

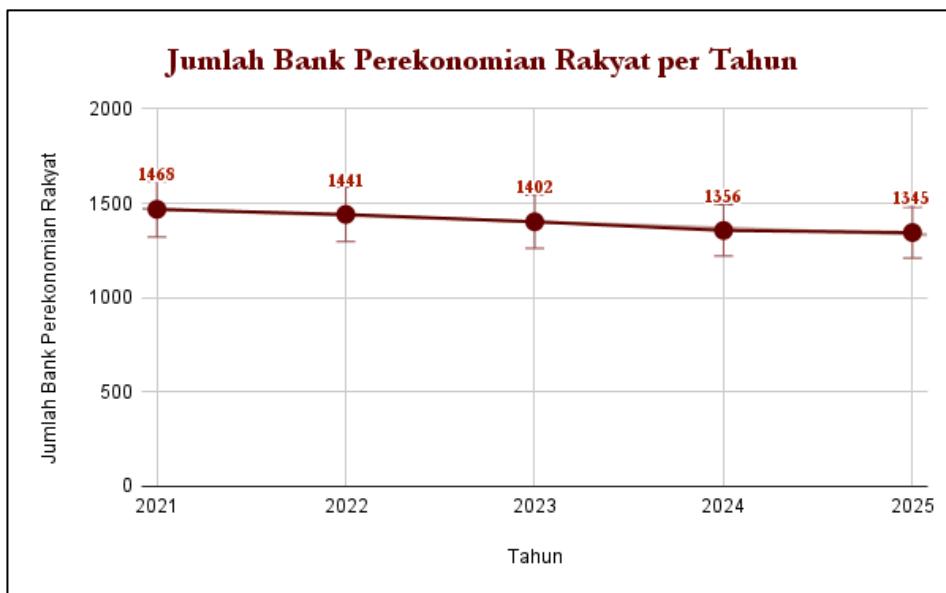
# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Bank Perekonomian Rakyat (BPR), sebelumnya dikenal sebagai Bank Perkreditan Rakyat, merupakan lembaga keuangan yang beroperasi secara terbatas dengan fokus utama melayani masyarakat lokal, khususnya dalam bentuk penghimpunan dana dan penyaluran kredit skala mikro. Perubahan nama ini sesuai dengan Undang-Undang No. 4 Tahun 2023 tentang Pengembangan dan Penguatan Sektor Keuangan (UU P2SK), yang mengamanatkan perubahan penamaan BPR untuk meningkatkan peranannya dalam perekonomian masyarakat [1]. Tujuan dari perubahan nama ini adalah untuk menjadikan BPR sebagai lembaga keuangan yang lebih kuat dan dapat meningkatkan kepercayaan masyarakat. Fungsi utama BPR mencakup penghimpunan simpanan masyarakat, pemberian kredit untuk usaha mikro, kecil, dan menengah (UMKM), serta penyediaan jasa keuangan yang lebih mudah diakses di tingkat regional. Peran BPR dalam perekonomian Indonesia cukup signifikan, karena turut mendukung inklusi keuangan, memperkuat sektor UMKM, dan menjaga stabilitas ekonomi lokal melalui akses kredit yang lebih mudah [2].

Menurut data Otoritas Jasa Keuangan (OJK), industri Bank Perekonomian Rakyat (BPR) di Indonesia menghadapi perkembangan yang signifikan dalam beberapa tahun terakhir. Dari sisi jumlah, BPR menunjukkan tren penurunan secara konsisten sebagaimana ditunjukkan pada Gambar 1.1. Pada akhir 2021 tercatat sebanyak 1.468 BPR, kemudian berkurang menjadi 1.441 BPR pada 2022. Tren penurunan ini berlanjut pada 2023 dengan jumlah 1.402 BPR, lalu menjadi 1.356 BPR pada akhir 2024, dan pada Maret 2025 tercatat hanya 1.345 BPR. Penurunan ini umumnya terjadi akibat konsolidasi maupun pencabutan izin usaha [3][4][5].



