



010101110000110
10101010100101
010101011100101
10100101010101

Analitik Prediktif dalam Keuangan

— Mengelola Risiko dengan Data Science —

Oleh: Rudy C Tarumingkeng



1 FOKUS UTAMA

- Prediksi
- Risiko
- Keputusan

2 RISIKO KEUANGAN

- Kredit
- Pasar
- Likuiditas
- Fraud

3 METODE

- Statistika
- Machine Learning
- AI
- Time Series

4 MANFAAT

- Early Warning
- Akurasi
- Efisiensi
- Mitigasi

5 PRINSIP

- Transparan
- Etis
- Andal
- Berbasis Data



*Data Science untuk keputusan keuangan
yang lebih cerdas, adaptif, dan bertanggung jawab*



Rudy C Tarumingkeng: **ANALITIK PREDIKTIF DALAM KEUANGAN -
MENGELOLA RISIKO DENGAN DATA SCIENCE**

Oleh:

[Prof Ir Rudy C Tarumingkeng, PhD](#)

Professor of Management NUP: 9903252922

Professor Emeritus, IPB-University

Anggota MPR-RI, 1982-1987

Dekan Fak. Kehutanan IPB University. 1967-1968, 1968-1971, 1977-1978

Rektor, Universitas Cenderawasih, Papua (1978-1988, dan

Rektor, Kampus AGRO Manokwari sekarang Universitas Papua Manokwari

Coordinator, CIDA/DIKTI SFU Burnaby BC Canada 1988-1991

Rektor, Universitas Kristen Krida Wacana, Jakarta (1991-2000)

Ketua, Dewan Guru Besar, IPB-University, Bogor (2005-2006)

AI - Data Analyst, dan Ketua, Senat Akademik, IBM-ASMI, Jakarta 2024-

© RudyCT Academic Series

rudyclt75@gmail.com

14 Juli 2026

ANALITIK PREDIKTIF DALAM KEUANGAN: MENGELOLA RISIKO DENGAN DATA SCIENCE

Abstrak

Industri keuangan beroperasi dalam lingkungan yang ditandai oleh ketidakpastian, perubahan pasar yang cepat, kompleksitas hubungan ekonomi, pertumbuhan transaksi digital, serta meningkatnya risiko siber dan kejahatan keuangan. Dalam lingkungan tersebut, kemampuan untuk memperkirakan kemungkinan kejadian pada masa mendatang menjadi bagian penting dari manajemen risiko. Analitik prediktif menawarkan pendekatan sistematis untuk menggunakan data historis dan data terkini guna mengestimasi probabilitas gagal bayar, kerugian kredit, volatilitas pasar, kebutuhan likuiditas, anomali transaksi, potensi fraud, tekanan keuangan perusahaan, serta berbagai risiko lain yang dihadapi lembaga keuangan.

Analitik prediktif dalam keuangan mengintegrasikan statistika, ekonometrika, machine learning, kecerdasan buatan, teknologi basis data, komputasi awan, dan pengetahuan domain keuangan. Pendekatan ini tidak menggantikan teori keuangan atau pertimbangan profesional, melainkan memperkuat kemampuan organisasi untuk mengidentifikasi pola, mengenali sinyal peringatan dini, mengukur ketidakpastian, dan

merancang tindakan mitigasi. Model prediktif dapat berupa regresi logistik, model survival, pohon keputusan, random forest, gradient boosting, support vector machine, jaringan saraf, model deret waktu, natural language processing, analisis jaringan, dan metode ensemble.

Meskipun model machine learning sering menawarkan peningkatan daya diskriminasi dan kemampuan menangkap hubungan nonlinier, penerapannya dalam keuangan menghadapi tantangan yang khas. Data keuangan bersifat dinamis, tidak seimbang, sensitif, serta dipengaruhi perubahan ekonomi, kebijakan, perilaku, dan regulasi. Model yang memiliki akurasi tinggi dalam periode pelatihan belum tentu stabil pada periode krisis. Selain itu, model yang sulit dijelaskan dapat menimbulkan persoalan tata kelola, kepatuhan, keadilan, dan kepercayaan. Oleh karena itu, evaluasi model tidak cukup hanya menggunakan akurasi atau area under the curve. Model juga harus dinilai berdasarkan kalibrasi probabilitas, kestabilan, ketahanan terhadap perubahan distribusi, biaya kesalahan, keadilan, interpretabilitas, serta dampak ekonominya.

