

KOMPARASI ALGORITMA *BOOSTING* UNTUK PREDIKSI GANGGUAN TIDUR

Ade Bagus Mawardi¹, Risqy Siwi Pradini^{2*}, M. Syauqi Haris³

^{1,2,3}Informatika, Fakultas Sains dan Teknologi, Institut Teknologi, Sains, dan Kesehatan RS. DR. Soepraoen Kesdam V/BRW

Keywords:

Komparasi, Algoritma *Boosting*, Gangguan tidur, Prediksi

Correspondent Email:

risqypradini@itsk-soepraoen.ac.id

Abstrak. Gangguan tidur merupakan salah satu permasalahan kesehatan yang dapat berdampak pada kualitas hidup seseorang. Dalam upaya meningkatkan akurasi prediksi gangguan tidur, teknologi kecerdasan buatan telah banyak dimanfaatkan, khususnya melalui pendekatan algoritma *machine learning*. Penelitian ini bertujuan untuk melakukan komparasi terhadap lima algoritma *boosting*, yaitu *AdaBoost*, *CatBoost*, *LightGBM*, *Gradient Boosting*, dan *XGBoost* menggunakan *dataset Sleep Health and Lifestyle*. Adapun tahap penelitian yang dilakukan meliputi pengumpulan data, prapemrosesan data, normalisasi, serta evaluasi model. Berdasarkan hasil evaluasi, algoritma *CatBoost* menunjukkan performa paling unggul dibandingkan dengan algoritma lainnya. Hasil evaluasi menunjukkan bahwa algoritma *CatBoost* memberikan performa terbaik dengan akurasi sebesar 97,37%, presisi 96,29%, *recall* 95,83%, dan *F1-score* 95,82%. Hasil analisis menunjukkan bahwa keunggulan *CatBoost* berasal dari kemampuannya dalam menangani fitur kategorikal secara langsung tanpa memerlukan *encoding* tambahan, serta kemampuannya dalam mengurangi *overfitting* dibandingkan dengan metode *boosting* lainnya. Temuan ini menunjukkan bahwa model prediksi berbasis *boosting* khususnya *CatBoost* dapat dijadikan alat bantu yang efektif dalam deteksi gangguan tidur secara lebih akurat.



JITET is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Abstract. *Sleep disorders are one of the health issues that can significantly impact a person's quality of life. To improve the accuracy of sleep disorder prediction, artificial intelligence technology has been widely utilized, particularly through machine learning algorithms. This study aims to compare the performance of five boosting algorithms AdaBoost, CatBoost, LightGBM, Gradient Boosting, and XGBoost using the Sleep Health and Lifestyle dataset. The research stages include data collection, preprocessing, normalization, and model evaluation. Based on the evaluation results, the CatBoost algorithm demonstrated the best performance among all algorithms tested, achieving an accuracy of 97.37%, precision of 96.29%, recall of 95.83%, and an F1-score of 95.82%. The analysis suggests that CatBoost's advantage lies in its ability to handle categorical features directly without requiring additional encoding, as well as its effectiveness in reducing overfitting compared to other boosting methods. These findings indicate that boosting-based prediction*

models, particularly CatBoost, can serve as an effective tool for more accurate detection of sleep disorders.

1. PENDAHULUAN

Menjaga kondisi tidur yang baik adalah salah satu faktor terpenting dalam upaya memelihara sistem kesehatan tubuh manusia [1][2]. Tidur yang berkualitas terbukti memberikan pengaruh positif pada fisik, kognitif, dan emosional seseorang [3]. Sebaliknya, jika kualitas tidur terganggu maka akan menjadi masalah kesehatan yang signifikan, yang dapat menurunkan kualitas hidup dan kesejahteraan individu secara keseluruhan [4]. Faktor-faktor yang dapat mempengaruhi gangguan tidur meliputi variabel seperti durasi tidur dan intensitas stres yang dihadapi. Gangguan tidur dapat di klasifikasi menjadi beberapa jenis, seperti *insomnia*, *sleep apnea*, atau tidak adanya gangguan tidur [5]. *Insomnia* merupakan kondisi gangguan tidur yang dimana seseorang sulit untuk tidur. Sedangkan, *sleep apnea* merupakan kondisi medis yang ditandai dengan kesulitan pernapasan selama tidur, yang berpotensi mengurangi kualitas istirahat seseorang [6]. Melalui pemahaman yang lebih mendalam tentang gangguan tidur seperti *insomnia* dan *sleep apnea*, adanya teknologi terkini memberikan peluang inovasi dalam membangun metode prediksi yang akurat.

Salah satu terobosan terbaru yang memanfaatkan teknologi adalah penerapan *machine learning* di bidang kesehatan. Di antara berbagai pendekatan *machine learning*, algoritma *boosting* merupakan salah satu metode yang baik untuk melakukan prediksi [7]. Algoritma *boosting* merupakan metode yang menggabungkan beberapa model untuk meningkatkan performa dan akurasi [8]. Algoritma *boosting* memiliki berbagai jenis pendekatan. Salah satu pendekatan menggunakan algoritma *boosting* diantaranya adalah *AdaBoost*, *CatBoost*, *LightGBM*, *GradientBoost*, dan *XGBoost* [9][10].