Gambar 1. 1 Jumlah Bank Perekonomian Rakyat per Tahun

Sumber: [3][4][5]

Meskipun jumlah BPR menurun, total aset industri justru menunjukkan pertumbuhan yang signifikan sebagaimana terlihat pada Gambar 1.2. Pada 2020, total aset BPR sebesar Rp155,07 triliun, kemudian meningkat menjadi Rp168,44 triliun pada 2021, Rp182,30 triliun pada 2022, Rp194,98 triliun pada 2023, dan Rp204,63 triliun pada 2024. Bahkan, hingga Maret 2025 total aset BPR masih terjaga tinggi sebesar Rp203,68 triliun. Hal ini menunjukkan bahwa meskipun jumlah BPR semakin sedikit, industri secara keseluruhan tetap tumbuh dari sisi aset dan memiliki daya tahan yang kuat [6].



Gambar 1. 2 Perkembangan Total Aset BPR

Sumber : [6]

Dalam memastikan kinerja BPR tetap sehat, Otoritas Jasa Keuangan (OJK) menetapkan ketentuan mengenai penilaian tingkat kesehatan bank melalui SEOJK.03/2022 tentang Penilaian Tingkat Kesehatan BPR dan BPRS, yang merupakan aturan pelaksanaan dari POJK Nomor 3/POJK.03/2022. Penilaian ini didasarkan pada pendekatan *Risk Profile, Good Corporate Governance, Earnings, dan Capital* (RGEC) yang menekankan evaluasi menyeluruh terhadap risiko, tata kelola, profitabilitas, dan kecukupan modal bank [7].

Dalam upaya meningkatkan efektivitas pengawasan industri perbankan, khususnya BPR, Otoritas Jasa Keuangan (OJK) melalui satuan kerja Departemen Manajemen Risiko dan Pengendalian Kualitas (DRPK) di unit Departemen Manajemen Risiko (DMRS) pada bagian *Data Analytic* telah mengimplementasikan sistem *Monitoring Tools* sebagai bagian dari tata kelola risiko yang terintegrasi dari berbagai aplikasi di OJK. *Monitoring Tools* menjalankan pemantauan secara berkelanjutan, memanfaatkan data yang tersedia untuk menghasilkan informasi yang akurat, relevan, dan terkini. Sistem ini membantu tim *Data Analytics* dalam melakukan analisis data yang bersifat *compliance* dan *insight* [8].

Meskipun *Monitoring Tools* menghadirkan kerangka kerja yang canggih, implementasinya masih memiliki keterbatasan, terutama dalam klasifikasi tingkat

kesehatan BPR. Data keuangan memang tersedia dalam sistem, tetapi analis tetap harus melakukan query manual melalui SQLPad untuk memperoleh hasil evaluasi. Proses manual ini memakan waktu, rawan inkonsistensi, dan dapat menunda pengambilan keputusan. Keterbatasan ini semakin krusial mengingat jumlah BPR yang diawasi terus bertambah, sehingga sistem membutuhkan mekanisme yang lebih efisien, otomatis, dan adaptif. Dari perspektif *Data Analytics*, kondisi tersebut mencerminkan celah dalam integrasi data dan otomasi analisis, yang dapat mengurangi efektivitas pengawasan dan kecepatan merespons dari pihak regulator. Oleh karena itu, pendekatan berbasis *machine learning* berpotensi menjadi solusi untuk mengotomatisasi klasifikasi tingkat kesehatan bank. *Machine Learning* (ML) secara operasional, memanfaatkan data dan algoritma untuk mengenali pola, menghasilkan prediksi, dan mengambil keputusan secara mandiri tanpa perlu diprogram secara eksplisit, serupa dengan proses pembelajaran manusia. [9][10].

