

PENERAPAN ALGORITMA RANDOM FOREST DALAM PENCARIAN POLA KLAIM BPJS RUMAH SAKIT UMUM KUMALA SIWI MIJEN KUDUS

Oleh:

¹Irwan Syah, ²Sutono, ³Mulyanto

^{1,2,3}Universitas Muria Kudus

Jl. Lingkar Utara UMK, Gondangmanis, Bae, Kudus, Jawa Tengah 59327

e-mail: irwansyahsipa2025@gmail.com¹, sutono@umk.ac.id², mulyanto@umk.ac.id³

ABSTRACT

The persistently high number of pending BPJS Kesehatan claims in hospitals impacts cash flow and financial risk, necessitating data-driven solutions to improve the effectiveness of claims management. This study aims to analyze the factors influencing BPJS claims status and test the performance of the Random Forest algorithm in classifying eligible and pending claims. The results show that feature analysis identified total hospital charges and LOS as the most dominant variables, followed by diagnosis and treatment codes. The Random Forest model, with 30 decision trees, achieved 96.88% accuracy with balanced precision and recall, despite imbalanced data (520 eligible claims and 24 pending). Claim patterns indicate that claims with high costs, long LOS, and complex diagnoses and treatments are at higher risk of delay. This study demonstrates Random Forest's technical superiority and managerial benefits as an early warning system, supporting data-driven decisions, and improving the efficiency of claims management and hospital cash flow.

Keywords: *Random Forest, BPJS Claims, Hospital Management, Data Mining*

ABSTRAK

Masih tingginya klaim pending BPJS Kesehatan di rumah sakit berdampak pada arus kas dan risiko finansial, sehingga diperlukan solusi berbasis data untuk meningkatkan efektivitas pengelolaan klaim. Penelitian ini bertujuan menganalisis faktor-faktor yang memengaruhi status klaim BPJS serta menguji kinerja algoritma Random Forest dalam klasifikasi klaim layak dan pending. Hasil penelitian menunjukkan Analisis feature importance mengidentifikasi total tarif RS dan LOS sebagai variabel paling dominan, diikuti diagnosis dan kode tindakan. Model Random Forest dengan 30 pohon keputusan mencapai akurasi 96,88% dengan precision dan recall seimbang, meskipun data tidak seimbang (520 klaim layak dan 24 pending). Pola klaim mengindikasikan bahwa klaim dengan biaya tinggi, LOS panjang, serta diagnosis dan tindakan kompleks lebih berisiko pending. Penelitian ini membuktikan Random Forest tidak hanya unggul secara teknis, tetapi juga bermanfaat secara manajerial sebagai sistem peringatan dini, mendukung keputusan berbasis data, serta meningkatkan efisiensi pengelolaan klaim dan arus kas rumah sakit.

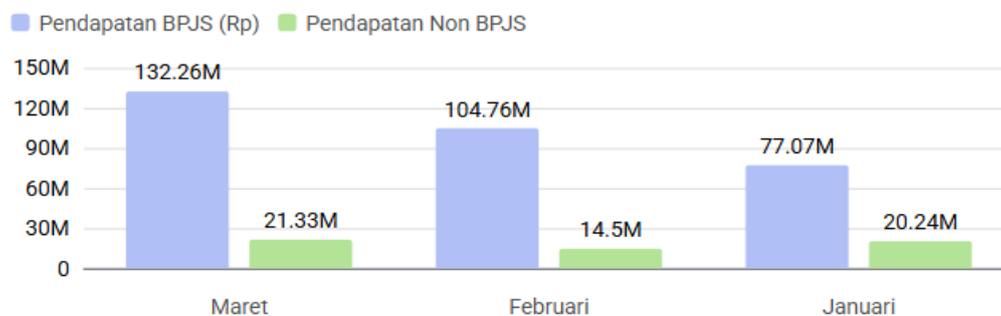
Kata Kunci: *Random Forest, Klaim BPJS, Manajemen Rumah Sakit, Data Mining*

PENDAHULUAN

Rumah sakit merupakan institusi pelayanan kesehatan yang memiliki peran penting dalam sistem kesehatan nasional, baik sebagai tempat penyelenggaraan pelayanan kuratif maupun promotif dan preventif. Peningkatan jumlah rumah sakit di Indonesia dari tahun ke tahun menunjukkan tingginya permintaan masyarakat terhadap layanan kesehatan yang berkualitas. Berdasarkan data Kementerian Kesehatan, pada tahun 2023 jumlah rumah sakit di Indonesia telah mencapai 3.155 unit, meningkat dibandingkan tahun-tahun sebelumnya. Pertumbuhan ini tidak hanya dipengaruhi oleh peningkatan jumlah penduduk, urbanisasi, dan transisi epidemiologi penyakit, tetapi juga oleh perkembangan teknologi kedokteran serta kebijakan pemerintah dalam memperluas akses pelayanan kesehatan melalui program Jaminan Kesehatan Nasional (JKN).