Penelitian sebelumnya menunjukkan bahwa salah satu jenis algoritma *boosting* yaitu, *Gradient Boosting* memiliki kinerja unggul, terutama dalam hal keakuratan prediksi, sehingga menjadi alat yang efektif untuk

analisis gangguan tidur [11]. Selain itu, metode *AdaBoost* berhasil mencapai akurasi sebesar 90,1% dalam klasifikasi data [12]. Dalam evaluasi performa algoritma klasifikasi, *AdaBoost* juga menunjukkan kinerja yang lebih baik dibandingkan algoritma *K-Nearest Neighbors (K-NN)* [13]. Algoritma *boosting* lainnya, seperti *CatBoost*, juga memberikan hasil yang signifikan. Dalam sebuah penelitian, *CatBoost* dikomparasikan dengan algoritma lain seperti *Extra Trees*, *Support Vector Regression*, dan *Random Forest*, dan berhasil menunjukkan akurasi yang lebih baik [14]. *LightGBM* yang merupakan, algoritma *boosting* lainnya, juga menarik perhatian dengan kemampuannya mencapai akurasi lebih tinggi dibandingkan metode klasifikasi lainnya [15]. Selanjutnya, *XGBoost* juga mampu melampaui akurasi metode *Support Vector Machine (SVM)*, sehingga menjadikannya salah satu algoritma paling efektif untuk tugas klasifikasi [16]. Berdasarkan berbagai penelitian ini, maka algoritma *boosting* secara konsisten menunjukkan performa yang baik dalam mencapai akurasi yang baik.

Dengan demikian penelitian ini bertujuan untuk membandingkan beberapa algoritma *boosting*, yaitu *AdaBoost*, *CatBoost*, *LightGBM*, *GradientBoost*, dan *XGBoost*, dalam membangun model prediksi yang akurat untuk mendiagnosis gangguan tidur. Dengan menguji kinerja masing-masing algoritma, penelitian ini berfokus untuk membandingkan metode yang paling efektif dalam mengoptimalkan akurasi prediksi. Diharapkan hasil penelitian ini memberikan kontribusi signifikan dalam pengembangan algoritma prediksi gangguan tidur yang lebih tepat dan efisien.

2. TINJAUAN PUSTAKA

Algoritma pertama yang digunakan untuk komparasi menggunakan studi kasus prediksi *sleep disorder* adalah *XGBoost (Extreme Gradient Boosting)*. Algoritma ini menggunakan pohon keputusan sebagai model dasar dan menerapkan teknik *boosting* untuk meningkatkan kinerja model [17]. Algoritma

kedua yang digunakan ialah *CatBoost* (*Categorical Boosting*) merupakan algoritma *machine learning* berbasis *gradient boosting* yang memiliki kemampuan unik untuk menangani data kategorik secara efisien tanpa perlu melakukan proses *encoding* yang umum dilakukan pada kebanyakan algoritma lainnya [18].

Algoritma ketiga yang digunakan ialah *Adaboost* yang sering digunakan dengan *classifier* algoritma lainnya untuk meningkatkan performa klasifikasi [19]. *AdaBoost* adalah algoritma pembelajaran berulang yang melatih dan merakit beberapa pengklasifikasi lemah menjadi pengklasifikasi yang kuat untuk mencapai klasifikasi yang sangat akurat [20]. Algoritma keempat yang digunakan ialah *LightGBM* yang termasuk dalam algoritma *boosting* yang menggunakan pendekatan pohon keputusan [21]. Algoritma *LightGBM* banyak digunakan karena efisiensi dan kecepatannya untuk menangani kumpulan data yang cukup besar [22]. Algoritma kelima yang digunakan *Gradient Boosting* yang dilatih secara bertahap, di mana pada setiap iterasinya, model membangun pohon keputusan baru dengan mengacu pada nilai gradien negatif dari fungsi *loss* untuk meminimalkan kesalahan prediksi secara progresif [23].

3. METODE PENELITIAN

Seluruh proses penelitian dilaksanakan menggunakan *Jupyter Notebook* untuk keperluan pengolahan data, analisis, serta penerapan algoritma. Tahapan penelitian yang dilakukan divisualisasikan pada Gambar 1.



Gambar 1. Tahapan penelitian

Tahap awal penelitian dimulai dengan pengumpulan *dataset* yang relevan, yaitu *Sleep Health and Lifestyle Dataset* yang diunduh dari *Kaggle*. *Dataset* ini berisi 374 baris dan 12 kolom yang merepresentasikan berbagai fitur yang berkaitan dengan pola tidur dan 1 kelas target. Data yang diperoleh kemudian melalui tahap pra-pemrosesan data guna memastikan kualitas dan kesiapan data untuk analisis lebih lanjut [24]. Adapun fitur *dataset* yang digunakan bisa dilihat dari Tabel 1 berikut ini.

Tabel 1. Fitur *dataset*

Fitur	Tipe data	Rentang nilai	Ket
<i>Person ID</i>	Int64	1-374	Pengenal individu
<i>Gender</i>	Kateg orikal	<i>Male and Female</i>	Jenis kelamin
<i>Age</i>	Int64	27 – 59	Usia
<i>Occupation</i>	Kateg orikal	11 jenis pekerjaan	Jenis pekerjaan
<i>Sleep Duration</i>	Float64	5.8 – 8.5 jam	Jumlah jam tidur per hari
<i>Quality of Sleep</i>	Int64	4 – 9 (skala kualitas)	Penilaian subjektif terhadap kualitas tidur
<i>Physical Activity Level</i>	Int64	30 – 90 (indikator aktivitas)	Jumlah menit aktivitas fisik

<i>Strees Level</i>	Int64	3 – 8 (skala stres)	Peringkat subjektif dari tingkat stres yang dialami
<i>BMI Category</i>	Kateg orikal	<i>Overweight, Normal, Obese, Normal Weight</i>	Kategori BMI
<i>Blood Pressure</i>	Kateg orikal	126/83, 125/80,140/90	Tekanan darah
<i>Heart Rate</i>	Int64	65 – 86 detak/menit	Denyut jantung
<i>Daily Steps</i>	Int64	3000 – 10000 langkah	Jumlah langkah
<i>Sleep Disorder</i>	Kateg orikal	<i>None, Sleep Apnea, Insomnia</i>	Gangguan tidur (<i>None, Sleep Apnea, Insomnia</i>)

Pada pengecekan *missing value*, data diperiksa untuk mengidentifikasi adanya nilai kosong pada setiap kolom. Tujuannya adalah untuk memastikan tidak ada data yang hilang atau data yang kurang atau tidak lengkap [25]. Dengan demikian, kualitas data dapat terjaga sehingga model prediksi yang dibangun dapat menghasilkan hasil yang lebih akurat.