Artikel ini membahas konsep, metodologi, arsitektur, aplikasi, evaluasi, implementasi, dan tata kelola analitik prediktif dalam manajemen risiko keuangan. Pembahasan mencakup risiko kredit, risiko pasar, risiko likuiditas, risiko operasional, fraud, anti-money laundering, tekanan keuangan perusahaan, stress testing, serta risiko model. Artikel ini juga memberikan kerangka implementasi bagi lembaga keuangan, termasuk pengembangan data pipeline, validasi independen, explainable artificial intelligence, human oversight, pemantauan model, dan integrasi prediksi dengan keputusan manajerial.

Kesimpulan utama artikel ini adalah bahwa nilai analitik prediktif tidak terletak pada kecanggihan algoritma semata. Nilai tersebut muncul ketika prediksi yang akurat, terkalibrasi, dapat dijelaskan, dan relevan diterjemahkan menjadi keputusan yang tepat waktu dan bertanggung jawab. Dengan tata kelola yang memadai, Data Science dapat membantu lembaga keuangan bergerak dari manajemen risiko yang reaktif menuju

manajemen risiko yang antisipatif, adaptif, dan berorientasi pada ketahanan jangka panjang.

Kata Kunci: *analitik prediktif, Data Science, manajemen risiko keuangan, risiko kredit, risiko pasar, risiko likuiditas, fraud detection, machine learning, explainable artificial intelligence, model risk management.*

1. Pendahuluan

Keuangan pada hakikatnya berhubungan dengan pengambilan keputusan mengenai masa depan. Bank memberikan kredit sekarang dengan harapan pembayaran akan diterima kemudian. Investor membeli aset hari ini berdasarkan perkiraan arus kas dan harga masa depan. Perusahaan asuransi menerima premi dengan kewajiban menanggung kerugian yang mungkin terjadi pada periode mendatang. Manajer keuangan menentukan kebutuhan modal kerja berdasarkan estimasi penerimaan, pembayaran, permintaan, harga, dan kondisi pasar.

Semua keputusan tersebut mengandung ketidakpastian. Tidak ada lembaga keuangan yang dapat mengetahui secara pasti apakah seorang debitur akan gagal bayar, apakah harga saham akan meningkat, apakah nasabah akan menarik simpanannya secara bersamaan, atau apakah suatu transaksi merupakan transaksi yang sah. Keputusan keuangan karena itu tidak hanya mempertimbangkan nilai yang diharapkan, tetapi juga distribusi kemungkinan hasil, besarnya kerugian, dan kemampuan organisasi menanggung dampaknya.

Secara tradisional, lembaga keuangan mengelola ketidakpastian melalui laporan keuangan, rasio, pengalaman profesional, aturan kebijakan, model statistik, limit, agunan, diversifikasi, dan cadangan kerugian. Berbagai instrumen tersebut tetap penting. Namun, pertumbuhan data digital, peningkatan kemampuan komputasi, dan perkembangan algoritma

membuka kemungkinan baru untuk mengelola risiko secara lebih granular dan dinamis.

Setiap transaksi pembayaran menghasilkan data mengenai waktu, nilai, lokasi, merchant, perangkat, dan pola perilaku. Aktivitas perbankan digital menghasilkan data mengenai frekuensi akses, perubahan perangkat, permintaan layanan, dan pola penggunaan produk. Pasar keuangan menghasilkan aliran data harga, volume, order book, berita, laporan perusahaan, dan sentimen. Data alternatif dapat berasal dari logistik, perdagangan elektronik, citra, serta hubungan rantai pasok.

Data tersebut memungkinkan lembaga keuangan tidak hanya melihat apa yang telah terjadi, tetapi juga memperkirakan apa yang mungkin terjadi. Proses memperkirakan kemungkinan hasil berdasarkan pola data disebut analitik prediktif.

Analitik prediktif tidak sama dengan ramalan yang bersifat intuitif. Ia merupakan proses ilmiah yang mencakup perumusan masalah, pengumpulan data, pengembangan model, evaluasi, validasi, implementasi, dan pemantauan. Prediksi yang dihasilkan selalu mengandung ketidakpastian. Oleh karena itu, hasil analitik prediktif idealnya dinyatakan sebagai probabilitas, rentang, skenario, atau distribusi, bukan sebagai kepastian mutlak.

