalkan Kesehatan Remaja: Studi Kasus. *MAHESA: Malahayati Health Student Journal*, 3(8), 2561–2574.
- Yudistira, E. N., & Kom, S. (2024). *Deep Learning: Teori, Contoh Perhitungan, dan Implementasi*. Deepublish.