Pada pengecekan data duplikat, *dataset* dicek untuk memastikan tidak terdapat data duplikat. Adanya data duplikat menyebabkan bias pada proses pelatihan model. Selain tidak juga meningkatkan waktu komputasi, menghabiskan memori, dan dapat menyebabkan ketidakseimbangan dalam data [26]. Pada penghapusan fitur tidak relevan, fitur yang tidak berpengaruh signifikan terhadap prediksi dihapus dari *dataset*.

Pada data *binning*, beberapa variabel numerik dikategorikan ke dalam beberapa kelompok. Data *binning* merupakan metode pengorganisasian data yang melibatkan pengelompokan nilai-nilai numerik ke dalam kategori seperti kelas rendah sedang dan tinggi [27]. Salah satu contoh penerapan data *binning* pada penelitian ini adalah tekanan darah yang dikelompokkan dalam kategori normal, *elevelted*, dan *high*.

Tahap berikutnya pada pra-pemrosesan data adalah normalisasi data. Hal ini perlu dilakukan karena variabel dalam data sering kali memiliki rentang nilai yang berbeda secara signifikan

[28]. Teknik normalisasi yang digunakan adalah *MinMax Scaler*, yang merubah rentang nilai fitur ke dalam interval [0, 1].

Setelah menyelesaikan tahap pra-pemrosesan data, langkah selanjutnya adalah melakukan pembagian *dataset*. Proses ini bertujuan untuk membagi data menjadi dua, data *testing* dan data *training* [29]. Sebanyak 90% baris *dataset* dipilih secara acak untuk pelatihan dan 10% sisanya digunakan untuk pengujian.

Tahap selanjutnya adalah pembangunan dan pelatihan model menggunakan lima algoritma *boosting* yang dikomparasi dalam penelitian ini, yaitu *AdaBoost*, *CatBoost*, *LightGBM*, *Gradient Boosting*, dan *XGBoost*. Pada tahap ini, setiap model dilatih menggunakan data *training* yang telah disiapkan sebelumnya. Adapun *pseudocode* masing-masing algoritma yang dikomparasikan ditunjukkan oleh Gambar 2 sampai dengan Gambar 6.

```

Input:
 $F_{normalized} (X_1^{norm}, \dots, X_n^{norm})$ ; set of financial ratios

Output:  $F_{optimal}$  : Set of selected financial ratios
Step 1: Load the normalized financial ratios
Step 2: Create an empty dictionary  $E$  to save the scores of financial ratios
Step 3: Instantiate an Extreme Gradient Boosting Classifier as EGB
Step 4: fit EGB
Step 4: Generate FIs
Step 5: Determine the FI threshold  $FI_h$ 
Step 6:
for  $t$  from  $F_{normalized}$  do
  if  $(FI(f^t) \geq FI_h)$  then
    add  $FI(f^t)$  into  $E$ 
  end if
end for
Step 5: Utilize the scores in  $E$  to produce  $F_{optimal}$ 
    
```

Gambar 2. *Pseudocode XGBoost* [30]

```

Inputs:  $\{(X, Y)\}_n^s, L, a, L, s, Mode$ 
 $\sigma_r \leftarrow$  random permutation of  $\{1, n\}, r = 0, \dots, s$ ;
 $M_0(i) \leftarrow 0$  for  $i = 0, \dots, n$ ;
If  $Mode = Plain$  then
   $M_r(i) \leftarrow 0$  for  $r = 1, \dots, s, i: \sigma_r(i) \leq 2^{r+1}$ ;
If  $Mode = Ordered$  then
  For  $j \leftarrow 1$  to  $\lceil \log 2n \rceil$  do
     $M_{r,j}(i) \leftarrow 0$  for  $r = 1, \dots, s, i = 1, \dots, 2^{j+1}$ 
  For  $t \leftarrow 1$  to  $L$  do
     $T_t, \{M_r\}_{r=1}^s \leftarrow Tree(\{M_r\}_{r=1}^s, \{(x, y)\}_{n=1}^n, a, L, \{\sigma_r\}_{r=1}^s, Mode)$ ;
     $Leaf_0^t(i) \leftarrow GetLeaf(x_i, T_t, \sigma_0)$  for  $i = 1, \dots, n$ ;
     $g_0 \leftarrow Gradient(L, M_0, y)$ ;
    Foreach  $Leaf^j$  in  $T_t$  do
       $b_j^t \leftarrow -avg(g_0(i) \text{ for } i: Leaf_0^t(i) = j)$ ;
     $M_0(i) \leftarrow M_0(i) + ab_{Leaf_0^t(i)}^t$  for  $i = 1, \dots, n$ 
  Return:  $F(x) = \sum_{t=1}^L \sum_{i=1}^n ab_{Leaf^t(x, T_t, Mode)}^t$ 
    
```

Gambar 3. *Pseudocode Catboost* [31]

Algorithm 1. AdaBoost.

