

KONSISTENSI METODE *INTERPRETABLE MACHINE LEARNING* DAN IMPLEMENTASINYA UNTUK EKSPLORASI FAKTOR RISIKO GANGGUAN MENTAL

AMRI LUTHFI NAJIH



PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Konsistensi Metode *Interpretable Machine Learning* dan Implementasinya untuk Eksplorasi Faktor Risiko Gangguan Mental” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Amri Luthfi Najih
M0501241057

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

RINGKASAN

AMRI LUTHFI NAJIH. Konsistensi Metode *Interpretable Machine Learning* dan Implementasinya untuk Eksplorasi Faktor Risiko Gangguan Mental. Dibimbing oleh BAGUS SARTONO dan SEPTIAN RAHARDIANTORO.

Model machine learning berbasis *gradient boosting*, seperti XGBoost dan LightGBM, banyak digunakan karena memiliki kemampuan prediksi yang tinggi pada berbagai jenis data. Namun, kedua model tersebut bersifat kompleks sehingga hasil prediksinya sulit dijelaskan secara langsung. Penelitian ini menggunakan *SHapley Additive exPlanations* (SHAP) sebagai metode interpretable machine learning untuk mengukur kontribusi peubah penjelas terhadap prediksi. Meskipun SHAP banyak digunakan untuk menghasilkan peringkat kepentingan peubah penjelas, perlu dianalisis sejauh mana peringkat tersebut tetap stabil dan sesuai ketika diterapkan pada data dengan karakteristik yang berbeda, khususnya pada kondisi proporsi kelas minoritas yang sangat tidak seimbang, korelasi antarpeubah yang tinggi, dan jumlah peubah penjelas yang bervariasi. Oleh karena itu, penelitian ini bertujuan mengevaluasi stabilitas peringkat kepentingan peubah penjelas berbasis SHAP pada model XGBoost dan LightGBM, serta menerapkannya untuk eksplorasi faktor risiko gangguan mental menggunakan data klaim BPJS Kesehatan.

Penelitian dilakukan melalui dua pendekatan, yaitu studi simulasi dan analisis data empiris. Pada studi simulasi, data dibangkitkan secara terkontrol berdasarkan kombinasi tiga faktor, yaitu jumlah peubah penjelas sebesar 7, 15, 25, dan 35; korelasi antarpeubah sebesar 0; 0,2; 0,5; dan 0,95; serta proporsi kelas minoritas sebesar 0,5; 0,3; 0,1; dan 0,01. Kombinasi tersebut menghasilkan 64 skenario data simulasi. Pada setiap skenario, model XGBoost dan LightGBM dibangun secara berulang sebanyak 100 kali. Peubah penjelas yang benar-benar berpengaruh terhadap peubah respon ditetapkan melalui fungsi pembentuk data, yaitu x_1 , x_2 , x_3 , x_6 , dan x_7 , sehingga hasil peringkat SHAP dapat dibandingkan dengan peringkat peubah penjelas sebenarnya. Stabilitas peringkat peubah penjelas dievaluasi menggunakan *Sequential Rank Agreement* (SRA).

Hasil simulasi menunjukkan bahwa SHAP mampu mengidentifikasi peubah penjelas dengan kontribusi kuat secara konsisten, terutama x_1 , x_2 , x_3 , dan x_6 . Namun, peubah penjelas dengan kontribusi sangat kecil seperti x_7 lebih sulit dibedakan dari peubah penjelas yang tidak berpengaruh. Stabilitas dan kesesuaian peringkat menurun secara nyata ketika data memiliki korelasi antarpeubah yang sangat tinggi dan proporsi kelas minoritas yang ekstrem. Kondisi korelasi antarpeubah 0,95 dan proporsi kelas minoritas 0,01 merupakan kombinasi yang paling menurunkan kesesuaian peringkat SHAP. Hasil Uji Friedman menunjukkan adanya perbedaan signifikan pada kesesuaian peringkat di antara 128 kombinasi kondisi model dan struktur data. Uji *Wilcoxon* yang memfokuskan perbedaan antartaraf pada setiap perlakuan menunjukkan bahwa korelasi rendah 0 belum memberikan perubahan yang konsisten, sedangkan korelasi sangat tinggi 0,95 menjadi faktor yang paling kuat menurunkan kesesuaian peringkat. Peningkatan jumlah peubah penjelas juga dapat memengaruhi peringkat, tetapi pengaruhnya tidak seragam dan bergantung pada kombinasi model, korelasi, dan proporsi kelas minoritas.