Sejak diberlakukannya JKN yang dikelola oleh Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan pada tahun 2014, hampir seluruh rumah sakit di Indonesia sangat bergantung pada sistem klaim sebagai sumber utama pendapatan. Melalui sistem ini, biaya pelayanan pasien ditanggung BPJS sesuai dengan mekanisme tarif INA-CBGs (Indonesian Case-Based Groups). Walaupun sistem ini dimaksudkan untuk meningkatkan efisiensi, transparansi, dan keadilan pembiayaan kesehatan, dalam praktiknya masih terdapat berbagai kendala, terutama terkait keterlambatan pembayaran klaim dan adanya klaim pending yang menyebabkan rumah sakit harus menanggung beban finansial dalam jangka waktu tertentu. Klaim pending umumnya terjadi karena perbedaan pencatatan diagnosis, kesalahan pengkodean tindakan medis, inkonsistensi dokumen administratif, atau ketidaksesuaian dengan aturan BPJS.

Perbandingan Pendapatan BPJS vs Non-BPJS per Bulan



Gambar 1. Perbandingan Pendapatan BPJS vs Non-BPJS Per Bulan

Kondisi tersebut berimplikasi langsung terhadap manajemen keuangan rumah sakit. Klaim yang tertunda dapat mengganggu stabilitas arus kas, menghambat pengadaan obat, logistik, hingga pembayaran tenaga kesehatan. Rumah Sakit Kumala Siwi, sebagai salah satu penyedia layanan kesehatan yang bekerja sama dengan BPJS, menghadapi tantangan serupa. Data awal tahun 2025 menunjukkan bahwa meskipun arus kas rumah sakit relatif positif, masih terdapat klaim pending dengan nilai signifikan yang berpotensi mengurangi likuiditas dan mengganggu operasional jangka pendek. Hal ini menegaskan bahwa pengelolaan klaim bukan hanya masalah administratif, tetapi juga menyangkut keberlangsungan pelayanan kesehatan.

Di sisi lain, ketersediaan data klaim dalam jumlah besar sebenarnya dapat dimanfaatkan untuk mendukung pengambilan keputusan manajerial berbasis data. Data

sekunder berupa dataset klaim BPJS, yang mencakup variabel kelas rawat, length of stay (LOS), total tarif RS, total klaim, laba, diagnosis, kode tindakan, INA-CBGs, serta label klaim layak atau pending, merupakan sumber informasi yang sangat kaya.

Pada periode Januari–Maret 2025, kinerja keuangan rumah sakit menunjukkan tren positif dengan peningkatan konsisten pada arus kas bersih: Rp34,3 juta di Januari, Rp41,4 juta di Februari, dan Rp52,4 juta di Maret. Pertumbuhan ini didorong oleh peningkatan pendapatan BPJS yang mendominasi lebih dari 75% total pendapatan, sementara kontribusi non-BPJS relatif kecil dan fluktuatif. Walaupun beban operasional ikut meningkat seiring volume layanan, rumah sakit masih mampu menjaga surplus kas, menandakan kondisi finansial yang sehat. Namun demikian, terdapat risiko besar dari ketergantungan pada BPJS. Komposisi pendapatan yang mencapai 88% pada Februari memperlihatkan kerentanan terhadap keterlambatan pencairan klaim. Hal ini terbukti dengan adanya pending klaim, yakni Rp18,4 juta pada Januari, turun menjadi Rp3,9 juta di Februari, tetapi melonjak kembali ke Rp20,4 juta pada Maret. Pending klaim ini membuat arus kas riil lebih rendah dari potensi pendapatan yang sebenarnya. Kondisi ini menegaskan bahwa meskipun arus kas menunjukkan tren membaik, risiko laten tetap ada akibat ketidakpastian pencairan klaim. Oleh karena itu, manajemen rumah sakit perlu melakukan diversifikasi pendapatan non-BPJS sekaligus memperkuat strategi efisiensi biaya operasional. Lebih jauh, penerapan metode prediksi berbasis algoritma seperti Random Forest menjadi relevan untuk mendeteksi klaim yang berpotensi pending. Dengan langkah ini, rumah sakit dapat meminimalisasi risiko keterlambatan klaim, menjaga stabilitas likuiditas, serta memastikan keberlanjutan keuangan dalam jangka panjang.