Input: A set of labeled training instances:
 $S = \{(x_n, y_n), n = 1, 2, \dots, N\}$, where $y_n \in \{-1, 1\}$, maximum number of base classifiers T .
Initialization: $w_n^{(1)} = 1/N, n = 1, 2, \dots, N; t = 1$.
while $t \leq T$ **do**
 Resampling a training subset $TR_t = \{(x_n, y_n), n = 1, 2, \dots, N\}$ with replacement;
 Train the TR_t using a base learning algorithm to obtain weak learner h_t ;
 Calculate the weight error of training instances
 $\epsilon_t = \sum_{n=1}^N w_n^{(t)} I(h_t(x_n) \neq y_n)$;
if $\epsilon_t > 0.5$ **or** $\epsilon_t = 0$ **then**
 Generate uniformly distributed weights $w_n^{(t)} = 1/N, n = 1, 2, \dots, N$, continue;
 Calculate $\alpha_t = \frac{1}{2} \ln(\frac{1-\epsilon_t}{\epsilon_t})$;
 Update the weights of training instances with $w_n^{(t+1)} = w_n^{(t)} \exp(-\alpha_t h_t(x_n) y_n)$;
 Normalization : $w_n^{(t+1)} = w_n^{(t+1)} / \sum_{n=1}^N w_n^{(t+1)}$;
 $t = t + 1$;
 Final classifier h_T is obtained by weighted majority voting for a testing instance x_n^* : $h_T(x_n^*) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x_n^*))$.

Gambar 4. Pseudocode Adaboost [32]

Input: Traffic-level data T , Packet-level data P , Model parameters $\theta_{LGB}, \theta_{MN}$
Output: Detection results $R = \{R_i\}_{i=1}^N$
foreach $T_i \in T$ **do**
 Phase 1: Traffic-Level Anomaly Detection
 Compute $P_{LGB} = \sigma(\sum_{j=1}^M w_j \cdot g(T_i; \theta_j))$
if $P_{LGB} < T_{LGB}$ **then**
 Assign $R_i \leftarrow \text{Normal}$
 Continue
 Phase 2: Packet-Level Fine-Grained Classification
 Extract P_i and perform depthwise separable convolution:

$$Z_{i,j,k} = \sum_{m=1}^M X_{i+s_1,j+s_2,m} \cdot W_{k,m}$$

 Perform pointwise convolution and compute:

$$P_{MN} = \text{Softmax}(Y)$$

 Assign $R_i \leftarrow \arg \max_k P_{MN,k}$
return Detection results $R = \{R_i\}_{i=1}^N$.

Gambar 5. Pseudocode LightGBM

Algorithm 2. The GB pseudocode.

1 **Input:** $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}, L(y, O(x))$
 2 **Begin**
 3 Initialize: $O_0(x) = \underset{w}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, w)$
 4 **for** $m = 1 : M$
 5 $r_{im} = -\frac{\partial L(y_i, O(x_i))}{\partial O(x_i)}$
 6 Train weak learner $C_m(x)$ on training data.
 7 Calculate $w : w_m = \underset{w}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, O_{m-1}(x_i) + w C_m(x_i))$
 8 Update: $O_m(x) = O_{m-1}(x) + w_m C_m(x)$
 9 **End for**
 10 **End**
 11 **Output:** $O_m(x)$

Gambar 6. Pseudoce Gradient Boosting [33]

Langkah terakhir adalah evaluasi kinerja model dengan menggunakan data *testing*. Berbagai metrik evaluasi digunakan untuk menilai performa masing-masing algoritma, di antaranya *akurasi*, *presisi*, *recall*, dan *F1-score*. Presisi adalah rasio jumlah data prediksi yang benar terhadap jumlah data yang diprediksi positif, dihitung dengan persamaan (1) [34]:

$$Recall \frac{tp}{tp+fp} \quad (1)$$

Recall, yakni rasio jumlah data prediksi

yang benar terhadap jumlah data yang sebenarnya positif, dihitung dengan persamaan (2) [34]:

$$Recall \frac{tp}{tp+fn} \quad (2)$$

Akurasi, yakni rasio jumlah data prediksi yang benar terhadap jumlah total data, dihitung dengan persamaan (3) [34]:

$$Accuracy \frac{TP+TN}{TP+FP+FN+TN} \quad (3)$$

F1-score adalah kombinasi dari presisi dan recall, dihitung dengan persamaan (4) [34]:

$$F1 \text{ Score} = 2x \frac{recall \times precision}{recall+precision} \quad (4)$$

4. HASIL DAN PEMBAHASAN

4.1 Pengecekan Missing Value

Hasil dari pengecekan *missing value* yang ditunjukkan oleh Tabel 2. Hasilnya tidak ditemukannya nilai yang hilang pada semua kolom yang diamati. Hal ini menunjukkan bahwa *dataset* yang digunakan memiliki kualitas data yang baik dan lengkap.

Tabel 2. Hasil Pengecekan *missing values*

Fitur	Jumlah
Person ID	0
Gender	0
Age	0
Occupation	0
Sleep Duration	0
Quality of Sleep	0
Physical Activity Level	0
Stress Level	0
BMI Category	0
Blood Pressure	0
Heart Rate	0
Daily Steps	0

4.2. Pengecekan Data Duplikat

Selanjutnya, dilakukan identifikasi duplikasi data untuk menghindari bias dalam model. Berdasarkan hasil pengecekan menggunakan fungsi `df.duplicated().sum()`, tidak ditemukan adanya data duplikat dalam *dataset*. Hal ini ditunjukkan dengan nilai 0 yang dihasilkan, mengindikasikan bahwa setiap baris data dalam *dataset* adalah unik.

4.3 Pengecekan Fitur yang Tidak Relevan

Terdapat 12 fitur dan 1 kelas target yang digunakan sebagaimana ditunjukkan oleh kolom fitur pada Tabel 1. Tidak semua fitur dalam *dataset* memiliki dampak yang signifikan terhadap target prediksi *sleep disorder*. Pada penelitian ini, fitur *Person Id* dihapus karena tidak memberikan kontribusi berarti terhadap model prediksi, sedangkan semua fitur lainnya tetap digunakan.

4.4. Data Binning

Dalam upaya meningkatkan interpretasi model dan menyederhanakan proses analisis, beberapa variabel numerik dikelompokkan ke dalam kategori tertentu menggunakan metode *binning*. Metode ini sangat berguna dalam menangani variabel numerik dengan rentang nilai yang luas dan kompleks. Tabel 3 menunjukkan sebelum dan sesudah dilakukannya data *binning*.