Untuk mewujudkan solusi tersebut, penelitian ini menggunakan data *dummy* yang disediakan oleh tim *Data Analytics* OJK, terdiri atas rasio-rasio keuangan yang mencerminkan kerangka *Risk Profile*, *Good Corporate Governance*, *Earnings*, dan *Capital* (RGEC). Data tersebut kemudian dilakukan proses pembersihan, validasi, dan pengolahan untuk memastikan kualitas dan konsistensi sebelum diterapkan pada metode *clustering*. Dengan pendekatan ini, penelitian bertujuan untuk mengotomatisasi klasifikasi tingkat kesehatan BPR, sehingga tim pengawas di unit DMRS dapat mengidentifikasi kondisi bank secara lebih akurat, dan berbasis data, mendukung pengambilan keputusan yang lebih tepat. Setelah data siap, tahap berikutnya adalah penerapan metode *clustering*. Hasil *clustering* ini kemudian dimanfaatkan sebagai input dalam tahap klasifikasi.

Penelitian ini menggunakan bahasa pemrograman Python dengan membandingkan kinerja dua algoritma *clustering*, yaitu *K-Means* dan *K-Medoids*, untuk menghasilkan segmentasi tingkat kesehatan BPR berdasarkan rasio RGEC. *K-Means* dipilih karena mampu melakukan segmentasi dengan kompleksitas komputasi [11][12], sehingga efisien dalam menangani data berukuran besar seperti laporan keuangan BPR. Namun, kelemahan utama *K-Means* adalah sensitivitas terhadap

*outlier* yang dapat memengaruhi posisi centroid dan menurunkan kualitas *cluster* [13]. Oleh karena itu, penelitian ini juga menguji performa *K-Medoids*, yang menggantikan *centroid* dengan *medoid* (salah satu objek dalam *cluster*), sehingga lebih *robust* terhadap *outlier* [14]. Dengan membandingkan kedua algoritma, penelitian ini bertujuan mengevaluasi antara efisiensi komputasi (*K-Means*) dan stabilitas hasil *clustering* (*K-Medoids*). Evaluasi performa *clustering* akan dilakukan dengan menggunakan metrik *Silhouette Score* dan *Davies-Bouldin Index* (DBI) untuk mengukur kualitas pemisahan *cluster* [15].

Hasil *clustering* kemudian digunakan sebagai input dalam tahap klasifikasi menggunakan algoritma *Random Forest*. *Random Forest* dipilih karena kemampuannya menangani data *multivariat*, ketahanannya terhadap *overfitting* dan memberikan hasil prediksi yang konsisten dengan tingkat akurasi tinggi [16][17]. Pada penelitian terdahulu [18] melakukan pengukuran menggunakan algoritma *random forest* menghasilkan akurasi sebesar 96,77%, Selanjutnya terdapat penelitian lain yang membuktikan bahwa *random forest* memiliki kinerja yang baik mencapai akurasi 98.4% dalam memprediksi kegagalan bank [19]. Hal ini dapat diartikan bahwa *Random Forest* terbukti berpotensi unggul dalam mengklasifikasikan tingkat kesehatan bank. Pada penelitian ini, kinerja *Random Forest* akan dievaluasi menggunakan akurasi, *precision*, *recall*, *F1-score*, serta *AUC-ROC* [20].

Berbagai penelitian sebelumnya membuktikan bahwa metode ML efektif dalam menganalisis risiko dan kesehatan bank [18][19], [22]–[29]. Namun, sebagian besar penelitian memiliki sejumlah keterbatasan. Pertama, sebagian besar penelitian sebelumnya hanya menggunakan rasio keuangan umum seperti ROA, Leverage, dan profitabilitas [21], tanpa mengacu pada standar RGEC sebagai kerangka resmi penilaian kesehatan bank oleh OJK [7]. Kedua, perbandingan langsung antara *K-Means* dan *K-Medoids* berbasis rasio RGEC masih sangat terbatas, sedangkan data keuangan BPR rentan terhadap *outlier*. Ketiga, mayoritas studi dilakukan pada perbankan luar negeri, sementara penelitian yang secara khusus membahas BPR di Indonesia masih jarang ditemukan. Keempat, penelitian terdahulu umumnya berhenti pada tahap *clustering* tanpa melanjutkannya ke proses klasifikasi prediktif [22]–[29].

