ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 15, Nomor 2, Tahun 2026, Halaman 483 - 494
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

PERBANDINGAN PERFORMA ALGORITMA XGBOOST DAN CATBOOST DENGAN OPTIMASI *RANDOM SEARCH* UNTUK KLASIFIKASI KEPESERTAAN BPJS PBI

Leni Pamularsih[1*], **Husnul Amira**[1], **Rahmi Nurul Ainun Fitrah**[1], **Aulia Rizki Firdawanti**[1]
[1]Program Studi Statistika dan Sains Data, Sekolah Sains Data, Matematika, dan Informatika, Institut Pertanian Bogor, Bogor, Jawa Barat 16680
***e-mail**: lenipamularsih@apps.ipb.ac.id

DOI: 10.14710/j.gauss.15.2.483-494

Article Info:
Received: 2026-07-10
Accepted: 2026-11-16
Available Online: 2026-11-25

Keywords:
BPJS PBI; CatBoost; XGBoost; *Random Search*; SHAP

Abstract: The National Health Insurance (JKN) program through BPJS Kesehatan aims to provide health protection for all Indonesian citizens. However, identifying eligible Contribution Assistance Beneficiaries (PBI) remains challenging due to complex socioeconomic conditions. This study compares the performance of Extreme Gradient Boosting (XGBoost), native Categorical Boosting (CatBoost), and CatBoost with manual one-hot encoding, optimized using Random Search, for classifying BPJS PBI membership in East Java Province. The study used 2024 National Socioeconomic Survey (SUSENAS) data consisting of 104,599 individuals and 27 socioeconomic predictor variables. Model performance was evaluated using accuracy, sensitivity, specificity, precision, F1-score, balanced accuracy, and ROC-AUC. The results show that all three approaches achieved relatively similar classification performance. XGBoost slightly outperformed the other approaches in overall accuracy and balanced accuracy, achieving 0.7505 and 0.7507, respectively. Meanwhile, CatBoost with manual one-hot encoding achieved the highest sensitivity (0.7463), F1-score (0.7504), and ROC-AUC (0.8216). SHAP analysis based on XGBoost identified district/city of residence as the most influential predictor, followed by receipt of BPNT and PKH assistance, marital status, and car ownership. These findings indicate that geographic and socioeconomic characteristics contribute to distinguishing BPJS PBI membership status. The results may provide supporting information for evaluating the targeting accuracy of health insurance contribution assistance programs.

1. PENDAHULUAN

BPJS Kesehatan merupakan program jaminan sosial pemerintah yang bertujuan memberikan perlindungan kesehatan bagi seluruh masyarakat Indonesia. Berdasarkan Peraturan Pemerintah Nomor 101 Tahun 2012, kepesertaan BPJS Kesehatan mencakup peserta Penerima Bantuan Iuran (PBI), yaitu fakir miskin dan masyarakat tidak mampu yang iurannya ditanggung pemerintah, serta peserta Non-PBI yang membayar iuran secara mandiri atau melalui pemberi kerja. Provinsi Jawa Timur dipilih sebagai wilayah penelitian karena memiliki keragaman kondisi sosial ekonomi antarkabupaten/kota. Berdasarkan data Badan Pusat Statistik (BPS), persentase penduduk miskin Jawa Timur pada September 2024 mencapai 9,56%, lebih tinggi dibandingkan angka nasional sebesar 8,57% (BPS Provinsi Jawa Timur, 2025). Selain itu, cakupan Jaminan Kesehatan Nasional (JKN) di Jawa Timur mencapai 95,98% per 1 Desember 2024, masih di bawah target nasional sebesar 98% (Dinas Kesehatan Provinsi Jawa Timur, 2025). Kondisi tersebut menunjukkan bahwa perluasan cakupan jaminan kesehatan dan ketepatan identifikasi penerima bantuan iuran masih menjadi perhatian. Beragamnya kondisi sosial ekonomi rumah tangga juga menjadi tantangan dalam mengidentifikasi peserta PBI secara akurat dan tepat sasaran (Agustina *et*

al., 2019; Kosasih *et al*., 2022). Oleh karena itu, diperlukan pendekatan klasifikasi berbasis data untuk mengidentifikasi status kepesertaan BPJS PBI secara objektif.

Klasifikasi kepesertaan BPJS PBI dan Non-PBI dipengaruhi oleh kombinasi berbagai karakteristik sosial ekonomi rumah tangga yang saling berhubungan (Kosasih *et al*., 2022; Putri *et al*., 2023). Hubungan antarvariabel tersebut berpotensi membentuk pola nonlinear dan interaksi yang kompleks. Salah satu pendekatan yang relevan adalah *Machine Learning*, khususnya *ensemble learning* dengan teknik *boosting*, yang membangun model secara berurutan untuk memperbaiki kesalahan prediksi model sebelumnya (Ptr *et al*., 2024; Efendi dan Suharjo, 2026). Algoritma *boosting* yang digunakan dalam penelitian ini adalah *Extreme Gradient Boosting* (XGBoost) dan *Categorical Boosting* (CatBoost). XGBoost merupakan algoritma berbasis pohon keputusan yang mampu mempelajari pola nonlinear dan interaksi antarvariabel serta efisien dalam mengolah data berukuran besar (Bentéjac *et al*., 2019). Sementara itu, CatBoost memiliki mekanisme khusus untuk menangani variabel kategorik (Efendi dan Suharjo, 2026). Pemilihan kedua algoritma didasarkan pada karakteristik data SUSENAS 2024 yang terdiri atas 104.599 observasi individu dan 27 variabel prediktor, yaitu 25 variabel kategorik dan 2 variabel numerik. Variabel tersebut mencakup karakteristik demografi, kondisi tempat tinggal, kepemilikan aset, dan penerimaan bantuan sosial. Dominasi variabel kategorik dan keragaman karakteristik prediktor tersebut menjadikan XGBoost dan CatBoost relevan untuk mempelajari pola klasifikasi status kepesertaan BPJS PBI dan Non-PBI. Perbandingan keduanya diperlukan untuk mengetahui algoritma yang memberikan kinerja terbaik pada karakteristik data tersebut.