Penggunaan machine learning dalam manajemen risiko perbankan telah berkembang pada risiko kredit, pasar, operasional, dan likuiditas. Namun, kedalaman penerapannya tidak selalu merata. Risiko kredit dan fraud cenderung memperoleh perhatian lebih besar karena ketersediaan data dan kejelasan target prediksinya. Risiko likuiditas, risiko sistemik, serta hubungan antara model individual dan stabilitas keuangan masih membutuhkan pengembangan lebih lanjut. ([MDPI](#))

Analitik prediktif juga mengubah sifat manajemen risiko. Pendekatan konvensional sering bersifat retrospektif. Manajemen menerima laporan yang menjelaskan kerugian atau penyimpangan yang sudah terjadi. Analitik

prediktif mengarahkan perhatian kepada sinyal yang muncul sebelum kerugian terealisasi.

Perubahan ini dapat dijelaskan melalui tiga tahap:

1. **Manajemen risiko deskriptif**, menjelaskan apa yang telah terjadi.
2. **Manajemen risiko prediktif**, memperkirakan apa yang mungkin terjadi.
3. **Manajemen risiko preskriptif**, menentukan tindakan yang sebaiknya diambil.

Namun, perubahan menuju manajemen risiko prediktif bukan semata-mata perubahan teknologi. Organisasi perlu menyesuaikan proses keputusan, kompetensi, budaya, struktur tanggung jawab, dan sistem pengendalian.

Model yang mampu memprediksi gagal bayar tidak akan memberikan manfaat apabila hasilnya tidak digunakan dalam persetujuan kredit, penetapan harga, pemantauan portofolio, atau strategi penagihan. Sistem yang mendeteksi fraud tidak akan efektif apabila tidak terhubung dengan mekanisme verifikasi. Prediksi kebutuhan likuiditas tidak akan membantu apabila treasury tidak memiliki instrumen atau kewenangan untuk menyesuaikan posisi pendanaan.

Oleh karena itu, analitik prediktif perlu dipahami sebagai sistem sosio-teknis. Sistem tersebut mencakup data, algoritma, teknologi, manusia, kebijakan, organisasi, dan lingkungan regulasi.

2. Pengertian Analitik Prediktif

Analitik prediktif adalah penggunaan data, metode statistik, ekonometrika, machine learning, dan teknik komputasi untuk memperkirakan probabilitas atau nilai suatu kejadian pada masa mendatang.

Dalam konteks keuangan, analitik prediktif dapat digunakan untuk memperkirakan:

- Kemungkinan debitur gagal bayar.
- Besarnya kerugian apabila gagal bayar terjadi.
- Eksposur kredit pada saat gagal bayar.
- Perubahan kualitas kredit.
- Harga dan volatilitas aset.
- Potensi penurunan nilai portofolio.
- Arus kas masuk dan keluar.
- Kebutuhan likuiditas.
- Kemungkinan transaksi merupakan fraud.
- Potensi terjadinya pencucian uang.
- Risiko perusahaan mengalami kesulitan keuangan.
- Kemungkinan nasabah meninggalkan lembaga keuangan.
- Kemungkinan klaim asuransi.
- Dampak skenario ekonomi terhadap portofolio.

Secara konseptual, model prediktif berusaha mempelajari fungsi:

$$\begin{bmatrix} Y = f(X) + \epsilon \end{bmatrix}$$

di mana:

- (Y) adalah hasil yang ingin diprediksi.
- (X) adalah sekumpulan variabel prediktor.
- (f) adalah pola hubungan yang dipelajari model.
- (ϵ) adalah unsur ketidakpastian atau kesalahan.

Dalam risiko kredit, (Y) dapat berupa status gagal bayar. Dalam risiko pasar, (Y) dapat berupa return atau volatilitas. Dalam risiko likuiditas, (Y) dapat berupa arus kas bersih. Dalam fraud detection, (Y) dapat berupa kategori transaksi sah atau tidak sah.

Model prediktif tidak selalu bertujuan memberikan satu angka. Model dapat menghasilkan:

- Probabilitas.
- Skor risiko.
- Kategori.
- Nilai kontinu.
- Interval prediksi.
- Peringkat.
- Sinyal peringatan.
- Skenario.