Dengan analisis yang tepat, pola-pola tertentu dapat ditemukan untuk memprediksi klaim yang berpotensi pending. Hal ini sejalan dengan perkembangan teknologi analisis data dalam bidang kesehatan yang semakin berkembang pesat melalui pendekatan data mining dan machine learning. Salah satu metode yang terbukti efektif dalam menganalisis data kompleks dan beragam adalah Random Forest. Metode ini bekerja dengan membangun sejumlah pohon keputusan (decision tree) dan menggabungkannya untuk menghasilkan prediksi yang lebih akurat dan stabil. Keunggulan Random Forest terletak pada kemampuannya dalam mengatasi overfitting, menangani data dengan jumlah variabel besar, serta memberikan hasil berupa variabel-variabel yang paling berpengaruh terhadap keputusan klasifikasi. Dalam konteks penelitian ini, Random Forest sangat relevan untuk digunakan karena dataset klaim BPJS memiliki struktur variabel yang kompleks dan saling berhubungan. Dengan demikian, penelitian ini diarahkan untuk menganalisis dan memprediksi kelayakan klaim BPJS di Rumah Sakit Kumala Siwi menggunakan metode Random Forest. Hasil penelitian diharapkan dapat memberikan kontribusi praktis bagi manajemen rumah sakit dalam meminimalisir klaim pending, mempercepat proses pengajuan klaim, dan menjaga kestabilan arus kas. Selain itu, penelitian ini juga berkontribusi pada aspek akademik dengan memperkuat literatur mengenai penerapan machine learning dalam manajemen keuangan rumah sakit, sekaligus memberikan bukti empiris mengenai efektivitas Random Forest dalam konteks prediksi klaim kesehatan di Indonesia.

METODE PENELITIAN

Penelitian ini menggunakan data sekunder sebagai sumber utama, yaitu data klaim BPJS Kesehatan pada layanan rawat inap yang diperoleh dari Rumah Sakit Kumala Siwi dengan total sebanyak 544 data klaim. Pemanfaatan data sekunder dipilih karena efisien, relevan dengan tujuan penelitian, serta mampu menggambarkan kondisi nyata di lapangan.

Dataset ini memiliki karakteristik multidimensional yang mencakup aspek medis (diagnosis, kode tindakan), administratif (kelas rawat, LOS, INA-CBG), dan finansial (tarif rumah sakit, total klaim, laba). Variabel-variabel tersebut memberikan landasan yang kuat bagi penerapan algoritma Random Forest untuk memprediksi status klaim, baik layak maupun pending. Melalui analisis ini, penelitian diharapkan dapat membantu rumah sakit dalam mendeteksi klaim bermasalah sejak dini, menyusun pola pengajuan klaim yang lebih efektif, serta menjaga kestabilan arus kas agar pelayanan pasien tidak terganggu. Seluruh proses pengumpulan dan pengolahan data dilakukan dengan memperhatikan keamanan dan etika penelitian, termasuk anonimisasi identitas pasien serta persetujuan resmi dari pihak rumah sakit. Dengan demikian, hasil penelitian tidak hanya memiliki nilai akademik, tetapi juga praktis bagi pengelolaan klaim BPJS dan manajemen keuangan rumah sakit.

Kerangka penelitian ini disusun untuk memahami faktor-faktor yang memengaruhi kelayakan klaim BPJS di rumah sakit dengan memanfaatkan algoritma Random Forest sebagai metode klasifikasi. Dalam konteks kerja sama rumah sakit dengan BPJS Kesehatan, sistem pembayaran INA-CBGs membuat klaim tidak selalu sesuai dengan biaya aktual sehingga berpotensi menimbulkan klaim pending akibat kesalahan administrasi, kode diagnosis, maupun ketidaksesuaian data. Kondisi tersebut berdampak pada arus kas rumah sakit yang harus menanggung biaya operasional tanpa kepastian pembayaran klaim. Oleh karena itu, penelitian ini merancang model prediksi dengan menempatkan variabel kelas rawat, lama rawat inap (LOS), total tarif rumah sakit, total klaim, laba, diagnosis, kode tindakan, dan INA-CBGs sebagai prediktor, sedangkan label klaim (layak atau pending) sebagai variabel target. Data yang digunakan berupa data sekunder dari 544 klaim rawat inap di RS Kumala Siwi periode Januari–Maret 2025, yang diperoleh dari bagian rekam medis dan keuangan rumah sakit.