Tabel 3. Sebelum dan sesudah data *binning*

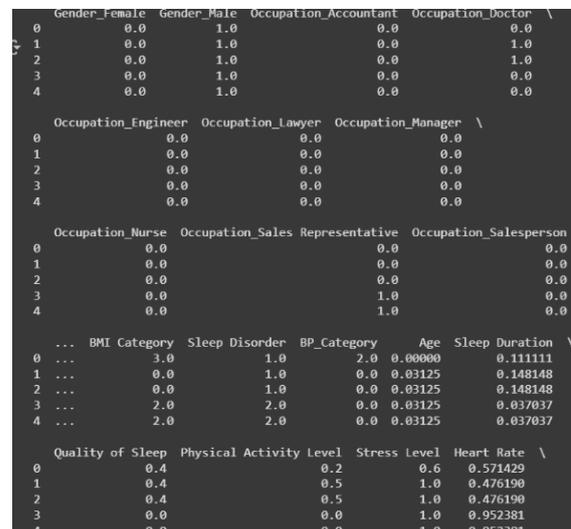
No.	Fitur	Sebelum	Sesudah
1	Blood Pressure	126/83	Prehypertension
2	Blood Pressure	125/80	Hypertension

Berdasarkan data di Tabel 3, salah satu fitur yang dikelompokkan adalah *blood pressure*. Fitur ini awalnya direpresentasikan dalam format numerik dengan dua nilai, yaitu sistolik dan diastolik. Untuk menyederhanakan analisis, tekanan darah diklasifikasikan ke dalam tiga kategori utama, yaitu Normal, *Elevated (Prehypertension)*, dan *High (Hypertension)*. Pengelompokan ini dilakukan berdasarkan nilai sistolik dan diastolik menggunakan fungsi *categorize_bp* yang telah didefinisikan dalam kode. Hasil dari pengelompokan ini kemudian disimpan dalam variabel baru yang diberi nama *BP Category*, yang memungkinkan analisis lebih lanjut terhadap hubungan tekanan darah dengan variabel lain dalam penelitian.

4.5. Normalisasi Data

Dalam penelitian ini, dilakukan normalisasi data untuk memastikan bahwa setiap fitur dalam *dataset* berada dalam skala yang seragam. Hal ini diperlukan karena *dataset* memiliki skala yang berbeda untuk setiap fitur, yang dapat menyebabkan model *machine learning* lebih condong pada fitur dengan nilai yang lebih besar, sehingga mempengaruhi

kinerja algoritma secara keseluruhan. Metode yang digunakan dalam proses normalisasi adalah *MinMax Scaler*. Hasil penerapan normalisasi data ada pada Gambar 7.



Gambar 7. Hasil normalisasi data menggunakan *Minmax Scaler*

4.6. Evaluasi Model

Pada penelitian ini, lima algoritma *boosting*, yaitu *AdaBoost*, *Gradient Boosting*, *CatBoost*, *LightGBM*, dan *XGBoost* digunakan untuk membangun model prediksi gangguan tidur. Setiap model dilatih menggunakan data *training* dan dinilai menggunakan data *testing* dengan metrik evaluasi berupa *akurasi*, *presisi*, *recall*, dan *F1-score*. Tabel 4 menyajikan hasil evaluasi performa kelima algoritma tersebut.

Tabel 4. Hasil evaluasi performa

No	Model	Precision	Recall	Accuracy	F1-Score
1	CatBoost	0.9629	0.958	0.973	0.958
2	GradientBoost	0.9333	0.943	0.947	0.932
3	AdaBoost	0.9333	0.943	0.947	0.932
4	LightGBM	0.8674	0.912	0.894	0.881
5	XGBoost	0.8674	0.912	0.894	0.881

Berdasarkan hasil pada Tabel 4, algoritma *CatBoost* menunjukkan performa yang sangat baik di antara kelima model dengan *akurasi* sebesar 97,37%, *precision* 96,29%, *recall* 95,83%, dan *F1-score* 95,82%. Setelah dianalisis melalui studi literatur, keunggulan

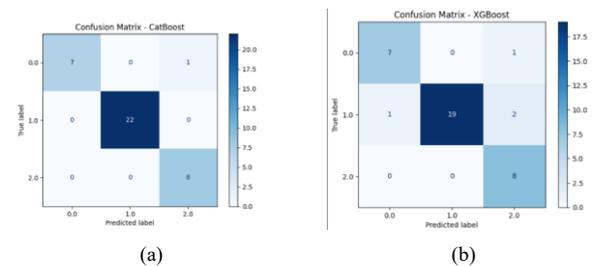
CatBoost berasal dari kemampuannya mengelola fitur kategorikal secara efisien tanpa perlu proses *encoding* tambahan, serta kemampuannya dalam meminimalkan *overfitting* dibandingkan algoritma boosting lain. Selain itu, *CatBoost* juga menawarkan efisiensi komputasi yang tinggi, sehingga cocok digunakan untuk pengolahan data berukuran besar (Pahlevi et al., 2024).

Selain *CatBoost*, algoritma *AdaBoost* dan *GradientBoosting* juga menunjukkan performa yang baik dengan nilai akurasi dan metrik lain di atas 94%. Keduanya memiliki tingkat *precision* dan *recall* yang seimbang, menunjukkan kemampuan yang konsisten dalam mengidentifikasi kelas positif dan negatif. Namun, performa *AdaBoost* dan *GradientBoosting* masih sedikit di bawah *CatBoost*, terutama dalam mengoptimalkan penanganan data kategorikal secara otomatis.