Sebagai contoh, model kredit dapat menghasilkan probabilitas gagal bayar sebesar 8 persen dalam 12 bulan. Model fraud dapat memberikan skor risiko 0 sampai 1. Model likuiditas dapat memberikan rentang kemungkinan penarikan dana harian.

Prediksi perlu dibedakan dari penjelasan kausal. Sebuah model dapat mengetahui bahwa variabel tertentu sangat berguna untuk memprediksi gagal bayar tanpa membuktikan bahwa variabel tersebut menyebabkan gagal bayar. Perbedaan ini sangat penting.

Misalnya, frekuensi keterlambatan pembayaran dapat menjadi prediktor kuat. Namun, intervensi untuk mengubah frekuensi tersebut tidak otomatis menghilangkan penyebab mendasar, seperti penurunan pendapatan atau tekanan arus kas.

Analitik prediktif menjawab pertanyaan:

Apa yang kemungkinan akan terjadi?

Analitik kausal menjawab:

Apa yang akan berubah apabila organisasi melakukan suatu tindakan?

Keduanya dapat digunakan bersama. Model prediktif dapat mengidentifikasi nasabah berisiko, sedangkan analisis kausal dapat menentukan jenis intervensi yang efektif bagi nasabah tersebut.

3. Analitik Prediktif sebagai Bagian dari Data Science

Data Science mencakup proses yang lebih luas daripada pemodelan prediktif. Data Science melibatkan:

- Perumusan masalah.
- Pengumpulan dan integrasi data.
- Pembersihan data.
- Eksplorasi.
- Rekayasa fitur.
- Pemodelan.
- Evaluasi.
- Interpretasi.
- Implementasi.
- Pemantauan.
- Komunikasi.
- Tata kelola.

Analitik prediktif merupakan salah satu hasil utama dari proses tersebut. Dengan demikian, model prediktif bukan produk yang berdiri sendiri. Kualitasnya bergantung pada keseluruhan ekosistem data.

Sebuah algoritma yang canggih tidak dapat memperbaiki data yang salah secara sistematis. Apabila tanggal gagal bayar tidak dicatat dengan konsisten, model survival akan memberikan estimasi yang tidak dapat dipercaya. Apabila transaksi fraud hanya dilabeli setelah penyelidikan panjang dan banyak kasus tidak terdeteksi, model akan belajar dari label yang tidak lengkap.

Data Science dalam keuangan juga membutuhkan pengetahuan domain. Definisi gagal bayar, eksposur, kerugian, likuiditas, dan materialitas tidak dapat ditentukan oleh algoritma. Definisi tersebut berkaitan dengan kebijakan, akuntansi, regulasi, dan strategi lembaga.

Analitik prediktif yang baik karena itu memerlukan integrasi empat kemampuan:

1. **Kemampuan statistik dan metodologis.**
2. **Kemampuan teknologi dan rekayasa data.**
3. **Pengetahuan keuangan dan manajemen risiko.**
4. **Kemampuan tata kelola, komunikasi, dan etika.**

Ketiadaan salah satu kemampuan tersebut dapat mengurangi nilai model.

Tim yang kuat secara teknis tetapi lemah dalam keuangan dapat membangun target yang salah. Tim yang memahami keuangan tetapi tidak memiliki rekayasa data dapat menghasilkan model yang tidak dapat dioperasikan. Tim yang menghasilkan model akurat tetapi tidak mampu menjelaskan hasilnya dapat menghadapi penolakan pengguna dan regulator.

4. Risiko dalam Keuangan

Risiko keuangan merupakan kemungkinan terjadinya hasil yang berbeda dari harapan dan menimbulkan kerugian atau menghambat pencapaian tujuan organisasi.

Dalam perbankan, risiko yang umum dikelola mencakup:

- Risiko kredit.
- Risiko pasar.
- Risiko likuiditas.
- Risiko operasional.
- Risiko hukum.
- Risiko reputasi.
- Risiko strategis.
- Risiko kepatuhan.

Di Indonesia, kerangka manajemen risiko bank umum mengatur penerapan manajemen risiko atas delapan jenis risiko tersebut. Lembaga keuangan nonbank dapat memiliki klasifikasi tambahan sesuai karakteristik kegiatannya. ([OJK Portal](#))

4.1 Risiko Kredit

Risiko kredit adalah risiko kerugian akibat kegagalan debitur atau pihak lain memenuhi kewajiban.