Metodologi penelitian dilakukan secara kuantitatif dengan pendekatan data mining. Tahapan penelitian meliputi identifikasi masalah, perumusan tujuan, studi literatur, pengumpulan dan praproses data, pembangunan model klasifikasi menggunakan Random Forest, evaluasi kinerja model, hingga analisis hasil. Pada tahap praproses, data diperiksa dari duplikasi, missing value, dan inkonsistensi sebelum dikonversi dan dibagi menjadi data latih dan data uji. Selanjutnya, metode Random Forest diterapkan dengan membentuk ratusan pohon keputusan berbasis subset data (bagging) dan pemilihan atribut acak untuk meningkatkan akurasi serta mencegah overfitting. Prediksi akhir dilakukan melalui voting mayoritas antar-pohon sehingga diperoleh klasifikasi status klaim layak atau pending.

Evaluasi model dilakukan menggunakan Confusion Matrix dan metrik evaluasi seperti akurasi, presisi, recall, F1-score, serta Area Under Curve (AUC) untuk menilai kinerja prediksi. Hasil evaluasi tidak hanya menunjukkan akurasi model, tetapi juga mengidentifikasi faktor-faktor dominan (feature importance) yang memengaruhi klaim pending. Dengan demikian, penelitian ini tidak hanya menghasilkan model prediksi yang akurat, tetapi juga memberikan informasi manajerial yang bermanfaat bagi rumah sakit untuk mengantisipasi risiko klaim pending, memperbaiki proses administratif, dan menjaga stabilitas keuangan melalui pengelolaan klaim yang lebih efektif.

HASIL DAN PEMBAHASAN

Implementasi algoritma Random Forest pada penelitian ini dilakukan melalui tiga tahap utama. Pertama, persiapan data yang mencakup pembersihan dataset klaim BPJS sebanyak 544 data rawat inap dari duplikasi, data kosong, dan kesalahan input, kemudian dilakukan transformasi variabel ke bentuk numerik serta pelabelan klaim menjadi dua kategori: layak (1) dan pending (0). Kedua, pembagian data ke dalam data latih (70%) dan

data uji (30%) agar model dapat mempelajari pola dari sebagian data dan diuji secara objektif pada data baru, sehingga menghindari overfitting. Ketiga, pembangunan model Random Forest yang memanfaatkan 30 pohon keputusan (decision tree) untuk menghasilkan prediksi lebih stabil melalui voting mayoritas. Hasil implementasi menunjukkan bahwa Random Forest mampu mengidentifikasi klaim layak dan pending dengan akurasi tinggi, berkat keragaman pohon dan mekanisme ensemble yang menyeimbangkan antara kompleksitas model dan kecepatan prediksi. Pola-pola aturan klaim yang terbentuk dari analisis 30 pohon keputusan dalam Random Forest dapat digambarkan secara naratif sebagai berikut :

1. Aturan Berdasarkan Lama Rawat (LOS)

- a. Jika **LOS ≤ 5 hari** → klaim cenderung **Layak**.
- b. Jika **LOS > 5 hari** → perlu dilihat faktor lain seperti tarif dan diagnosis.

2. Aturan Berdasarkan Total Tarif RS

- a. Jika **Tarif < Rp10.000.000** → klaim lebih sering **Layak**, karena sesuai dengan paket INA-CBGs standar.
- b. Jika **Tarif ≥ Rp10.000.000** → klaim berpotensi **Pending**, terutama bila ada perbedaan tarif antara RS dan standar INA-CBGs.

3. Aturan Berdasarkan Diagnosis dan Kode Tindakan (INA-CBGs)

- a. **Diagnosis umum dan standar** (misalnya penyakit dengan perawatan rutin) → klaim biasanya **Layak**.
- b. **Diagnosis kompleks/penyakit kronis** dengan banyak tindakan → klaim lebih rentan **Pending**, apalagi jika coding tidak sesuai standar BPJS.

4. Aturan Berdasarkan Laba RS dari Klaim

- a. Jika **Laba positif (RS masih untung)** → klaim umumnya **Layak**, karena biaya sesuai paket klaim.
- b. Jika **Laba negatif (RS merugi dari klaim)** → klaim lebih sering **Pending**, karena BPJS cenderung melakukan verifikasi ulang.