Pada data empiris, penelitian menggunakan data klaim BPJS Kesehatan periode 2018–2023 yang terdiri atas 586.733 peserta. Dari jumlah tersebut, 7.826 peserta atau 1,33% teridentifikasi pernah masuk kelompok INACBGs “*Mental Health and Behavioral Groups*”, sedangkan 578.907 peserta lainnya tidak teridentifikasi mengalami gangguan mental. Kondisi ini menunjukkan ketidakseimbangan kelas yang sangat ekstrem dengan rasio sekitar 1:74,5. Untuk menangani masalah tersebut, digunakan metode *oversampling* SMOTE dan ADASYN pada data latih. Hasil pemodelan menunjukkan bahwa metode *oversampling* mampu meningkatkan kemampuan model dalam mendeteksi kelas minoritas. Kombinasi LightGBM dengan SMOTE menghasilkan *balanced accuracy* sebesar 0,8139 dan *sensitivity* sebesar 0,6377 pada data uji, sedangkan LightGBM dengan ADASYN menghasilkan *balanced accuracy* sebesar 0,8058 dan *sensitivity* sebesar 0,6209. Nilai tersebut lebih tinggi dibandingkan model tanpa penanganan ketidakseimbangan data.

Selain meningkatkan kinerja prediksi, penanganan ketidakseimbangan data juga meningkatkan stabilitas interpretasi peubah penjelas. Pada model LightGBM, ADASYN menghasilkan rata-rata SRA paling kecil, yaitu 0,5350, diikuti SMOTE sebesar 0,5950. Sebaliknya, model LightGBM tanpa penanganan ketidakseimbangan data menghasilkan rata-rata SRA sebesar 2,2534, yang menunjukkan peringkat peubah penjelas paling tidak stabil. Interpretasi SHAP pada kombinasi LightGBM dengan SMOTE menunjukkan bahwa peubah penjelas paling berkontribusi terhadap prediksi gangguan mental adalah jumlah anggota keluarga yang pernah mengakses poli jiwa, jenis kelamin, jenis fasilitas kesehatan, kelas rawat, segmen peserta, jumlah rawat jalan, hubungan keluarga, status perkawinan, jumlah anggota keluarga, dan akses poli anestesi.

Secara keseluruhan, penelitian ini menunjukkan bahwa interpretasi kepentingan peubah penjelas berbasis SHAP tidak hanya dipengaruhi oleh model yang digunakan, tetapi juga oleh struktur data. Korelasi antarpeubah yang sangat tinggi dan ketidakseimbangan kelas ekstrem dapat menurunkan stabilitas serta kesesuaian peringkat peubah penjelas. Oleh karena itu, evaluasi stabilitas interpretasi perlu dilakukan sebelum hasil SHAP digunakan sebagai dasar pengambilan keputusan. Pada data empiris BPJS, penanganan ketidakseimbangan data terbukti tidak hanya meningkatkan kemampuan deteksi kelas minoritas, tetapi juga menghasilkan interpretasi peubah penjelas yang lebih stabil dan lebih dapat diandalkan.

Kata kunci: Kepentingan peubah penjelas, LightGBM, SHAP, SRA, XGBoost.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

SUMMARY

AMRI LUTHFI NAJIH. Consistency of Interpretable Machine Learning Methods and Their Application to Mental Health Risk Factor Exploration. Supervised by BAGUS SARTONO and SEPTIAN RAHARDIANTORO.