Dengan mempertimbangkan berbagai gap tersebut, penelitian ini bertujuan untuk memberikan pemahaman yang lebih mendalam mengenai efektivitas algoritma *K-Means* dan *K-Medoids* dalam segmentasi tingkat kesehatan BPR berbasis RGEC, sekaligus mengembangkan model klasifikasi prediktif yang relevan dengan kondisi perbankan di Indonesia.

## 1.2 Rumusan Masalah

Dari latar belakang di atas dapat diambil beberapa rumusan masalah seperti berikut:

- 1) Algoritma manakah antara *K-Means* dan *K-Medoids* yang memiliki kinerja terbaik dalam proses clustering tingkat kesehatan BPR, ditinjau dari nilai *Silhouette Score* dan *Davies–Bouldin Index* (DBI)?
- 2) Bagaimana hasil clustering terbaik dapat dimanfaatkan untuk membangun model klasifikasi menggunakan algoritma *Random Forest* untuk memprediksi kategori tingkat kesehatan BPR berdasarkan Rasio RGEC?
- 3) Bagaimana model prediksi diimplementasikan dalam sebuah sistem yang dikembangkan secara *web-based*?

## 1.3 Batasan Masalah

Dalam penelitian ini, batasan masalah ditetapkan untuk menjaga fokus dan ruang lingkup penelitian agar tetap jelas. Adapun batasan masalah yang diterapkan adalah sebagai berikut:

- 1) Penelitian ini dibatasi pada penggunaan data keuangan Bank Perekonomian Rakyat (BPR) berupa data *dummy* yang bersumber dari sistem *Monitoring Tools* Otoritas Jasa Keuangan (OJK), dengan periode pengamatan mulai Juni 2024 hingga Agustus 2025.
- 2) Rasio yang dianalisis terbatas pada komponen *Risk Profile*, *Good Corporate Governance*, *Earnings*, dan *Capital* (RGEC).
- 3) Penelitian ini menggunakan dua algoritma *clustering*, yaitu *K-Means* dan *K-Medoids*, untuk melakukan segmentasi tingkat kesehatan BPR. Hasil

*clustering* digunakan sebagai input untuk tahap *classification* dengan algoritma *Random Forest*.

- 4) Penelitian ini membatasi implementasi model prediksi ke dalam sistem yang dikembangkan secara *web-based*, tanpa membahas integrasi ke aplikasi lain di luar platform web.

## 1.4 Tujuan dan Manfaat Penelitian

Berdasarkan tujuan yang ingin dicapai dalam penelitian ini, selanjutnya dapat ditentukan manfaat yang diharapkan, baik dari segi akademik maupun praktis, untuk memberikan kontribusi yang lebih nyata bagi pengembangan ilmu dan praktik pengawasan BPR.

### 1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, tujuan dari penelitian ini adalah sebagai berikut:

- 1) Membandingkan kinerja algoritma *K-Means* dan *K-Medoids* dalam mengelompokkan tingkat kesehatan Bank Perekonomian Rakyat (BPR) berdasarkan rasio keuangan *Risk Profile*, *Good Corporate Governance*, *Earnings*, dan *Capital* (RGEC).
- 2) Menentukan hasil klasterisasi terbaik antara kedua algoritma tersebut dan memanfaatkan label hasil klasterisasi sebagai dasar untuk membangun model klasifikasi menggunakan algoritma Random Forest dalam memprediksi kategori tingkat kesehatan BPR.
- 3) Mengimplementasikan sistem prediksi tingkat kesehatan BPR berbasis web menggunakan *Streamlit*, yang mampu menerima data baru dan memberikan hasil prediksi secara otomatis.