Kinerja kedua algoritma juga dipengaruhi oleh pengaturan *hyperparameter* yang menentukan proses pembentukan model. Pengaturan yang kurang tepat dapat memengaruhi kemampuan prediksi dan generalisasi model (Handyanto *et al*., 2026). Oleh karena itu, digunakan metode *Random Search* untuk mengeksplorasi kombinasi *hyperparameter* secara acak dari ruang pencarian yang ditentukan. Metode ini memungkinkan eksplorasi berbagai kombinasi parameter dengan jumlah percobaan terbatas (Hartono *et al*., 2024; Handyanto *et al*., 2026). Selain ketepatan klasifikasi, diperlukan interpretasi untuk mengetahui variabel yang berkontribusi terhadap hasil prediksi. Metode *Shapley Additive exPlanations* (SHAP) digunakan untuk mengukur kontribusi setiap variabel terhadap hasil prediksi berdasarkan konsep nilai Shapley dalam teori permainan (Lundberg dan Lee, 2017). SHAP memungkinkan identifikasi variabel paling berpengaruh secara global serta penjelasan kontribusi variabel pada setiap observasi secara lokal.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan membandingkan kinerja XGBoost dan CatBoost yang dioptimasi menggunakan *Random Search* dalam klasifikasi status kepesertaan BPJS PBI dan Non-PBI di Provinsi Jawa Timur. Penelitian ini juga bertujuan mengidentifikasi variabel sosial ekonomi rumah tangga yang paling berkontribusi terhadap hasil klasifikasi menggunakan SHAP. Hasil penelitian diharapkan dapat memberikan rekomendasi model klasifikasi terbaik serta informasi pendukung untuk evaluasi ketepatan sasaran program bantuan iuran kesehatan.

2. TINJAUAN PUSTAKA

XGBoost merupakan pengembangan dari *Gradient Boosting* yang bekerja dengan membangun model secara bertahap untuk memperbaiki kesalahan prediksi model sebelumnya, dengan mengoptimalkan fungsi objektif berupa *loss* dan regularisasi sehingga mampu meningkatkan akurasi sekaligus mengurangi *overfitting* (Bentéjac *et al*., 2019).
Fungsi objektif XGBoost dirumuskan pada Persamaan 1:

$$\mathcal{L} = \sum_{i=1}^{n} \ell\left(y_i, \hat{y}_i\right) + \sum_{m=1}^{M} \Omega\left(f_m\right) \quad (1)$$

dengan $\hat{y}_i = \sum_{m=1}^{M} f_m(x_i)$ dan $\Omega(f_m) = \gamma T_m + \frac{1}{2}\lambda \sum_{j=1}^{T_m} w_j^2$, dimana n adalah jumlah pengamatan, M adalah jumlah pohon keputusan, serta f_m adalah fungsi pohon keputusan ke-m. Untuk klasifikasi biner, fungsi *loss* menggunakan *binary logistic loss* yang dirumuskan dalam Persamaan 2:

$$\ell(y_i, \hat{p}_i) = -y_i \ln(\hat{p}_i) - (1 - y_i) \ln(1 - \hat{p}_i) \quad (2)$$

dengan $\hat{p}_i = \frac{1}{1+\exp(-\hat{y}_i)}$.

CatBoost (*Categorical Boosting*) merupakan algoritma *boosting* yang dirancang khusus untuk menangani variabel kategorik secara efisien tanpa memerlukan proses *one-hot encoding* yang kompleks (Prokhorenkova *et al*., 2018). CatBoost menggunakan teknik *ordered boosting* untuk mengurangi bias dan meningkatkan kemampuan generalisasi model. Sebagai *base learner*, prediksi satu pohon keputusan dinyatakan pada Persamaan 3:

$$H(x_i) = \sum_{j=1}^{J} c_j \, 1\{x_i \in R_j\} \quad (3)$$

dengan $H(x_i)$ merupakan prediksi pohon untuk pengamatan ke-i, J merupakan jumlah daun, c_j merupakan nilai prediksi pada daun ke-j, dan R_j merupakan wilayah pengamatan pada daun ke-j. Prediksi keseluruhan model diperoleh melalui agregasi hasil prediksi setiap pohon dituliskan pada Persamaan 4:

$$F_M(x_i) = F_0(x_i) + \eta \sum_{m=1}^{M} H_m(x_i) \quad (4)$$

dengan $F_M(x_i)$ merupakan prediksi akhir model, $F_0(x_i)$ merupakan prediksi awal, η merupakan *learning rate*, M merupakan jumlah pohon keputusan, dan $H_m(x_i)$ merupakan prediksi pohon ke-m. *Hyperparameter tuning* dilakukan untuk memperoleh kombinasi parameter yang mampu menghasilkan performa model terbaik. Pada penelitian ini digunakan metode *random Search*, yaitu teknik optimasi yang melakukan pencarian kombinasi hyperparameter secara acak dari ruang pencarian yang telah ditentukan (Mulyo & Zyen, 2025).

Pada XGBoost, *random search* digunakan untuk memperoleh kombinasi rentang parameter yang disesuaikan, meliputi tingkat pembelajaran (*learning rate*), kedalaman maksimum pohon (*max depth*), bobot minimum anak simpul (*min child weight*), serta parameter regularisasi L1 (*alpha*) dan L2 (*lambda*). Sedangkan CatBoost mengoptimasi parameter *learning rate*, *depth*, dan *L2 leaf regularization*. Setelah mendapatkan kombinasi *hyperparameter* terbaik, dilakukan penentuan *threshold* optimal untuk mengkonversi probabilitas dari klasifikasi menjadi dua kelas biner.Evaluasi model dilakukan untuk menilai kemampuan kedua model dalam mengklasifikasikan status kepesertaan BPJS PBI dan Non-PBI, dengan mengacu pada *confusion matrix*. Untuk memberikan gambaran menyeluruh mengenai ketepatan klasifikasi serta kemampuan model dalam membedakan kedua kelas secara optimal, kinerja kedua algoritma diukur menggunakan serangkaian metrik klasifikasi biner yang komprehensif, meliputi *accuracy*, *sensitivity*, *specificity*, *F1-score*, *kappa*, dan *ROC-AUC* (Tharwat, 2021).