Penelitian ini menggunakan algoritma Random Forest dengan 30 tree untuk memprediksi klaim BPJS di rumah sakit, terbukti menghasilkan kinerja yang sangat baik dengan akurasi 96,88%.

accuracy: 96.88%

	true Layak	true Pending	class precision
pred. Layak	434	17	96.23%
pred. Pending	0	93	100.00%
class recall	100.00%	84.55%	

Gambar 2. Hasil Confusion Matrix

Mayoritas klaim berhasil diklasifikasikan dengan tepat ke dalam kategori klaim layak maupun pending, dengan hanya sedikit kesalahan klasifikasi. Jumlah 30 tree dipilih karena mampu memberikan keseimbangan antara akurasi dan efisiensi komputasi.

Hasil analisis menunjukkan bahwa variabel seperti kelas rawat, lama hari rawat (LOS), total tarif rumah sakit, total klaim, diagnosis, kode tindakan, serta INA-CBGs berperan penting dalam menentukan kelayakan klaim. Pada kategori klaim layak, model menunjukkan performa optimal dengan recall 100% dan precision 96,23%, sehingga hampir seluruh klaim

layak teridentifikasi dengan benar. Sementara itu, pada klaim pending, model memiliki precision 100% dengan recall 84,55%, meskipun masih ada sebagian klaim pending yang salah terklasifikasi sebagai layak.

Tabel 1. Perbandingan Laba RS Sebelum dan Sesudah Implementasi Random Forest

Periode	Lab a Sebelum Implementasi Random Forest	Lab a Setelah Implementasi Random Forest
Jan-25	Rp 77.067.470	Rp 95.546.797
Feb-25	Rp 104.757.132	Rp 108.691.710
Mar-25	Rp 132.264.128	Rp 152.690.016

Pada metode Random Forest yang diimplementasikan dalam sistem pengelolaan klaim BPJS di rumah sakit, maka akan dihasilkan data perbandingan kinerja keuangan yang menggambarkan perbedaan laba sebelum dan sesudah penerapan metode tersebut. Data ini tidak hanya berfungsi sebagai ilustrasi numerik, tetapi juga sebagai representasi empiris mengenai sejauh mana efektivitas algoritma Random Forest dalam meningkatkan akurasi prediksi klaim yang pada akhirnya berdampak langsung terhadap stabilitas arus kas rumah sakit. Dengan kata lain, penerapan metode ini memberikan gambaran strategis mengenai bagaimana rumah sakit dapat mengoptimalkan proses manajemen klaim, meminimalisasi kesalahan prediksi, serta mengantisipasi potensi klaim pending atau dispute yang selama ini sering menimbulkan hambatan finansial.

Tabel 1 menyajikan data perbandingan laba pada periode Januari hingga Maret 2025, baik sebelum maupun sesudah implementasi Random Forest. Dari data ini terlihat bahwa terdapat peningkatan laba yang cukup signifikan, sehingga dapat ditarik kesimpulan awal bahwa penerapan Random Forest berpotensi menjadi solusi yang efektif dalam mendukung pengelolaan klaim BPJS secara lebih akurat dan efisien, serta memberikan kontribusi positif terhadap keberlanjutan finansial rumah sakit.

Implementasi algoritma Random Forest terbukti memberikan dampak positif terhadap kinerja keuangan rumah sakit, khususnya dalam hal pendapatan BPJS dan arus kas bersih. Pada Januari 2025, pendapatan BPJS meningkat dari Rp77,07 juta menjadi Rp95,55 juta atau naik sebesar 24 persen, dengan arus kas bersih yang juga melonjak dari Rp34,34 juta menjadi Rp52,81 juta atau meningkat hingga 53,8 persen. Pada Februari 2025, meskipun kenaikannya lebih moderat, pendapatan BPJS tetap bertumbuh dari Rp104,76 juta menjadi Rp108,69 juta atau naik 3,8 persen, diikuti dengan peningkatan arus kas bersih dari Rp41,46 juta menjadi Rp45,39 juta atau naik sebesar 9,5 persen. Sementara itu, pada Maret 2025, terjadi peningkatan pendapatan BPJS dari Rp132,26 juta menjadi Rp152,69 juta atau naik 15,5 persen, serta kenaikan arus kas bersih dari Rp52,44 juta menjadi Rp72,86 juta atau meningkat sebesar 39 persen. Secara keseluruhan, penerapan Random Forest berkontribusi pada rata-rata peningkatan arus kas bersih sekitar 34,1 persen dalam tiga bulan. Hal ini menunjukkan bahwa penggunaan metode pembelajaran mesin mampu mempercepat verifikasi klaim, mengurangi klaim pending, serta memperkuat stabilitas keuangan rumah sakit secara signifikan.