Machine learning models based on gradient boosting, such as XGBoost and LightGBM, are widely used because they provide strong predictive performance across various types of data. However, these models have complex structures, which makes their predictions difficult to interpret directly. To address this issue, this study uses SHapley Additive exPlanations (SHAP) as an interpretable machine learning method to measure the contribution of each predictor to model predictions. Although SHAP is commonly used to produce importance rankings of predictors, the extent to which these rankings remain stable and aligned with the true predictor structure under different data conditions still needs to be evaluated systematically. This issue is particularly important when the data have a highly imbalanced proportion of the minority class, strong correlation among predictors, and different numbers of predictors. Therefore, this study aims to evaluate the stability of SHAP importance rankings of predictors in XGBoost and LightGBM models, and to apply this approach to explore risk factors for mental disorders using BPJS Kesehatan claims data.

This study was conducted using two approaches: a simulation study and an empirical data analysis. In the simulation study, data were generated under controlled conditions based on combinations of three factors: the number of predictors, set at 7, 15, 25, and 35; the correlation among predictors, set at 0, 0.2, 0.5, and 0.95, and the proportion of the minority class, set at 0.5, 0.3, 0.1, and 0.01. These combinations produced 64 simulation scenarios. For each scenario, XGBoost and LightGBM models were trained repeatedly with 100 replications. The predictors that truly influenced the response variable were specified through the function used to generate the data, namely x_1 , x_2 , x_3 , x_6 , and x_7 . This design made it possible to compare the rankings produced by SHAP with the true predictor rankings. The stability of predictor rankings was evaluated using Sequential Rank Agreement (SRA).

The simulation results show that SHAP consistently identified predictors with strong contributions, particularly x_1 , x_2 , x_3 , and x_6 . However, a predictor with a very weak contribution, such as x_7 , was more difficult to distinguish from predictors with no influence. The stability and agreement of the rankings declined substantially when very strong correlation among predictors occurred together with extreme class imbalance. The combination of a predictor correlation level of 0.95 and a minority class proportion of 0.01 produced the largest decline in the agreement of SHAP rankings. The Friedman test indicated significant differences in ranking agreement across the 128 combinations of model and data conditions. The Wilcoxon signed rank tests, which focused on differences between levels within each factor, showed that a low correlation level of 0.2 did not lead to consistent changes, whereas a very high correlation level of 0.95 was the strongest condition associated with reduced ranking agreement. Increasing the number of predictors also affected the rankings, but the effect varied across conditions

depending on the model, the level of predictor correlation, and the proportion of the minority class.

For the empirical analysis, this study used BPJS Kesehatan claims data from 2018 to 2023, consisting of 586,733 participants. Among them, 7,826 participants, or 1.33 percent, were identified as having been included in the INACBGs “Mental Health and Behavioral Groups,” while the remaining 578,907 participants were not identified as having mental disorders. This distribution indicates an extremely imbalanced classification problem, with an approximate ratio of 1:74.5 between the minority and majority classes. To address this problem, the SMOTE and ADASYN oversampling methods were applied to the training data. The modeling results show that oversampling improved the ability of the models to detect the minority class. The combination of LightGBM and SMOTE achieved a balanced accuracy of 0.8139 and a sensitivity of 0.6377 on the test data, while the combination of LightGBM and ADASYN achieved a balanced accuracy of 0.8058 and a sensitivity of 0.6209. These values were higher than those obtained from models trained without handling class imbalance.

In addition to improving predictive performance, handling class imbalance also improved the stability of predictor interpretation. In the LightGBM model, ADASYN produced the lowest average SRA value, at 0.5350, followed by SMOTE at 0.5950. In contrast, LightGBM without class imbalance handling produced an average SRA value of 2.2534, indicating the least stable predictor rankings. The SHAP interpretation of the model combining LightGBM and SMOTE showed that the most influential predictors in predicting mental disorders were the number of family members who had accessed psychiatric outpatient services, sex, type of primary health facility, treatment class, participant segment, number of outpatient visits, family relationship status, marital status, number of family members, and access to anesthesiology services.

Overall, this study shows that the interpretation of predictor importance using SHAP is influenced not only by the machine learning model but also by the underlying structure of the data. Very strong correlation among predictors and extreme class imbalance can reduce both the stability and the agreement of predictor rankings. Therefore, the stability of SHAP interpretations should be evaluated before the results are used as a basis for decision making. In the empirical BPJS data, handling class imbalance not only improved the detection of the minority class but also produced predictor interpretations that were more stable and more reliable.