Untuk meningkatkan interpretabilitas model, penelitian ini menggunakan metode SHAP (*SHapley Additive Explanations*). SHAP merupakan metode *Explainable Artificial Intelligence* (XAI) yang didasarkan pada konsep nilai *Shapley* dalam teori permainan untuk mengukur kontribusi setiap variabel terhadap prediksi model. Melalui SHAP, model yang bersifat *black box* dapat diinterpretasikan dengan lebih baik sehingga variabel-variabel yang berpengaruh terhadap keputusan model dapat diidentifikasi secara lebih jelas (Lundberg & Lee, 2017). Persamaan SHAP dituliskan dalam Persamaan 5 (Lundberg & Lee, 2017):

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i \, z'_i \tag{5}$$

dengan $g(z')$ merupakan keluaran model yang dijelaskan, ϕ_0 merupakan nilai dasar (*baseline*) berupa nilai ekspektasi keluaran model pada *background data*, ϕ_i merupakan kontribusi fitur ke-i, dan $z'_i \in \{0,1\}$ menunjukkan keberadaan fitur ke-i dalam representasi tersederhanakan. Pada penelitian ini, nilai SHAP diinterpretasikan pada skala *raw margin* (log-odds), sehingga nilai SHAP positif menunjukkan bahwa suatu prediktor meningkatkan kecenderungan individu untuk diklasifikasikan sebagai penerima BPJS PBI, sedangkan nilai SHAP negatif menunjukkan kecenderungan sebaliknya.

3. METODE PENELITIAN

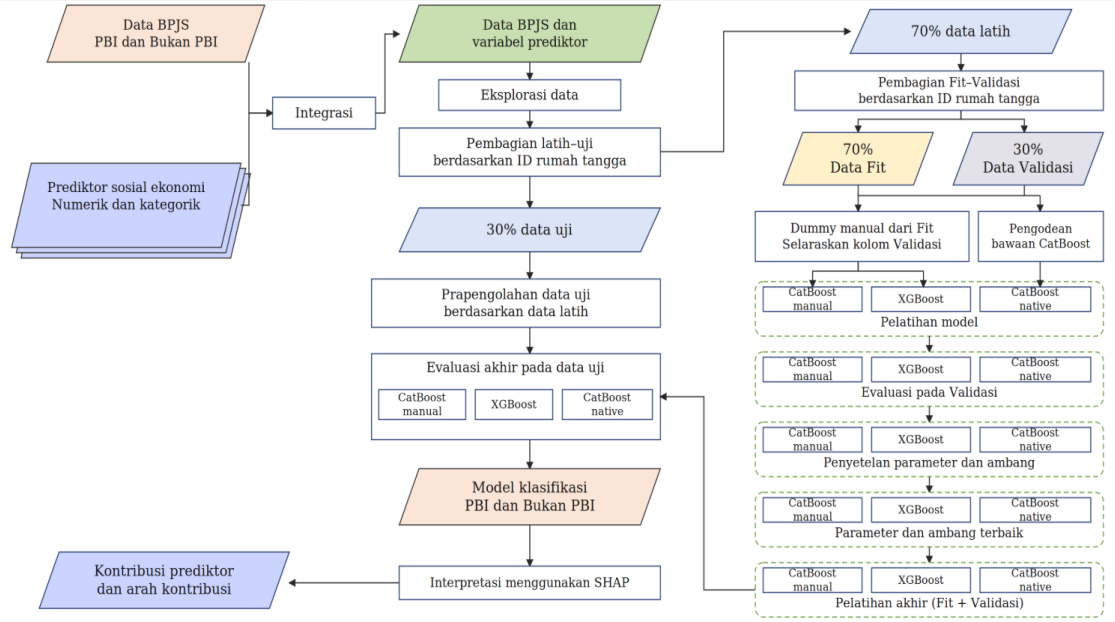
Data dalam penelitian ini menggunakan data Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2024 di Provinsi Jawa Timur. Satuan amatan yang digunakan adalah anggota rumah tangga (individu) sebanyak 104.599 jiwa.Variabel dependen adalah status kepesertaan BPJS Kesehatan yang diklasifikasikan menjadi PBI (1) dan Non-PBI (0), sedangkan variabel independen terdiri atas 27 prediktor numerik dan kategorik sebagaimana disajikan pada Tabel 1. Pemilihan prediktor didasarkan pada karakteristik sosial ekonomi rumah tangga yang merepresentasikan tingkat kesejahteraan serta digunakan dalam identifikasi rumah tangga miskin dan rentan miskin (BPS, 2019; Kementerian PPN/Bappenas, 2023).

Tabel 1. Variabel Penelitian

Variabel	Tipe Data	Variabel	Tipe Data
R1101A-Kepesertaan PBI dan Non-PBI (Y)	Kategorik	**R2001B**-Kepemilikan lemari es/kulkas (X14)	Kategorik
R102-Kabupaten/Kota (X1)	Kategorik	**R2001C**-Kepemilikan AC (X15)	Kategorik
R105-Tipe daerah (perkotaan/pedesaan) (X2)	Kategorik	**R2001F**-Kepemilikan laptop (X16)	Kategorik
R301-Banyaknya Anggota Rumah Tangga (X3)	Numerik	**R2001G**-Kepemilikan emas ≥ 10 gram (X17)	Kategorik
R404-Status perkawinan (X4)	Kategorik	**R2001H**-Kepemilikan sepeda motor (X18)	Kategorik
R1102-Mempunyai keluhan kesehatan (X5)	Kategorik	**R2001K**-Kepemilikan mobil (X19)	Kategorik
R1802-Status kepemilikan rumah (X6)	Kategorik	**R2001M**-Kepemilikan tanah/lahan (X20)	Kategorik
R1804-Luas lantai rumah (X7)	Numerik	**R2202**-Penerima Kartu Keluarga Sejahtera (X21)	Kategorik
R1807-Bahan bangunan utama dinding rumah terluas (X8)	Kategorik	**R2203**-Penerima Program Keluarga Harapan (X22)	Kategorik
R1808-Bahan bangunan utama lantai rumah terluas (X9)	Kategorik	**R2207**-Penerima Bantuan Pangan Non Tunai (X23)	Kategorik
R1809A-Fasilitas tempat buang air besar (X10)	Kategorik	**R2209A**-Penerima BLT Desa (X24)	Kategorik
R1810A-Sumber air utama yang digunakan sebagai air minum (X11)	Kategorik	**R2209C**-Penerima Bantuan Pangan/Beras (X25)	Kategorik
R1816-Sumber penerangan (X12)	Kategorik	**R2211A**-Penerima Bantuan Tunai Rutin dari Pemda (X26)	Kategorik
R2001A-Kepemilikan tabung gas 5,5 kg atau lebih (X13)	Kategorik	**R2211B**-Penerima Bantuan Tunai Tidak Rutin dari Pemda (X27)	Kategorik