Sementara itu, *LightGBM* dan *XGBoost* memperoleh *akurasi* yang lebih rendah, yaitu sebesar 89,47%. Meskipun kedua model ini cukup populer dalam berbagai penelitian, pada kasus ini performanya tidak setinggi model-model lain. Hal ini dapat disebabkan oleh karakteristik *dataset* yang mengandung cukup banyak fitur kategorikal, sehingga penanganan data tersebut pada *LightGBM* dan *XGBoost* kurang optimal tanpa *preprocessing* tambahan. Namun, kedua algoritma ini tetap memiliki keunggulan dalam hal kecepatan pelatihan dan efisiensi komputasi, sehingga tetap menjadi pilihan yang layak pada kasus dengan data berukuran besar [35].

Secara keseluruhan, hasil penelitian ini menegaskan bahwa pemilihan algoritma *boosting* yang tepat sangat berpengaruh terhadap akurasi dan efektivitas model prediksi gangguan tidur. *CatBoost* direkomendasikan sebagai algoritma terbaik untuk aplikasi prediksi gangguan tidur berbasis *machine learning* pada *dataset* serupa, terutama ketika terdapat banyak fitur kategorikal. Temuan ini sejalan dengan hasil penelitian lain yang menunjukkan bahwa penanganan fitur kategorikal yang efisien dan resiliensi terhadap *overfitting* menjadi keunggulan utama *CatBoost* [36]. Selain itu, hasil evaluasi yang diperoleh dari metrik *precision*, *recall*, dan *F1-score* yang tinggi pada *CatBoost* dan *Gradient Boosting* juga menunjukkan bahwa kedua model ini tidak hanya mampu meminimalkan kesalahan

prediksi, tetapi juga dapat menjaga keseimbangan antara deteksi kasus positif dan negatif. Temuan ini penting dalam konteks aplikasi nyata, di mana prediksi gangguan tidur yang akurat dan minim kesalahan sangat dibutuhkan untuk mendukung pengambilan keputusan dalam layanan kesehatan.



Gambar 8. (a) *confusion matrix* Catboost, (b) *confusion matrix* XGBoost

Untuk memperkuat pemahaman terhadap kesalahan prediksi yang dilakukan oleh model, dilakukan analisis menggunakan *confusion matrix* yang ditunjukkan Gambar 8(a) menampilkan *confusion matrix* dari model *CatBoost*, yang merupakan model dengan performa terbaik. Berdasarkan visualisasi tersebut, *CatBoost* menunjukkan kemampuan klasifikasi yang sangat baik untuk ketiga kelas target, yaitu *None*, *Sleep Apnea*, dan *Insomnia*. Model ini hanya melakukan satu kesalahan klasifikasi pada kelas *None* yang diprediksi sebagai *Insomnia*, sedangkan semua data kelas *Sleep Apnea* dan *Insomnia* diklasifikasikan secara sempurna.

Sebagai pembandingan, Gambar 8(b) menampilkan *confusion matrix* dari model *XGBoost* yang memiliki akurasi terendah. Meskipun *XGBoost* masih mampu mengklasifikasikan kelas *None* dan *Insomnia* dengan cukup akurat, pada kelas *Sleep Apnea* model menunjukkan penurunan performa dengan tiga kesalahan klasifikasi: satu kasus *Sleep Apnea* diklasifikasikan sebagai *None*, dan dua sebagai *Insomnia*. Kesalahan ini berdampak pada penurunan nilai *recall* dan *F1-score* model secara keseluruhan. Hal ini memperkuat kesimpulan bahwa *XGBoost* kurang optimal dalam menangani fitur kategorikal secara langsung.

5. KESIMPULAN

Penelitian ini bertujuan untuk membandingkan lima algoritma *boosting* yaitu *AdaBoost*, *CatBoost*, *LightGBM*, *Gradient Boosting*, dan *XGBoost* dalam memprediksi gangguan tidur berdasarkan faktor-faktor kesehatan dan gaya hidup. Dengan menggunakan *dataset Sleep Health and Lifestyle*, penelitian ini menerapkan berbagai teknik pra-pemrosesan data, termasuk normalisasi dan pemilihan fitur, guna memastikan kualitas data yang optimal sebelum digunakan dalam model pembelajaran mesin. Evaluasi dilakukan menggunakan metrik akurasi, *presisi*, *recall*, dan *F1-score* untuk menentukan performa terbaik dari masing-masing algoritma.

Berdasarkan hasil evaluasi, algoritma *CatBoost* menunjukkan performa paling unggul dengan akurasi 97,37%, *presisi* 96,29%, *recall* 95,83%, dan *F1-score* 95,82%. Keunggulan *CatBoost* ini berasal dari kemampuannya dalam menangani fitur kategorikal secara lebih efektif tanpa memerlukan *encoding* tambahan, serta kemampuannya dalam mengurangi *overfitting* dibandingkan dengan metode *boosting* lainnya. Selain itu, algoritma ini juga memiliki efisiensi komputasi yang baik, sehingga dapat digunakan untuk pemrosesan data dalam skala besar.

Dari hasil penelitian ini, dapat disimpulkan bahwa algoritma *boosting*, khususnya *CatBoost*, merupakan teknik yang sangat efektif dalam mendeteksi pola gangguan tidur. Penggunaan algoritma ini memungkinkan identifikasi dini terhadap individu yang berisiko mengalami gangguan tidur berdasarkan berbagai variabel kesehatan dan gaya hidup. Dengan demikian, model prediksi yang dikembangkan dapat menjadi alat bantu bagi tenaga medis dan peneliti dalam melakukan intervensi kesehatan yang lebih tepat waktu dan berbasis data.

Hasil penelitian ini membuka peluang bagi penelitian lanjutan untuk meningkatkan akurasi model prediksi gangguan tidur. Salah satu arah penelitian di masa depan adalah integrasi fitur tambahan seperti data biometrik dari sensor EEG dan ECG, yang dapat memberikan informasi lebih mendalam mengenai pola tidur seseorang. Data biometrik ini diyakini mampu meningkatkan sensitivitas dan spesifisitas model dalam membedakan berbagai jenis gangguan tidur, seperti *insomnia*, *sleep apnea*, dan gangguan tidur lainnya.

