UNIVERSITAS
GADJAH MADA
**Identifikasi Faktor Prediktor Sindroma Metabolik pada Populasi Indonesia: Analisis Pembelajaran Mesin dari Indonesia Family Life Survey (IFLS)**
DHITE BAYU NUGROHO, dr. Hemi Sinorita, SpPD, K-Emd.; dr. Heni Retnowulan, SpPD, K-P
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# INTISARI

Indonesia saat ini sedang menghadapi transisi epidemiologi dengan pergeseran beban penyakit dari penyakit menular ke penyakit tidak menular (PTM), di mana penyakit kardiovaskular dan diabetes melitus menjadi penyumbang utama mortalitas. Salah satu entitas klinis yang merepresentasikan akumulasi risiko kardiometabolik tersebut adalah Sindroma Metabolik (MetS). Deteksi dini MetS sangat krusial, namun saat ini masih sangat bergantung pada pemeriksaan laboratorium invasif yang tidak selalu mudah diakses oleh seluruh populasi. Selain itu, metode statistik konvensional memiliki keterbatasan dalam menangkap interaksi kompleks antarvariabel multidimensional yang memengaruhi kondisi ini.

Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor prediktor Sindroma Metabolik pada populasi dewasa Indonesia serta membangun model prediksi berbasis *Machine Learning* menggunakan dataset berskala nasional. Secara khusus, studi ini berupaya menggali prediktor relevan dari dimensi biomedis, fungsi fisik, kognitif, dan psikososial, serta mengevaluasi performa model prediktif yang dibangun.

Metode penelitian yang digunakan adalah kuantitatif analitik dengan desain potong lintang (*cross-sectional*) memanfaatkan data sekunder dari *Indonesian Family Life Survey* (IFLS) gelombang ke-5. Analisis dilakukan terhadap 2.931 subjek yang memiliki kelengkapan data multidimensional. Model prediksi dibangun menggunakan algoritma *Random Forest* yang dievaluasi melalui pengukuran diskriminasi (*Area Under the Curve*/AUC), akurasi, sensitivitas, dan spesifisitas.

Hasil penelitian menunjukkan prevalensi MetS sebesar 20,5% (600/2.931). Model *Random Forest* memiliki kemampuan diskriminasi yang baik dengan nilai AUC 0,836, akurasi 0,797, dan spesifisitas 0,875. Analisis *variable importance* menunjukkan bahwa status merokok merupakan prediktor dengan kontribusi tertinggi, diikuti kekuatan genggam tangan, kadar CRP, status pernikahan, pola makan tradisional, dan kapasitas paru yang keseluruhannya melampaui beberapa faktor klinis konvensional. Pada analisis univariat, jenis kelamin laki-laki menunjukkan asosiasi terkuat (OR 13,5; IK 95% 10,7–17,4; $p < 0,001$), diikuti riwayat penyakit jantung (OR 4,24; IK 95% 2,03–8,94; $p < 0,001$) dan tingkat pendidikan perguruan tinggi (OR 4,08; IK 95% 3,02–5,51; $p < 0,001$). Status merokok (pernah merokok) juga menunjukkan asosiasi bermakna dengan kejadian MetS (OR 5,0; $p < 0,001$), sementara usia dan kadar CRP tidak menunjukkan hubungan yang signifikan secara statistik. Temuan ini menunjukkan bahwa pendekatan *Machine Learning* berbasis data multidimensional dapat menjadi alat bantu yang valid untuk deteksi dini MetS pada populasi Indonesia, khususnya dalam konteks layanan kesehatan dengan keterbatasan akses pemeriksaan laboratorium.

UNIVERSITAS
GADJAH MADA
**Identifikasi Faktor Prediktor Sindroma Metabolik pada Populasi Indonesia: Analisis Pembelajaran Mesin dari Indonesia Family Life Survey (IFLS)**
DHITE BAYU NUGROHO, dr. Hemi Sinorita, SpPD, K-Emd.; dr. Heni Retnowulan, SpPD, K-P
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

Kata kunci: Sindroma Metabolik, *Machine Learning*, IFLS, *Random Forest*, Prediksi Risiko.

UNIVERSITAS
GADJAH MADA
**Identifikasi Faktor Prediktor Sindroma Metabolik pada Populasi Indonesia: Analisis Pembelajaran Mesin dari Indonesia Family Life Survey (IFLS)**
DHITE BAYU NUGROHO, dr. Hemi Sinorita, SpPD, K-Emd.; dr. Heni Retnowulan, SpPD, K-P
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# ABSTRACT

*Indonesia is currently undergoing an epidemiological transition marked by a shift in disease burden from communicable to non-communicable diseases (NCDs), with cardiovascular disease and diabetes mellitus as the leading contributors to mortality. Metabolic Syndrome (MetS) represents one key clinical entity that encapsulates the accumulation of cardiometabolic risk. Early detection of MetS is critical; however, it currently remains heavily reliant on invasive laboratory examinations that are not consistently accessible across the entire population. Furthermore, conventional statistical methods are limited in their capacity to capture the complex interactions among the multidimensional variables that influence this condition.*

*This study aimed to identify predictive factors for Metabolic Syndrome in the adult Indonesian population and to develop a Machine Learning-based prediction model using a national-scale dataset. Specifically, the study sought to explore relevant predictors across biomedical, physical function, cognitive, and psychosocial dimensions, and to evaluate the performance of the constructed predictive models.*

*A quantitative analytic approach with a cross-sectional design was employed, utilizing secondary data from the fifth wave of the Indonesian Family Life Survey (IFLS-5). Analysis was conducted on 2,931 subjects with complete multidimensional data. Prediction models were developed using Random Forest algorithm and evaluated through discrimination metrics including Area Under the Curve (AUC), accuracy, sensitivity, and specificity.*

*The study found a MetS prevalence of 20.5% (600/2,931). The Random Forest model demonstrated good discriminative ability with an AUC of 0.836, accuracy of 0.797, and specificity of 0.875. Variable importance analysis identified smoking status as the highest-contributing predictor, followed by grip strength, CRP level, marital status, traditional dietary pattern, and lung capacity — all of which surpassed several conventional clinical factors in predictive importance. In univariate analysis, male sex demonstrated the strongest association (OR 13.5; 95% CI 10.7–17.4; $p < 0.001$), followed by history of heart disease (OR 4.24; 95% CI 2.03–8.94; $p < 0.001$) and tertiary education level (OR 4.08; 95% CI 3.02–5.51; $p < 0.001$). Smoking history was also significantly associated with MetS occurrence (OR 5.0; $p < 0.001$), while age and CRP level did not demonstrate statistically significant associations. These findings suggest that a multidimensional data-driven Machine Learning approach represents a valid screening tool for early MetS detection in the Indonesian population, particularly in healthcare settings with limited access to laboratory examinations.*

*Keywords: Metabolic Syndrome, Machine Learning, IFLS, Random Forest, Risk Prediction.*