Prosedur analisis data meliputi eksplorasi data, prapemrosesan, pembentukan model, evaluasi, dan interpretasi. Penanganan variabel kategorik dilakukan menggunakan *one-hot encoding*, yaitu transformasi kategori menjadi kolom biner. Penelitian membandingkan XGBoost, CatBoost dengan *one-hot encoding* manual, dan CatBoost *native* untuk mengevaluasi pengaruh prapemrosesan terhadap kinerja model. Variabel R1804 dan R301 diperlakukan sebagai numerik, sedangkan prediktor lainnya sebagai kategorik. Pada XGBoost dan CatBoost manual digunakan kolom dummy, sementara CatBoost *native* menggunakan tipe *factor*. Perbandingan ini mengacu pada Ray *et al.* (2026), yang melaporkan perbedaan kinerja CatBoost pada kedua pendekatan relatif kecil.

Selanjutnya, data dibagi menjadi data latih (70%) dan data uji (30%). Data latih digunakan untuk pembentukan model, validasi, dan optimasi *hyperparameter* dengan prosedur serta anggaran pencarian yang setara. Model terbaik dievaluasi menggunakan data uji untuk menentukan metode dengan kinerja terbaik. Pada tahap akhir, dilakukan interpretasi menggunakan SHAP untuk mengidentifikasi variabel yang paling berkontribusi terhadap klasifikasi kepesertaan BPJS PBI. *Workflow* analisis data disajikan pada Gambar 1.



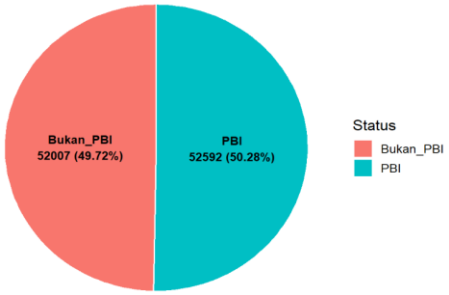
Gambar 1. *Workflow* Analisis Data

4. HASIL DAN PEMBAHASAN

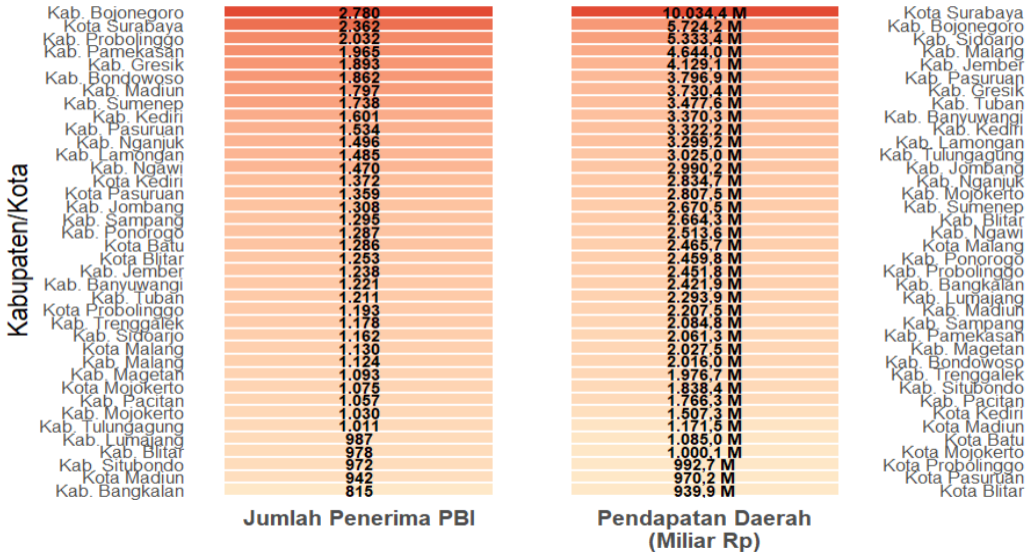
Tahap pertama sebelum pemodelan adalah melakukan eksplorasi data untuk memahami struktur, pola, distribusi, hubungan antarvariabel, serta mendeteksi anomali. Hasil eksplorasi menunjukkan bahwa data tidak mengandung *missing value*, sehingga seluruh observasi digunakan pada tahap pemodelan. Meskipun terdapat beberapa nilai ekstrem (*outlier*) pada variabel numerik, data tetap dipertahankan karena masih merepresentasikan kondisi sebenarnya pada data SUSENAS dan algoritma XGBoost maupun CatBoost relatif *robust* terhadap nilai ekstrem. Selanjutnya, proporsi penerima BPJS PBI dan BPJS Non-PBI divisualisasikan menggunakan *pie chart* pada Gambar 2.

Hasil visualisasi dapat dilihat pada Gambar 3, yang menunjukkan bahwa individu penerima BPJS PBI di Provinsi Jawa Timur tahun 2024 sebesar 52.592 (50,28%) dan

individu penerima bukan BPJS PBI sebesar 52.007 (49,72%). Selisih proporsi kedua kelas tersebut hanya sekitar 0,56%, sehingga data tidak mengalami ketimpangan kelas atau proporsi data seimbang. Selain itu, untuk melihat persebaran jumlah penerima BPJS PBI berdasarkan data susenas dan jumlah pendapatan daerah pada masing-masing kabupaten/kota di Provinsi Jawa Timur, dilakukan visualisasi menggunakan *heatmap*. Visualisasi heatmap menunjukkan adanya variasi jumlah penerima BPJS PBI dan pendapatan daerah antarkabupaten/kota di Jawa Timur. Variasi ini relevan untuk diamati karena sebagian pembiayaan PBI dapat bersumber dari APBD, sehingga pendapatan daerah dapat menggambarkan kapasitas fiskal daerah dalam mendukung kepesertaan PBI, meskipun jumlah penerima PBI juga dipengaruhi oleh kondisi sosial-ekonomi, jumlah penduduk, dan kebijakan penetapan peserta.