Keywords: Importance rankings of predictors, LightGBM, SHAP, SRA, XGBoost.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

KONSISTENSI METODE *INTERPRETABLE MACHINE LEARNING* DAN IMPLEMENTASINYA UNTUK EKSPLORASI FAKTOR RISIKO GANGGUAN MENTAL

AMRI LUTHFI NAJIH

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Magister Statistika dan Sains Data

**PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Tesis : Konsistensi Metode *Interpretable Machine Learning* dan Implementasinya untuk Eksplorasi Faktor Risiko Gangguan Mental

Nama : Amri Luthfi Najih

NIM : M0501241057

Disetujui oleh

Pembimbing 1:
Dr. Bagus Sartono, S.Si., M.Si.



Pembimbing 2:
Dr. Septian Rahardiantoro, S.Stat., M.Si.



Diketahui oleh

Ketua Program Studi:
Dr. Agus Mohamad Soleh, S.Si., MT.
NIP 197503151999031004



Dekan Sekolah Sains Data, Matematika, Dan Informatika:
Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.
NIP 196607021993021001



Tanggal Ujian: 21 Juli 2026

Tanggal Pengesahan:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak Agustus 2025 sampai bulan Maret 2026 ini ialah *Interpretable Machine Learning*, dengan judul “Konsistensi Metode *Interpretable Machine Learning* dan Implementasinya untuk Eksplorasi Faktor Risiko Gangguan Mental”. Penelitian ini penulis susun sebagai salah satu syarat untuk menyelesaikan studi Magister Sains pada Program Studi Statistika dan Sains Data. Penulisan tesis ini dapat terselesaikan tidak terlepas dari dukungan, bimbingan, dan bantuan berbagai pihak sehingga penulis menyampaikan ucapan terima kasih kepada:

1. Dr. Bagus Sartono, S.Si., M.Si. dan Dr. Septian Rahardiantoro, S.Stat., M.Si. selaku komisi pembimbing yang senantiasa bersedia meluangkan waktunya untuk memberikan bimbingan, masukan, dukungan, dan arahan kepada penulis;
2. Dr. Agus Mohamad Soleh, S.Si., MT. selaku Ketua Program Studi dan Prof. Dr. Ir. I Made Sumertajaya, M.Si. selaku penguji luar komisi atas masukan dan saran yang sangat bermanfaat untuk menyempurnakan tesis ini.
3. Kedua orang tua penulis Bapak Suhamto dan Alm. Ibu Anis Husainiyah atas dukungan dan doa sejak dari dalam kandungan hingga saat ini yang tak pernah putus.
4. Istri tercinta Afifah dan kedua anak penulis Ali dan Nizam yang selalu mendukung penulis dan selalu memberikan semangat untuk menyelesaikan studi Magister ini.
5. Seluruh dosen pengajar dan staf di program studi magister statistika dan sains data atas pengajaran, bimbingan, serta bantuan teknis dan nonteknis selama menempuh studi.
6. Rekan-rekan seperjuangan program studi magister statistika dan sains data tahun 2024 kelas reguler atas semangat, kekompakan, dan kebersamaan dalam berjuang serta bantuan kepada penulis baik secara langsung maupun tidak langsung.

Penulis mohon maaf atas kesalahan dan kekurangan dalam tesis ini. Semoga tesis ini memberikan manfaat kepada pihak terkait dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2026