Implikasi dari peningkatan arus kas bersih ini sangatlah penting. Dengan kondisi likuiditas yang lebih baik, rumah sakit dapat menjalankan kegiatan operasional tanpa hambatan yang berarti. Rumah sakit mampu memenuhi kewajiban rutin seperti membayar gaji tenaga kesehatan, biaya operasional harian, serta melakukan pemeliharaan infrastruktur. Selain itu, adanya dana lebih juga memungkinkan rumah sakit melakukan investasi pada pengembangan layanan kesehatan, misalnya dengan meningkatkan kualitas alat medis, memperluas fasilitas pelayanan, atau menambah program kesehatan preventif bagi

masyarakat. Dengan demikian, dampak positif dari penggunaan algoritma Random Forest tidak hanya dirasakan secara internal oleh pihak manajemen rumah sakit, tetapi juga oleh pasien dan masyarakat luas melalui peningkatan kualitas layanan.

Tabel 1. Arus Kas Setelah Implementasi dan Sebelum Implementasi Random Forest

Periode Bulan	Thn	Pendapatan Sebelum Random Forest (Rp)	BPJS Implementasi (Rp)	Pendapatan BPJS Implementasi (Rp)	Setelah Random Forest	Pendapatan Non BPJS
Januari	2025	Rp 77.067.470		Rp 95.546.797		Rp20.240.715
Februari	2025	Rp 104.757.132		Rp 108.691.710		Rp14.501.386
Maret	2025	Rp 132.264.128		Rp 152.690.016		Rp21.329.745

Periode Bulan	Thn	Total Pendapatan Sebelum Implementasi Random Forest (Rp)	Total Pendapatan Setelah Implementasi Random Forest (Rp)	Pengeluaran Operasional (Rp)	Arus Kas Bersih Sebelum Implementasi Random Forest (Rp)	Arus Kas Bersih Setelah Implementasi Random Forest (Rp)
Januari	2025	Rp97.308.185	Rp115.787.512	Rp62.971.415	34.336.770	52.816.097
Februari	2025	Rp119.258.518	Rp123.193.096	Rp77.801.959	41.456.559	45.391.137
Maret	2025	Rp153.593.873	Rp174.019.761	Rp101.157.337	52.436.536	72.862.424

Selain itu, keberhasilan implementasi algoritma Random Forest dalam meningkatkan arus kas bersih rumah sakit juga dapat dipandang sebagai bukti nyata bahwa penerapan teknologi berbasis data mining mampu memberikan nilai tambah dalam sektor kesehatan. Dalam konteks penelitian ini, algoritma Random Forest terbukti efektif dalam mengatasi permasalahan keterlambatan klaim BPJS yang selama ini sering menjadi kendala bagi rumah sakit. Dengan klaim yang lebih cepat dan tepat, potensi kerugian akibat keterlambatan pembayaran dapat diminimalkan. Hal ini pada akhirnya meningkatkan kepastian arus kas yang diterima, sekaligus memperkuat ketahanan finansial rumah sakit.

Temuan ini memiliki implikasi strategis bagi pengembangan sistem keuangan rumah sakit di masa depan. Penerapan teknologi prediktif seperti Random Forest dapat dijadikan model dalam pengelolaan data klaim dan arus kas di rumah sakit lain yang menghadapi permasalahan serupa. Dengan adanya bukti empiris berupa peningkatan arus kas bersih yang signifikan, pihak manajemen rumah sakit, regulator, maupun pemerintah dapat mempertimbangkan pemanfaatan algoritma data mining sebagai strategi pengelolaan keuangan yang lebih efektif.