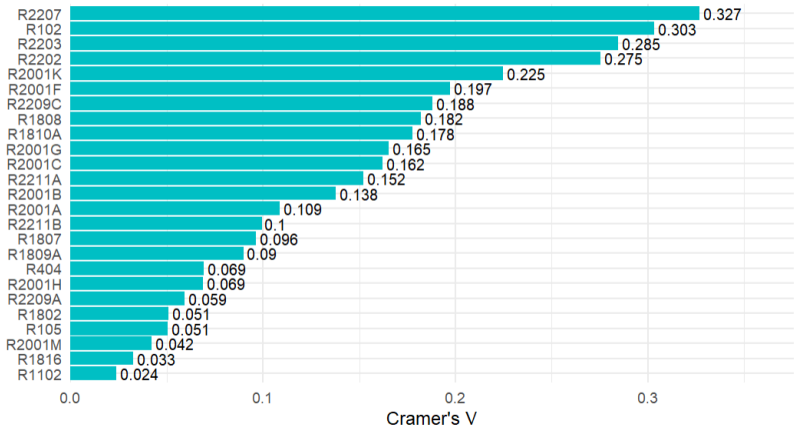


Gambar 2. Proporsi Individu Penerima BPJS PBI dan BPJS Non-PBI



Gambar 3. *Heatmap* Jumlah Penerima BPJS PBI dan Pendapatan Daerah di Jawa Timur Tahun 2024

Untuk mengetahui kekuatan hubungan dari 27 variabel prediktor dengan variabel target status kepersertaan BPJS PBI, dilakukan eksplorasi data dengan menggunakan Cramer's V. Ukuran ini digunakan untuk variabel kategorik. Berdasarkan Cramer's V pada Gambar 4, variabel kategorik yang memiliki hubungan paling kuat dengan status kepesertaan BPJS PBI adalah penerima Bantuan Pangan Non-Tunai (BPNT) (R2207) dengan nilai Cramer's V sebesar 0,327, diikuti oleh kabupaten/kota (R102) sebesar 0,303, penerima Program Keluarga Harapan (PKH) (R2203) sebesar 0,285, dan penerima Kartu Keluarga Sejahtera (KKS) (R2202) sebesar 0,275. Hasil ini menunjukkan bahwa variabel bantuan sosial dan karakteristik wilayah memiliki keterkaitan yang relatif lebih kuat dengan status kepesertaan BPJS PBI dibandingkan variabel lainnya.



Gambar 4. Cramer's V antara Variabel Kategorik dengan Variabel Target

Setelah eksplorasi data, dilakukan pemodelan menggunakan XGBoost, CatBoost *native*, dan CatBoost dengan *one-hot encoding* manual untuk mengklasifikasikan status kepesertaan BPJS PBI dan Non-PBI. Data dibagi berdasarkan ID rumah tangga menjadi 70% data latih dan 30% data uji. Data latih kemudian dibagi menjadi 70% data *fit* dan 30% *validation*, sehingga proporsi keseluruhan masing-masing sebesar 49%, 21%, dan 30%. Pembagian dilakukan secara terstratifikasi dengan *seed* 123 untuk memastikan anggota rumah tangga yang sama berada pada subset yang sama.

Optimasi *hyperparameter* dilakukan menggunakan *Random Search* sebanyak 30 kombinasi. Setiap kandidat dilatih hingga 2.500 iterasi dengan *early stopping* setelah 80 iterasi tanpa peningkatan AUC validasi. *Threshold* dioptimalkan pada rentang 0,20–0,80 dengan interval 0,01. Kandidat terbaik dipilih berdasarkan akurasi tertinggi, dengan *balanced accuracy*, *F1-score*, dan AUC sebagai kriteria tambahan. Ruang pencarian *hyperparameter* disajikan pada Tabel 2.

Tabel 2. Ruang Pencarian *Hyperparameter*

XGBoost	CatBoost(Manual dan *Native*)	Ruang Pencarian
eta	learning_rate	{0,005; 0,01; 0,02; 0,03; 0,05; 0,07; 0,10}
max_depth	depth	{3; 4; 5; 6; 7; 8}
lambda	l2_leaf_reg	0,1–30
colsample_bytree	rsm	0,70–1,00
min_child_weight	-	{1; 3; 5; 8; 10; 15; 20}
subsample	-	0,65–0,95
gamma	-	0–5
alpha	-	0–10
-	random_strength	0–10
-	bagging_temperature	0–5

Berdasarkan pencarian pada data validation, diperoleh kombinasi *hyperparameter* dan *threshold* terpilih untuk setiap pendekatan sebagaimana disajikan pada Tabel 3. Untuk menilai performa model, digunakan beberapa metrik evaluasi seperti *accuracy, sensitivity, specificity, precision, F1-score, balanced accuracy* dan ROC-AUC pada Tabel 4. Hasil menunjukkan bahwa ketiga pendekatan dengan ruang pencarian yang sama menghasilkan kinerja klasifikasi yang relatif serupa. XGBoost memperoleh *accuracy* data uji tertinggi sebesar 75,05%, diikuti CatBoost dengan *one-hot encoding* manual sebesar 74,99% dan CatBoost *native* sebesar 74,98%. Namun, CatBoost dengan *one-hot encoding* manual menghasilkan *sensitivity* dan *F1-score* tertinggi, masing-masing sebesar 74,63% dan 75,04%, yang menunjukkan kemampuan lebih baik dalam mengidentifikasi peserta PBI pada

threshold terpilih. Sebaliknya, XGBoost menghasilkan specificity tertinggi sebesar 78,54%. Dengan demikian, perbedaan kinerja terutama terlihat pada keseimbangan kemampuan mengenali kedua kelas.

Tabel 3. *Hyperparameter* dan *Threshold* Terpilih dari XGBoost dan CatBoost

Pengaturan	**CatBoost *Native***	**CatBoost *one-hot encoding* manual**	**XGBoost**
Iterasi	1	1	12
learning_rate/eta	0,1	0,1	0,07
depth/max_depth	5	5	4
l2_leaf_reg/lambda	3,6196	3,6196	15,1523
random_strength	3,3282	3,3282	-
bagging_temperature	4,4570	4,4570	-
rsm/colsample_bytree	0,8062	0,8062	0,7693
border_count	32	32	-
min_child_weight	-	-	15
subsample	-	-	0,6927
gamma	-	-	3,0901
alpha	-	-	3,9715
Best_nrounds	200	357	732
Threshold terpilih	0,51	0,47	0,5