DAFTAR PUSTAKA

- [1] K. Jasmine and F. Martdianty, "The Analisis Pengaruh Abusive Supervision terhadap Employee Creativity dengan Sleep Deprivation, Emotional Exhaustion, dan Self Efficacy sebagai Mediator," *Jurnal Manajemen dan Organisasi*, vol. 13, no. 1, pp. 23–35, 2022, doi: 10.29244/jmo.v13i1.34343.
- [2] J. L. Putra and W. F. Hidayat, "Prediksi Kualitas Tidur: Pendekatan Machine Learning yang Mengintegrasikan Faktor Kesehatan dan Lingkungan," *Computer Science (CO-SCIENCE)*, vol. 4, no. 2, pp. 157–162, 2024, doi: 10.31294/coscience.v4i2.4737.
- [3] K. J. Mehta, "Effect of sleep and mood on academic performance—at interface of physiology, psychology, and education," Dec. 01, 2022, *Springer Nature*. doi: 10.1057/s41599-021-01031-1.
- [4] I. U. Aisyah and N. Q. Wijayani, "Penggunaan Gadget Terhadap Kualitas Tidur Dan Kesehatan Mental Remaja," *Jurnal Multidisiplin Ilmu Sosial*, vol. 2, no. 9, pp. 31–40, 2023.
- [5] D. Sari, "Prediksi Gangguan Tidur pada Sleep Health and Lifestyle Menggunakan Support Vector Machine dan Neural Network," *JAVIT : Jurnal Vokasi Informatika*, pp. 36–42, 2024, doi: 10.24036/javit.v4i1.168.
- [6] S. S. Azzahra, "Obstructive Sleep Apnea (OSA) Sebagai Faktor Resiko Hipertensi," *Jurnal Ilmiah Kesehatan Sandi Husada*, vol. 10, no. 2, pp. 321–324, 2019, doi: 10.35816/jiskh.v10i2.180.
- [7] E. Mardiani, N. Rahmansyah, S. Ningsih, and ..., "Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma," *Innovative: Journal Of ...*, vol. 3, no. 2, pp. 13880–13892, 2023.
- [8] I. M. M. Matin, M. Agustin, B. Sugiarto, and A. N. Asri, "Deteksi Malware Menggunakan Machine Learning Dengan Metode Ensemble," *Prosiding Sains Nasional dan Teknologi*, vol. 13, no. 1, pp. 265–270, Nov. 2023, doi: 10.36499/psnst.v13i1.9224.
- [9] L. M. Cendani and A. Wibowo, "Perbandingan Metode Ensemble Learning pada Klasifikasi Penyakit Diabetes," *Jurnal Masyarakat Informatika*, vol. 13, no. 1, pp. 33–44, 2022, doi: 10.14710/jmasif.13.1.42912.
- [10] Jan Melvin Ayu Soraya Dachi and Pardomuan Sitompul, "Analisis Perbandingan Algoritma XGBoost dan Algoritma Random Forest Ensemble Learning pada Klasifikasi Keputusan Kredit," *Jurnal Riset Rumpun Matematika Dan Ilmu Pengetahuan Alam*, vol.