PENUTUP

Kesimpulan

Penelitian ini membuktikan bahwa penerapan algoritma Random Forest dalam klasifikasi klaim BPJS Kesehatan mampu memberikan manfaat strategis yang signifikan bagi manajemen rumah sakit. Dengan menggunakan 30 pohon keputusan, model berhasil mencapai akurasi tinggi sebesar 96,88%, precision yang baik pada kedua kelas klaim, serta recall optimal terutama pada klaim layak (100%). Meskipun data penelitian tidak seimbang, dengan 520 klaim layak dan 24 klaim pending, model tetap dapat mengenali pola klaim pending dengan recall 84,55%. Hal ini menunjukkan bahwa Random Forest cukup efektif

dalam menghadapi tantangan data yang tidak seimbang, sekaligus menegaskan perannya sebagai sistem prediksi yang andal dalam mendukung keputusan manajerial. Analisis feature importance mengungkapkan bahwa variabel paling dominan dalam klasifikasi klaim adalah total tarif rumah sakit, lama rawat inap (LOS), diagnosis, kode tindakan INA-CBGs, dan laba rumah sakit. Variabel-variabel ini terbukti berkaitan langsung dengan kelayakan klaim, di mana klaim dengan LOS pendek, biaya moderat, dan diagnosis standar cenderung disetujui, sementara klaim dengan LOS panjang, biaya tinggi, serta diagnosis dan tindakan kompleks lebih sering berakhir pending.

Dari perspektif manajemen, hasil penelitian ini menunjukkan bahwa Random Forest tidak hanya berfungsi sebagai alat analisis teknis, tetapi juga sebagai instrumen penting dalam penerapan manajemen berbasis data (data-driven management). Model prediktif ini dapat digunakan dalam perencanaan (planning) untuk memproyeksikan klaim pending dan dampaknya terhadap arus kas, dalam pengorganisasian (organizing) untuk mendukung alokasi sumber daya manusia pada proses verifikasi, dalam penggerakan (actuating) untuk mempercepat pengambilan keputusan terkait perbaikan dokumen atau coding, serta dalam pengendalian (controlling) untuk memantau efektivitas pengelolaan klaim secara berkelanjutan. Dengan demikian, penerapan algoritma ini selaras dengan prinsip POAC dan relevan dengan tuntutan digitalisasi layanan kesehatan.

Berdasarkan temuan penelitian, terdapat beberapa rekomendasi yang dapat diberikan. Pertama, rumah sakit disarankan untuk mengintegrasikan model prediksi Random Forest ke dalam SIMRS maupun sistem e-Klaim BPJS, agar verifikasi klaim dapat dilakukan otomatis dan real-time. Kedua, peningkatan kompetensi SDM menjadi langkah penting, melalui pelatihan bagi petugas admin, verifikator, dan tenaga medis untuk meminimalkan kesalahan coding dan administrasi yang sering menjadi penyebab klaim pending. Ketiga, hasil prediksi dapat digunakan untuk mendukung perencanaan keuangan, khususnya dalam mengantisipasi potensi keterlambatan pembayaran klaim, risiko defisit arus kas, serta strategi mitigasi keuangan rumah sakit. Keempat, penelitian ini juga memperkuat konsep manajemen berbasis analitik (analytics-based management), di mana pengambilan keputusan dapat dilakukan berbasis bukti (evidence-based decision making) dengan dukungan teknologi prediksi. Kelima, untuk penelitian selanjutnya disarankan memperluas dataset dengan mencakup klaim dari berbagai rumah sakit dan membandingkannya dengan algoritma lain seperti XGBoost, Gradient Boosting, maupun Neural Network, sehingga diperoleh gambaran yang lebih komprehensif terkait efektivitas berbagai teknik analitik. Dengan demikian, penelitian ini tidak hanya memberikan manfaat praktis bagi peningkatan efisiensi pengelolaan klaim BPJS, tetapi juga berkontribusi pada pengembangan teori manajemen modern di era digitalisasi kesehatan.

DAFTAR PUSTAKA

- Anitha, R., & Siva Sundhara Raja, D. (2018). Development of computer-aided approach for brain tumor detection using random forest classifier. *International Journal of Imaging Systems and Technology*, 28(1), 48–53. <https://doi.org/10.1002/ima.22255>
- Antonius Artanto EP. (2018). Faktor-Faktor Penyebab Klaim Tertunda BPJS Kesehatan RSUD Dr. Kanujoso Djatiwibowo Periode Januari-Maret 2016 Antonius Artanto EP. *Jurnal Administrasi Rumah Sakit*, 4.