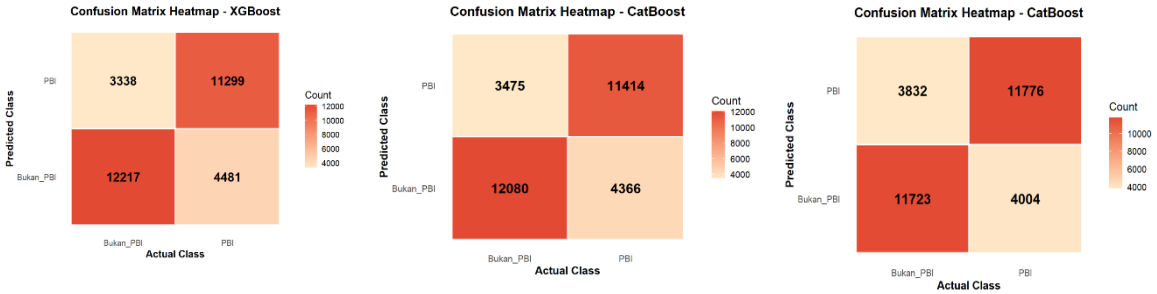
Catatan: Tanda "-" menunjukan parameter tidak digunakan pada model tersebut.

Tabel 4. Metrik Evaluasi Model XGBoost dan CatBoost

Metrik Evaluasi	XGBoost		CatBoost *native*		CatBoost *one-hot encoding* manual	
	Training	*Testing*	*Training*	*Testing*	*Training*	*Testing*
Accuracy	0,7594453	0,7504707	0,7611651	0,7497686	0,7669387	0,7499282
Sensitivity	0,7292187	0,7160330	0,7378029	0,7233207	0,7685809	0,7462611
Specificity	0,7899704	0,7854066	0,7847580	0,7765992	0,7652804	0,7536483
Precision	0,7780870	0,7719478	0,7758670	0,7666062	0,7678092	0,7544849
F1-Score	0,7528607	0,7429398	0,7563563	0,7443347	0,7681948	0,7503505
Balanced Accuracy	0,7595946	0,7507198	0,7612805	0,7499599	0,7669306	0,7499547
ROC-AUC	0,8335830	0,8215416	0,8382354	0,8206405	0,8453238	0,8215762

Selain metrik evaluasi di atas, digunakan juga nilai ROC-AUC untuk mengevaluasi model XGBoost dan CatBoost. ROC-AUC menunjukkan bahwa performa pada data uji cukup baik karena nilai AUC pada data uji tidak mengalami penurunan yang signifikan jika dibandingkan dengan nilai AUC pada data latih. Hal tersebut mengindikasikan tidak terjadinya *overfitting* yang kuat. Kedua model, baik XGBoost maupun CatBoost, memiliki nilai AUC testing ≥ 0,8, yang dapat disimpulkan bahwa model sudah bagus (Nahm, 2022). Untuk melihat lebih rinci mengenai jumlah observasi yang berhasil maupun gagal diklasifikasikan oleh model maka digunakan *confusion matrix*.

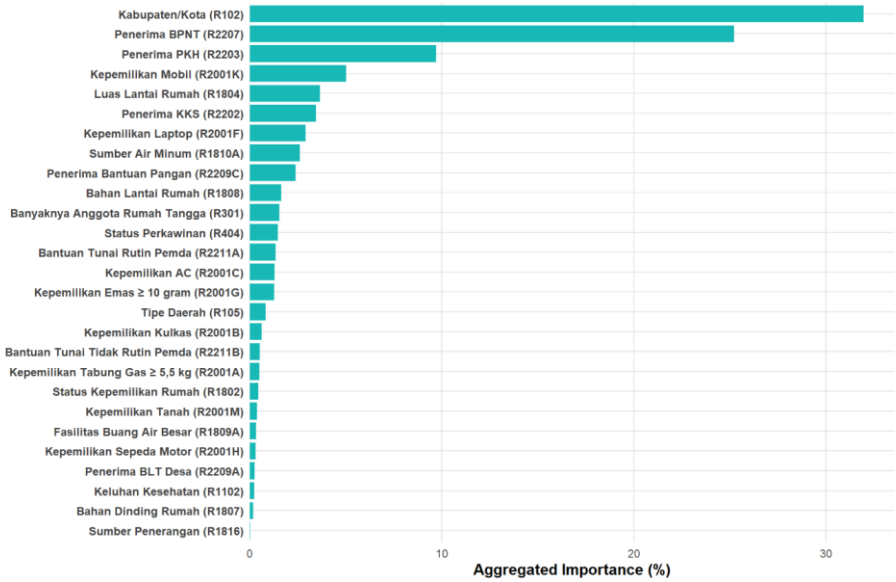
Gambar 5 menunjukkan bahwa dari 31.335 individu pada data uji, XGBoost mengklasifikasikan 12.217 individu Non-PBI dan 11.299 individu PBI secara benar. CatBoost *native* mengklasifikasikan 12.080 individu Non-PBI dan 11.414 individu PBI secara benar, sedangkan CatBoost dengan *one-hot encoding* manual masing-masing sebanyak 11.723 dan 11.776 individu. CatBoost dengan *one-hot encoding* manual menghasilkan kesalahan paling sedikit dalam mengenali peserta PBI, yaitu 4.004 individu, tetapi paling banyak mengklasifikasikan individu Non-PBI sebagai PBI, yaitu 3.832 individu. Dengan demikian, XGBoost lebih baik dalam mengenali kelas Non-PBI, sedangkan CatBoost dengan one-hot encoding manual lebih baik dalam mengenali kelas PBI pada threshold masing-masing model



Model XGBoost | Model CatBoost *Native* | Model CatBoost One-hot Encoding Manual

Gambar 5. *Confusion Matrix* dengan Model XGBoost dan CatBoost

Ketiga pendekatan dengan menggunakan ruang pencarian yang sama pada *hyperparameter* utama yang berpadanan menunjukkan kinerja klasifikasi yang relatif serupa, dengan selisih akurasi pada data uji kurang dari 0,1 poin persentase. Hal tersebut sejalan dengan penelitian Jaramillo et al. (2026), yang menerapkan kandidat nilai yang sama untuk kedua algoritma XGBoost dan CatBoost pada beberapa *hyperparameter* utama seperti jumlah pohon atau iterasi, laju pembelajaran, dan kedalaman pohon. Hasilnya memperoleh kinerja yang berdekatan, XGBoost dan CatBoost masing-masing menghasilkan AUC sebesar 0,952 dan 0,953, *recall* sebesar 89,2% dan 88,8%, serta *F1-score* sebesar 89,2% dan 89,5%. Pada penelitian ini, XGBoost memperoleh *accuracy, specificity, dan precision* tertinggi, sedangkan CatBoost dengan *one-hot encoding* manual menghasilkan *sensitivity* dan *F1-score* tertinggi. Dengan demikian, XGBoost sedikit unggul secara deskriptif dalam ketepatan klasifikasi keseluruhan, sementara CatBoost dengan *one-hot encoding* manual lebih mampu mengenali peserta PBI pada threshold terpilih. Untuk mengidentifikasi variabel yang berkontribusi terhadap hasil klasifikasi, dilakukan analisis *feature importance* dan SHAP menggunakan model XGBoost yang menunjukkan kinerja terbaik secara keseluruhan pada data pengujian.