Risiko kredit muncul pada:

- Kredit konsumsi.
- Kredit usaha.
- Pembiayaan korporasi.
- Obligasi.
- Derivatif.
- Trade finance.
- Pinjaman antarbank.

- Pembiayaan digital.

4.2 Risiko Pasar

Risiko pasar muncul akibat perubahan:

- Suku bunga.
- Nilai tukar.
- Harga saham.
- Harga komoditas.
- Volatilitas.
- Spread kredit.

4.3 Risiko Likuiditas

Risiko likuiditas terjadi ketika organisasi tidak mampu memenuhi kewajiban yang jatuh tempo tanpa menanggung biaya yang tidak dapat diterima atau mengganggu kondisi keuangannya.

4.4 Risiko Operasional

Risiko operasional berasal dari:

- Kegagalan proses.
- Kesalahan manusia.
- Gangguan sistem.
- Serangan siber.
- Fraud.
- Kejadian eksternal.
- Kegagalan pihak ketiga.

4.5 Risiko Model

Risiko model adalah risiko kerugian atau keputusan yang tidak tepat akibat:

- Model yang salah.
- Asumsi yang tidak sesuai.
- Data yang buruk.
- Implementasi yang keliru.
- Penggunaan di luar tujuan.
- Interpretasi yang salah.
- Perubahan lingkungan.

Semakin besar ketergantungan terhadap model, semakin penting risiko model. Pedoman manajemen risiko model menekankan pengembangan yang kuat, validasi independen, tata kelola, dokumentasi, dan pengendalian penggunaan model. ([Federal Reserve](#))

4.6 Risiko Sistemik

Risiko sistemik berkaitan dengan gangguan yang menyebar ke seluruh sistem keuangan. Model yang serupa dan strategi otomatis yang digunakan banyak institusi dapat menghasilkan perilaku seragam. Ketika kondisi memburuk, model dapat mendorong penjualan aset, pengetatan kredit, atau penarikan likuiditas secara bersamaan.

Analitik prediktif karena itu memiliki dua sisi. Ia dapat memperkuat manajemen risiko individu, tetapi penggunaan yang homogen dapat menciptakan risiko kolektif.

5. Dari Manajemen Risiko Reaktif Menuju Prediktif

Manajemen risiko reaktif merespons kejadian setelah sinyal risiko menjadi nyata. Kredit masuk proses penagihan setelah mengalami tunggakan.

Fraud diselidiki setelah nasabah melaporkan kehilangan. Likuiditas dicari setelah terjadi kekurangan kas.

Pendekatan prediktif berusaha mengenali perubahan sebelum kejadian utama terjadi.

Dalam risiko kredit, sinyal awal dapat berupa:

- Penurunan saldo.
- Peningkatan penggunaan fasilitas.
- Keterlambatan kecil yang berulang.
- Penurunan omzet.
- Perubahan pola pembayaran.
- Penurunan kualitas pemasok.
- Meningkatnya permintaan restrukturisasi.

Dalam fraud, sinyal dapat berupa:

- Perangkat baru.
- Lokasi tidak biasa.
- Kecepatan transaksi tinggi.
- Merchant yang tidak sesuai pola.
- Percobaan autentikasi berulang.
- Perubahan penerima dana.
- Hubungan jaringan mencurigakan.

Dalam likuiditas, sinyal dapat berupa:

- Konsentrasi depasan meningkat.
- Sensitivitas dana terhadap suku bunga.
- Peningkatan penarikan digital.

- Memburuknya kondisi pasar pendanaan.
- Penurunan nilai aset likuid.
- Peningkatan kebutuhan margin.

Analitik prediktif mengubah sinyal tersebut menjadi indikator kuantitatif. Namun, organisasi perlu menghindari anggapan bahwa semua sinyal selalu bermakna. Sebagian sinyal merupakan gangguan acak.

Tantangan utama adalah membedakan:

- Signal dari noise.
- Risiko nyata dari alarm palsu.
- Perubahan sementara dari perubahan struktural.
- Hubungan prediktif dari hubungan semu.

6. Data Keuangan sebagai Fondasi Prediksi

Kualitas prediksi sangat dipengaruhi oleh kualitas dan relevansi data.

6.1 Data Internal

Data internal dapat mencakup:

- Profil nasabah.
- Rekening.
- Kredit.
- Pembayaran.
- Transaksi.
- Agunan.
- Kolektibilitas.
- Keluhan.