- Ariyanti, F., & Gifari, M. T. (2019). Analisis Persetujuan Klaim BPJS Kesehatan pada Pasien Rawat Inap. *Jurnal Ilmu Kesehatan Masyarakat*, 8(04), 156–166. <https://doi.org/10.33221/jikm.v8i04.415>
- Frempong, N. K., Nicholas, N., & Boateng, M. A. (2017). Decision Tree as a Predictive Modeling Tool for Auto Insurance Claims. *International Journal of Statistics and Applications*, 7(2), 117–120. <https://doi.org/10.5923/j.statistics.20170702.07>
- Lan, T., Hu, H., Jiang, C., Yang, G., & Zhao, Z. (2020). A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. *Advances in Space Research*, 65(8), 2052–2061. <https://doi.org/10.1016/j.asr.2020.01.036>
- Li, Y., Yan, C., Liu, W., & Li, M. (2018). A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification. *Applied Soft Computing Journal*, 70, 1000–1009. <https://doi.org/10.1016/j.asoc.2017.07.027>
- Lin, W., Wu, Z., Lin, L., Wen, A., & Li, J. (2017). An ensemble random forest algorithm for insurance big data analysis. *IEEE Access*, 5, 16568–16575. <https://doi.org/10.1109/ACCESS.2017.2738069>
- Moon, M., & Lee, S. K. (2017). Applying of decision tree analysis to risk factors associated with pressure ulcers in long-term care facilities. *Healthcare Informatics Research*, 23(1), 43–52. <https://doi.org/10.4258/hir.2017.23.1.43>
- Mustika Kurniawati, Tuty Satrijawati, & Tuti Hariyanto. (2020). Faktor Yang Mempengaruhi Pengembalian Administrasi Klaim Bpjs Kesehatan Ditinjau Dari Syarat-Syarat Kelengkapan Administrasi Klaim. *Chmk Health Journal*, Volume 4 Nomor 3.
- Nuraini, N., Wijayanti, R. A., Putri, F., Alfiansyah, G., Deharja, A., & Santi, M. W. (2019). Optimalisasi Manajemen Penanganan Klaim Pending Pasien Bpjs Rawat Inap Di Rumah Sakit Citra Husada Jember. *Kesmas Indonesia*, 11, 24–35. <http://jos.unsoed.ac.id/index.php/kesmasindo/article/view/1314/1031>
- Nurzaman, F. (2017). Penerapan Algoritma Regresi Linier Untuk Prediksi Jumlah Klaim Pada Asuransi Kesehatan. *Seminar Nasional Teknologi Informasi*, 105–114.
- Nyoman, I., Semarajana, G., & Soewondo, P. (2019). *Icash-A102 Factors Related To Pending Claim In Indonesian National Health Insurance (Jkn): A Systematic Review* (Nomor 4).
- Pal, R., & Pal, S. (t.t.). Application of Data Mining Techniques in Health Fraud Detection. *International Journal of Engineering Research and General Science*, 3(5). www.ijergs.org
- Pal, R., & Pal, S. (2015). Application of Data Mining Techniques in Health Fraud Detection. *International Journal of Engineering Research and General Science*, 3(5), 129–137. www.ijergs.org

- Robson, B., & Boray, S. (2019). Studies in the use of data mining, prediction algorithms, and a universal exchange and inference language in the analysis of socioeconomic health data. *Computers in Biology and Medicine*, 112. <https://doi.org/10.1016/j.combiomed.2019.103369>
- Sari, M., Gulbandila, E., & Dalkilic, N. (2018). Risk Assessment with Decision Tree in Professional Liability Insurance: In Accounting. *Journal of Artificial Intelligence*, 12(1), 18–23. <https://doi.org/10.3923/jai.2019.18.23>
- Vekeman, F., Piña-Garza, J. E., Cheng, W. Y., Tuttle, E., Giguère-Duval, P., Oganisian, A., Damron, J., Sheng Duh, M., Shen, V., Saurer, T. B., Montouris, G. D., & Isojarvi, J. (2019). Development of a classifier to identify patients with probable Lennox–Gastaut syndrome in health insurance claims databases via random forest methodology. *Current Medical Research and Opinion*, 35(8), 1415–1420. <https://doi.org/10.1080/03007995.2019.1595552>
- Wang, C., Wang, H., Zhou, C., & Chen, H. (2020). ExperienceThinking: Constrained hyperparameter optimization based on knowledge and pruning. *Knowledge-Based Systems*, 106602. <https://doi.org/10.1016/j.knosys.2020.106602>
- Zhan, C., Zheng, Y., Zhang, H., & Wen, Q. (2021). Random-Forest-Bagging Broad Learning System with Applications for COVID-19 Pandemic. *IEEE Internet of Things Journal*, 8(21), 15906–15918. <https://doi.org/10.1109/JIOT.2021.3066575>