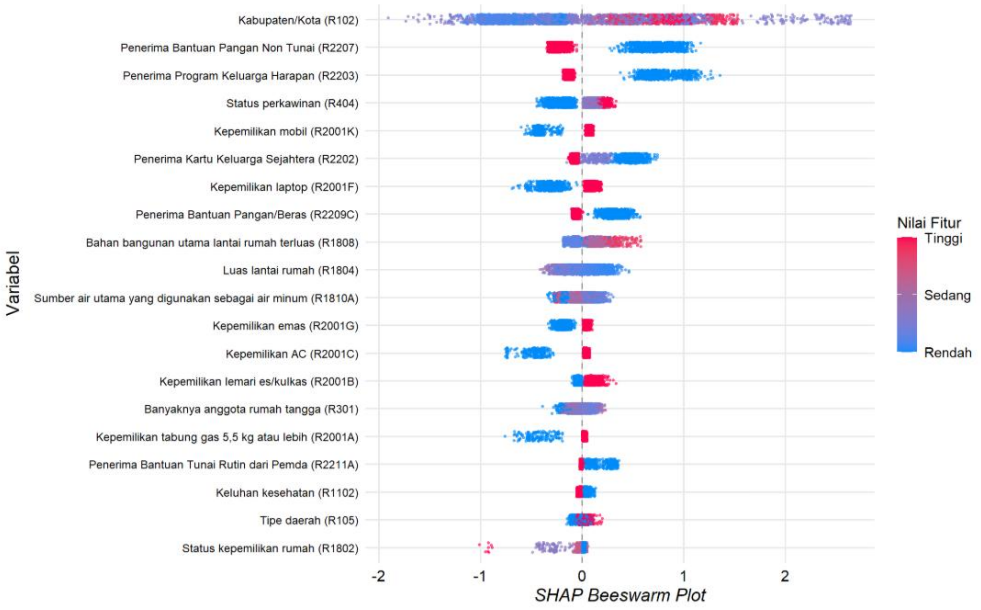


Gambar 6. *Feature Importance* Model XGBoost

Berdasarkan Gambar 6, variabel Kabupaten/Kota (R102) memiliki nilai *feature importance* tertinggi dibandingkan dengan variabel lainnya, hal ini selaras dengan Pasal 3

Peraturan Menteri Keuangan No. 51 tahun 2024 yang menjelaskan bahwa kontribusi pemerintah daerah dalam pembayaran iuran peserta PBI Jaminan Kesehatan dilakukan sesuai kapasitas fiskal daerah. Selanjutnya diikuti oleh variabel penerima BPNT (R2207), penerima PKH (R2203), kepemilikan mobil (R2001K), dan luas lantai rumah (R1804). Hasil ini menunjukkan bahwa faktor wilayah, kondisi tempat tinggal, karakteristik rumah tangga, serta penerimaan bantuan sosial memiliki peran penting dalam membedakan penerima dan bukan penerima BPJS PBI. Hal tersebut menunjukkan bahwa kepesertaan BPJS sangat dipengaruhi oleh kombinasi beberapa parameter sosial, ekonomi, dan penerimaan bantuan pemerintah (Kosasih *et al.*, 2022; Putri *et al*., 2023)

SHAP *beeswarm plot* pada Gambar 7 menunjukkan bahwa lokasi kabupaten/kota (R102) memiliki kontribusi terbesar terhadap prediksi kepesertaan BPJS PBI, dengan nilai SHAP yang tersebar pada sisi positif maupun negatif. Hal ini menunjukkan adanya perbedaan kontribusi antarwilayah terhadap prediksi model. Pada variabel BPNT (R2207) dan PKH (R2203), kategori penerima cenderung memberikan kontribusi positif terhadap skor prediksi model, sehingga meningkatkan kecenderungan individu untuk diklasifikasikan sebagai peserta PBI. Sebaliknya, pada kepemilikan mobil (R2001K), kategori memiliki mobil cenderung memberikan kontribusi negatif. Luas lantai rumah (R1804) yang lebih besar cenderung berkontribusi negatif, sedangkan jumlah anggota rumah tangga (R301) menunjukkan kontribusi yang relatif kecil dan terkonsentrasi di sekitar nol. Hal tersebut menunjukkan bahwa karakteristik sosial ekonomi rumah tangga berperan penting dalam menentukan status kepesertaan jaminan kesehatan seperti BPJS (Kosasih *et al*., 2022; Putri *et al*., 2023).



Gambar 7. SHAP *Beeswarm Plot* Model XGBoost

SHAP *beeswarm plot* pada Gambar 7 menunjukkan bahwa lokasi kabupaten/kota (R102) merupakan prediktor paling dominan dalam klasifikasi status kepesertaan BPJS PBI di Jawa Timur. Selain itu, penerimaan BPNT (R2207) dan PKH (R2203) berkontribusi terhadap peningkatan probabilitas prediksi PBI, sedangkan luas lantai rumah (R1804) dan jumlah anggota rumah tangga (R301) menunjukkan kontribusi yang cenderung menurunkan probabilitas prediksi PBI pada pola observasi tertentu. Berdasarkan temuan tersebut, pemerintah daerah perlu memperkuat pemutakhiran dan verifikasi data penerima PBI dengan mempertimbangkan karakteristik wilayah, kondisi tempat tinggal, serta keterkaitan data

penerimaan bantuan sosial. Hasil penelitian ini dapat menjadi informasi pendukung dalam evaluasi ketepatan sasaran program bantuan iuran kesehatan.