Amri Luthfi Najih

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	4
1.4 Manfaat	4
II TINJAUAN PUSTAKA	5
2.1 Penanganan Data Tidak Seimbang	5
2.2 <i>Gradient Boosting</i>	8
2.3 <i>Extreme Gradient Boosting (XGBoost)</i>	9
2.4 <i>Light Gradient Boosting Machine (LightGBM)</i>	11
2.5 <i>Shapley Additive Explanations (SHAP)</i>	12
2.6 <i>Sequential Rank Agreement (SRA)</i>	13
2.7 Analisis Nonparametrik untuk Membandingkan Kesesuaian Peringkat	16
III Metode	18
3.1 Data	18
3.2 Data Simulasi	18
3.3 Data Klaim BPJS	19
3.4 Prosedur Analisis	20
IV HASIL DAN PEMBAHASAN	22
4.1 Kajian Simulasi Data untuk melihat performa SHAP	22
4.2 Analisis Nonparametrik untuk Mengevaluasi Perbedaan Kesesuaian Peringkat Berdasarkan Model, Jumlah Peubah Penjelas, Korelasi Antarpeubah Penjelas, dan Proporsi Kelas Minoritas	28
4.3 Kesimpulan Hasil Pemeringkatan Peubah Penjelas oleh SHAP	34
4.4 Statistik Deskriptif Data Klaim BPJS Tahun 2018-2023	35
4.5 Hasil pemodelan prediksi kasus gangguan mental	39
4.6 Interpretasi model prediksi kasus gangguan mental	44
V SIMPULAN DAN SARAN	50
5.1 Simpulan	50
5.2 Saran	51
DAFTAR PUSTAKA	52
LAMPIRAN	57
RIWAYAT HIDUP	61

DAFTAR TABEL

1	Faktor yang memengaruhi Gangguan Mental Pada Penelitian Sebelumnya	19
2	Statistik deskriptif peubah numerik	37
3	Statistik deskriptif peubah penjelas kategorik	38
4	Statistik deskriptif peubah penjelas kategorik (lanjutan)	39
5	Rata-rata nilai ukuran kinerja model pada 100 ulangan	39
6	Rata-rata nilai SRA seluruh himpunan peringkat teratas	44

DAFTAR GAMBAR

1	Ilustrasi data tidak seimbang	5
2	Pohon keputusan XGBoost	10
3	Pohon keputusan LightGBM	11
4	Ilustrasi Perhitungan SRA	15
5	Alur Penelitian	21
6	Distribusi peringkat SHAP untuk seluruh kombinasi kondisi data (a) jumlah peubah 7, (b) jumlah peubah 15, (c) jumlah peubah 25, dan (d) jumlah peubah 35.	23
7	Rata-rata peringkat kondisi data yang memiliki (a) korelasi 0 dan (b) korelasi 0,2	24
8	Rata-rata peringkat kondisi data yang memiliki (a) korelasi 0,5; (b) korelasi 0,95.	25
9	Persentase ketepatan peringkat peubah penjelas x_1 , x_2 , x_3 , x_6 , dan x_7 pada 100 ulangan model LightGBM.	26
10	Nilai SRA pada himpunan peringkat teratas ke-10 (10 peubah penjelas peringkat tertinggi) pada kondisi data dengan 35 peubah penjelas dengan model LightGBM.	27
11	Median korelasi spearman pada semua kombinasi kondisi data	30
12	Uji lanjut Interaksi Model dengan korelasi antarpeubah penjelas	31
13	Uji lanjut Interaksi Model dengan Proporsi kelas minoritas	31
14	Uji lanjut interaksi jumlah peubah penjelas dengan korelasi antarpeubah penjelas	32
15	Uji lanjut Interaksi model XGBoost dan LightGBM	33
16	Perbandingan Kelas peubah respon (Gangguan Mental)	36
17	Distribusi data peubah penjelas numerik	38
18	Perbandingan hasil pemodelan pada masing-masing metode	40
19	Distribusi peringkat SHAP pada masing-masing peubah penjelas menggunakan model XGBoost	41
20	Distribusi peringkat SHAP pada masing-masing peubah penjelas menggunakan model LightGBM	42
21	Skor <i>Sequential Rank Agreement</i> (SRA)	43
22	Grafik (a) Nilai kepentingan peubah penjelas dan (b) arah prediksi model.	46

- 23 Contoh interpretasi lokal dari dua sampel yang berbeda, (a) prediksi tidak gangguan mental, (b) prediksi gangguan mental. 48

DAFTAR LAMPIRAN

- | | | |
|---|--|----|
| 1 | Lampiran 1 Keterangan peubah penjelas data BPJS | 58 |
| 2 | Lampiran 2 Persentase ketepatan peringkat peubah penjelas X1, X2, X3, X6, dan X7 pada 100 ulangan model XGBoost. | 59 |
| 3 | Lampiran 3 <i>Dependence plot</i> sepuluh peubah penjelas paling penting | 60 |





@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.