- 2, no. 2, pp. 87–103, 2023, doi: 10.55606/jurrimipa.v2i2.1470.
- [11] M. Maulidah and N. Hidayati, “Prediksi Kesehatan Tidur Dan Gaya Hidup Menggunakan Machine Learning,” *CONTEN : Computer and Network Technology*, vol. 4, no. 1, pp. 81–86, 2024.
- [12] I. Fahrudi, I. K. E. Purnama, and M. H. Purnomo, “Asesmen ECG-Apnea Satu Sadapan untuk Peningkatan Akurasi Klasifikasi Gangguan Tidur Berdasarkan AdaBoost,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, pp. 196–204, 2020, doi: 10.22146/jnteti.v9i2.159.
- [13] A. Maysa, S. P. A. Alkadri, and I. Istikomah, “Klasifikasi Tingkat Kepuasan di Maskapai Penerbangan: Studi Komparasi Algoritma K-NN dan Adaboost,” *Jurnal Informatika Polinema*, vol. 10, no. 3, pp. 405–412, 2024, doi: 10.33795/jip.v10i3.5166.
- [14] I. Maulana, A. M. Siregar, S. Arum, P. Lestari, S. Faisal, and U. B. Perjuangan, “OPTIMAL STUDY OF REAL-ESTATE PRICE PREDICTION MODELS USING STUDI OPTIMAL MODEL PREDIKSI HARGA REAL-ESTATE MENGGUNAKAN,” vol. 5, no. 4, pp. 1149–1164, 2024.
- [15] H. Yang, Z. Chen, H. Yang, and M. Tian, “Predicting Coronary Heart Disease Using an Improved LightGBM Model: Performance Analysis and Comparison,” *IEEE Access*, vol. 11, no. March, pp. 23366–23380, 2023, doi: 10.1109/ACCESS.2023.3253885.
- [16] M. R. Andryan, M. Fajri, and N. Sulistyowati, “Komparasi Kinerja Algoritma Xgboost Dan Algoritma Support Vector Machine (Svm) Untuk Diagnosis Penyakit Kanker Payudara,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 1, p. 1, 2022, doi: 10.26798/jiko.v6i1.500.
- [17] Muflih Ihza Rifatama, Mohammad Reza Faisal, Rudy Herteno, Irwan Budiman, and Muhammad Itqan Mazdadi, “Optimasi Algoritma K-Nearest Neighbor Dengan Seleksi Fitur Menggunakan Xgboost,” *Jurnal Informatika dan Rekayasa Elektronik*, vol. 6, no. 1, pp. 64–72, 2023, doi: 10.36595/jire.v6i1.723.
- [18] A. M. Eko Fitra Firmadani, A. Hudawi AS, and A. Tholib, “Optimasi Model CatBoost dengan Feature Selection dan Hyperparameter Tuning untuk Prediksi Nasabah Bank Potensial,” *Academic Journal of Computer Science Research*, vol. 6, no. 2, p. 118, 2024, doi: 10.38101/ajcsr.v6i2.15656.
- [19] L. Qadrini, A. Sepperwali, and A. Aina, “Decision Tree Dan Adaboost Pada Klasifikasi Penerima Program Bantuan Sosial,” *Jurnal Inovasi Penelitian*, vol. 2, no. 7, pp. 1959–1966, 2021.
- [20] M. Anjas Aprihartha, J. Prasetya, D. Sefri, and I. Fallo, “Implementasi CART-Real Adaboost dalam Memprediksi Minat Pelanggan Membeli Sepatu A B S T R A K INFORMASI ARTIKEL A B S T R A C T,” vol. 12, no. 1, pp. 35–46, 2024, [Online]. Available: <https://ejournal.upi.edu/index.php/JEM>
- [21] S. Diantika, “Penerapan Teknik Random Oversampling Untuk Mengatasi Imbalance Class Dalam Klasifikasi Website Phishing Menggunakan Algoritma Lightgbm,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 19–25, 2023, doi: 10.36040/jati.v7i1.6006.
- [22] A. Pramudyantoro, E. Utami, and D. Ariatmanto, “Penggabungan K-Nearest Neighbors dan Lightgbm Untuk prediksi Diabetes Pada Dataset Pima Indians: Menggunakan pendekatan Exploratory Data Analysis,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1133–1144, 2024, [Online]. Available: <https://jurnal.stkipgritlungagung.ac.id/index.php/jupi/article/view/4966/2114>
- [23] F. Firmansyah and A. Yulianto, “Pemodelan Pembelajaran Mesin untuk Prediksi Kesehatan Mental di Tempat Kerja,” *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 397–407, 2024, doi: 10.33395/jmp.v13i1.13674.
- [24] R. Leonardo, J. Pratama, and C. Chrisnatalis, “Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing,” *Jurnal Teknologi Dan Ilmu Komputer Prima (Jutikomp)*, vol. 3, no. 2, pp. 455–459, 2020, doi: 10.34012/jutikomp.v3i2.1321.
- [25] K. R. Supriyanti, B. A. Damiri, and W. N. Ramadhan, “Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di Provinsi Sumatera Utara Menggunakan Metode Fuzzy C-Means,” *Jurnal Statistika dan Komputasi*, vol. 3, no. 1, pp. 1–10, 2024, doi: 10.32665/statkom.v3i1.2623.
- [26] N. Gupta *et al.*, “Data Quality Toolkit: Automatic assessment of data quality and remediation for machine learning datasets,” 2021.
- [27] M. Halawa, “Normalisasi Data Penjualan Ternak Menggunakan Metode Binning Untuk Menentukan Pola Penjualan Ternak Menerapkan Metode Apriori,” vol. 2, no. 2, pp. 73–81, 2024.
- [28] Ary Prandika Siregar, Dwi Priyadi Purba, Jojo Putri Pasaribu, and Khairul Reza Bakara, “Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke,”

- Jurnal Penelitian Rumpun Ilmu Teknik*, vol. 2, no. 4, pp. 155–164, 2023, doi: 10.55606/juprit.v2i4.3039.
- [29] E. Muningsih, “Kombinasi Metode K-Means Dan Decision Tree Dengan Perbandingan Kriteria Dan Split Data,” *Jurnal Teknoinfo*, vol. 16, no. 1, p. 113, 2022, doi: 10.33365/jti.v16i1.1561.
- [30] S. Smiti, M. Soui, and K. Ghedira, “Tri-XGBoost model improved by BLSmote-ENN: an interpretable semi-supervised approach for addressing bankruptcy prediction,” *Knowl Inf Syst*, vol. 66, no. 7, pp. 3883–3920, 2024, doi: 10.1007/s10115-024-02067-w.
- [31] A. Samat, E. Li, P. Du, S. Liu, and J. Xia, “GPU-Accelerated CatBoost-Forest for Hyperspectral Image Classification Via Parallelized mRMR Ensemble Subspace Feature Selection,” *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, pp. 3200–3214, 2021, doi: 10.1109/JSTARS.2021.3063507.
- [32] J. Cao, S. Kwong, and R. Wang, “A noise-detection based AdaBoost algorithm for mislabeled data,” *Pattern Recognit*, vol. 45, no. 12, pp. 4451–4465, 2012, doi: 10.1016/j.patcog.2012.05.002.
- [33] M. Alqahtani, A. Gumaei, H. Mathkour, and M. M. Ben Ismail, “A genetic-based extreme gradient boosting model for detecting intrusions in wireless sensor networks,” *Sensors (Switzerland)*, vol. 19, no. 20, 2019, doi: 10.3390/s19204383.
- [34] S. Clara, D. Laksmi Prianto, R. Al Habsi, E. Friscila Lumbantobing, and N. Chamidah, “Implementasi Seleksi Fitur Pada Algoritma Klasifikasi Machine Learning Untuk Prediksi Penghasilan Pada Adult Income Dataset,” *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*, vol. 2, no. 1, pp. 741–747, 2021.
- [35] P. Septiana Rizky, R. Haiban Hirzi, and U. Hidayaturrohmah, “Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang,” *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 15, no. 2, pp. 228–236, 2022, doi: 10.36456/jstat.vol15.no2.a5548.
- [36] A. Darmawan, M. Muliadi, D. Kartini, T. H. Saragih, and R. A. Nugraha, “Implementasi Catboost Dengan Menggunakan Hyper-Parameter Tuning Bayesian Search Untuk Memprediksi Penyakit Diabetes,” *Jurnal Komputasi*, vol. 11, no. 2, pp. 148–156, 2023, doi: 10.23960/komputasi.v11i2.13746.