5. KESIMPULAN

Berdasarkan hasil penelitian, ketiga pendekatan, yaitu XGBoost, CatBoost *native*, dan CatBoost dengan *one-hot encoding* manual, menghasilkan kinerja klasifikasi yang relatif serupa dalam mengklasifikasikan status kepesertaan BPJS Penerima Bantuan Iuran (PBI) di Provinsi Jawa Timur. Pada data *testing*, XGBoost sedikit unggul dengan nilai *accuracy* sebesar 0,7505 dan *balanced accuracy* sebesar 0,7507. Sementara itu, CatBoost dengan *one-hot encoding* manual unggul pada *sensitivity* sebesar 0,7463 dan *F1-score* sebesar 0,7504. Hasil ini menunjukkan bahwa XGBoost memiliki ketepatan klasifikasi keseluruhan yang sedikit lebih baik, sedangkan CatBoost manual lebih baik dalam mengidentifikasi peserta PBI.

Berdasarkan analisis *feature importance* dan SHAP pada model XGBoost, variabel kabupaten/kota (R102) merupakan prediktor paling dominan dalam klasifikasi kepesertaan BPJS PBI. Variabel lain yang turut berkontribusi antara lain penerimaan BPNT (R2207), PKH (R2203), status perkawinan (R404), dan kepemilikan mobil (R2001K). Hasil SHAP menunjukkan bahwa kontribusi variabel dapat berbeda antarindividu, termasuk berdasarkan karakteristik wilayah dan kondisi sosial ekonomi rumah tangga. Temuan ini menunjukkan bahwa karakteristik wilayah, kondisi tempat tinggal, serta penerimaan bantuan sosial dapat menjadi informasi penting dalam membedakan status kepesertaan BPJS PBI. Oleh karena itu, hasil penelitian dapat digunakan sebagai informasi pendukung dalam evaluasi ketepatan sasaran program bantuan iuran kesehatan.

DAFTAR PUSTAKA

Agustina, R., Dartanto, T., Sitompul, R., Susiloretni, K. A., Suparmi, Achadi, E. L., Taher, A., Wirawan, F., Sungkar, S., Sudarmono, P., Shankar, A. H., & Thabrany, H. (2019). Universal health coverage in Indonesia: Concept, progress, and challenges. *The Lancet, 393*(10166), 75–102. https://doi.org/10.1016/S0140-6736(18)31647-7

Badan Pusat Statistik. (2019). *Poverty measurement in Indonesia*. BPS-Statistics Indonesia.

Badan Pusat Statistik Provinsi Jawa Timur. (2025). *Persentase Penduduk Miskin di Jawa Timur September 2024 Turun Menjadi 9,56 Persen*. Surabaya: BPS Provinsi Jawa Timur.

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Dinas Kesehatan Provinsi Jawa Timur. (2025). *Profil kesehatan Provinsi Jawa Timur Tahun 2024*. Dinas Kesehatan Provinsi Jawa Timur.

Efendi, M. A. N., & Suharjo, I. (2026). Perbandingan Algoritma XGBoost dan CatBoost dalam Prediksi Harga Rumah Berdasarkan Data Perumahan di Jabodetabek. *Jurnal Nasional Teknologi Komputer*, *6*(1), 51–63. https://doi.org/10.61306/jnastek.v6i1.345.

Handyanto, L., Arifin, W. A., & Syafri, H. (2026). Perbandingan Metode Hyperparameter Tuning Pada Model XGBoost Untuk Prediksi Multioutput. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, *10*(1), 50. https://doi.org/10.30872/jurti.v10i1.22958

Hartono, F. W., Muljono, & Fanani, A. Z. (2024). Improving the Accuracy of House Price Prediction using Catboost Regression with Random Search Hyperparameter Tuning:

A Comparative Analysis. *Advance Sustainable Science, Engineering and Technology*, *6*(3). https://doi.org/10.26877/asset.v6i3.602

Jaramillo, I.-F., Orozco-Iguasnia, W., Alcocer Quinteros, R. P., Villarroel-Molina, R. R., & Vilcacundo-Chiluisa, A. (2026). A machine learning-based predictive model for maintenance management of combustion engines in the agricultural sector. Algorithms, 19, Article 618. https://doi.org/10.3390/a19080618

Kementerian Keuangan Republik Indonesia. (2024). *Peraturan Menteri Keuangan Nomor 51 Tahun 2024 tentang Perubahan atas Peraturan Menteri Keuangan Nomor 78/PMK.02/2020.* Berita Negara Republik Indonesia Tahun 2024 Nomor 476.

Kementerian PPN/Bappenas. (2023). *Buku metadata indikator pembangunan tahun 2023*. Kementerian Perencanaan Pembangunan Nasional/Bappenas.

Kosasih, D. M., Adam, S., Uchida, M., Yamazaki, C., Koyama, H., & Hamazaki, K. (2022). Determinant factors behind changes in health-seeking behaviour before and after implementation of universal health coverage in Indonesia. *BMC Public Health*, *22*(1). https://doi.org/10.1186/s12889-022-13142-8

Lundberg, S. M., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. Section 2, 1–10.

Mulyo, H., & Zyen, A. K. (2025). Pengaruh Hyperparameter Tuning Gradient Boosting Terhadap Prediksi Pemilihan Program Studi Mahasiswa Baru. *Bulletin of Computer Science Research*, *5*(2), 131–137. https://doi.org/10.47065/bulletincsr.v5i2.454

Nahm, F. S. (2022). Receiver operating characteristic curve: overview and practical use for clinicians. *Korean Journal of Anesthesiology*, *75*(1), 25–36. https://doi.org/10.4097/kja.21209

Pemerintah Republik Indonesia. (2012). *Peraturan Pemerintah Republik Indonesia Nomor 101 Tahun 2012 tentang Penerima Bantuan Iuran Jaminan Kesehatan*. Jakarta: Sekretariat Negara.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. Section 4, 1–11.

Putri, N. K., Laksono, A. D., & Rohmah, N. (2023). Predictors of national health insurance membership among the poor with different education levels in Indonesia. *BMC Public Health*, *23*(1). https://doi.org/10.1186/s12889-023-15292-9

Ray, A., Li, X., Chen, Y., Yang, X., Zhang, W., & Hui, D. (2026). Predicting Biochar-Induced Changes in Soil Organic Carbon With Ensemble Machine Learning. *GCB Bioenergy*, *18*(5). https://doi.org/10.1111/gcbb.70112

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics, 17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003