### 1.4.2 Manfaat Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis. Secara teoritis:

- 1) Memberikan kontribusi terhadap pengembangan ilmu di bidang *data mining* dan *machine learning*, khususnya dalam penerapan dan perbandingan algoritma *clustering K-Means* dan *K-Medoids* pada analisis tingkat kesehatan perbankan.
- 2) Menjadi referensi akademik bagi penelitian selanjutnya yang berfokus pada pengembangan model prediksi kesehatan bank berbasis kombinasi *unsupervised learning* dan *supervised learning*.
- 3) Memperluas pemahaman tentang kinerja algoritma klasifikasi *Random Forest* dalam mengolah data keuangan hasil klasterisasi, sehingga dapat diterapkan pada berbagai kasus prediksi di bidang keuangan.

Selanjutnya, secara praktis:

- 1) Menyediakan model prediksi tingkat kesehatan BPR yang dapat membantu lembaga pengawas, seperti Otoritas Jasa Keuangan (OJK), dalam melakukan pemantauan secara cepat, akurat, dan berbasis data.
- 2) Mendukung otomatisasi proses evaluasi kesehatan bank, sehingga mengurangi risiko kesalahan akibat proses manual.
- 3) Memberikan dasar ilmiah dan teknis bagi pengembangan sistem penilaian kesehatan BPR berbasis teknologi, yang dapat diintegrasikan ke dalam sistem pengawasan digital di masa depan.

## 1.5 Sistematika Penulisan

Sistematika penulisan ini terdiri dari 5 bagian yaitu :

### 1) BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang permasalahan yang menjadi dasar penelitian, rumusan masalah yang terdiri dari pertanyaan inti penelitian, batasan ruang lingkup agar penelitian tetap fokus, serta tujuan dan manfaat penelitian baik secara teoretis maupun praktis. Bab ini ditutup dengan uraian sistematika penulisan sebagai petunjuk pembaca terhadap struktur pembahasan keseluruhan skripsi.

## 2) BAB II LANDASAN TEORI

Bab ini membahas teori-teori yang menjadi dasar penelitian, meliputi konsep Bank Perekonomian Rakyat (BPR), penilaian kesehatan bank berbasis RGEC, konsep dasar machine learning, algoritma *K-Means*, *K-Medoids*, *Random Forest*, framework CRISP-DM, penelitian terdahulu yang relevan, serta perangkat lunak pendukung. Teori-teori ini memberikan landasan ilmiah bagi proses analisis.

## 3) BAB III METODOLOGI PENELITIAN

Bab ini menguraikan pendekatan metodologis yang digunakan dalam penelitian. Pembahasan meliputi gambaran umum objek penelitian, metode penelitian, perbandingan framework, alur penelitian berdasarkan CRISP-DM, teknik pengumpulan data, identifikasi variabel, serta teknik analisis data yang digunakan untuk proses *clustering* dan klasifikasi. Bab ini memberikan penjelasan sistematis mengenai tahapan pengerjaan penelitian.

## 4) BAB IV ANALISIS DAN HASIL PENELITIAN

Bab ini menyajikan hasil implementasi dan analisis data sesuai tahapan CRISP-DM, dimulai dari business understanding hingga deployment. Pembahasan mencakup analisis struktur data, pembersihan dan transformasi rasio ke dalam bentuk penilaian RGEC, perbandingan algoritma *K-Means* dan *K-Medoids*, penilaian performa menggunakan *Silhouette Score* dan *Davies–Bouldin Index*, pembangunan model klasifikasi *Random Forest* berdasarkan hasil *cluster* terbaik, serta implementasi model prediksi berbasis web. Bab ini menjadi inti temuan penelitian.

## 5) BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan dari hasil penelitian serta saran bagi pengembangan penelitian maupun pengembangan sistem prediksi di masa mendatang.