- Interaksi layanan.
- Laporan keuangan.
- Sistem treasury.
- Kejadian operasional.
- Log keamanan.

Keunggulan data internal adalah kedekatannya dengan proses bisnis. Kelemahannya adalah fragmentasi dan ketidakkonsistenan.

6.2 Data Eksternal

Data eksternal dapat mencakup:

- Data biro kredit.
- Data pasar.
- Indikator makroekonomi.
- Data industri.
- Harga komoditas.
- Data perdagangan.
- Data cuaca.
- Data geografis.
- Berita.
- Media sosial.
- Data rantai pasok.

6.3 Data Terstruktur

Data terstruktur tersimpan dalam tabel dengan variabel yang jelas.

Contohnya:

- Nilai kredit.
- Suku bunga.
- Saldo.
- Tanggal.
- Pendapatan.
- Rasio keuangan.

6.4 Data Tidak Terstruktur

Data tidak terstruktur meliputi:

- Laporan tahunan.
- Berita.
- Catatan analis.
- Percakapan layanan.
- Dokumen kredit.
- Kontrak.
- Surat elektronik.
- Audio.
- Gambar.

Natural language processing memungkinkan data tersebut dimasukkan ke dalam model.

6.5 Data Alternatif

Data alternatif digunakan untuk melengkapi informasi tradisional.

Contohnya:

- Data transaksi digital.

- Pola pembayaran utilitas.
- Arus perdagangan elektronik.
- Data perangkat.
- Citra satelit.
- Data logistik.
- Hubungan rantai pasok.

Data alternatif dapat membantu menilai nasabah yang memiliki sejarah kredit terbatas. Namun, penggunaannya harus relevan, proporsional, dan tidak diskriminatif.

6.6 Data Real-Time

Data real-time penting bagi:

- Fraud detection.
- Intraday liquidity.
- Market risk.
- Cybersecurity.
- Algorithmic trading.
- Payment monitoring.

Pemrosesan real-time memerlukan arsitektur streaming, latensi rendah, dan aturan respons yang jelas.

7. Kualitas Data dan Risiko Informasi

Data keuangan sering dianggap objektif karena berbentuk angka. Namun, angka merupakan hasil dari proses pencatatan, definisi, kebijakan, dan keputusan manusia.

Dimensi kualitas data meliputi:

- Akurasi.
- Kelengkapan.
- Konsistensi.
- Ketepatan waktu.
- Validitas.
- Keunikan.
- Relevansi.
- Ketertelusuran.

7.1 Kesalahan Definisi

Istilah “gagal bayar” dapat memiliki definisi berbeda. Ketidakkonsistenan definisi menghasilkan label yang tidak seragam.

7.2 Data Hilang

Data hilang dapat terjadi karena:

- Sistem tidak mencatat.
- Nasabah tidak memberikan informasi.
- Variabel baru diperkenalkan.
- Data tidak dapat diintegrasikan.
- Data hilang secara selektif.

Menangani data hilang dengan mengganti nilai rata-rata tidak selalu tepat. Pola hilangnya data dapat mengandung informasi.

7.3 Duplikasi

Nasabah dapat tercatat dalam beberapa sistem dengan identitas berbeda. Duplikasi dapat menyebabkan estimasi eksposur dan perilaku yang salah.

7.4 Bias Historis

Data masa lalu dapat mencerminkan ketidakadilan masa lalu. Apabila kelompok tertentu memiliki akses kredit lebih terbatas, data historis dapat membuat model memperkirakan kelompok tersebut sebagai risiko tinggi.

7.5 Label Delay

Dalam fraud dan anti-money laundering, label baru tersedia setelah investigasi. Model dilatih menggunakan data yang terlambat dan tidak selalu lengkap.

7.6 Data Leakage

Data leakage terjadi ketika informasi yang seharusnya belum tersedia pada waktu prediksi masuk ke model.

Contohnya, model gagal bayar menggunakan variabel restrukturisasi yang baru muncul setelah kondisi debitur memburuk. Model tampak akurat, tetapi tidak berguna sebagai peringatan dini.

8. Siklus Hidup Analitik Prediktif

8.1 Perumusan Masalah

Proyek harus dimulai dari keputusan yang ingin didukung.

Pertanyaan penting:

- Risiko apa yang diprediksi?
- Siapa pengguna model?
- Kapan prediksi dibutuhkan?
- Apa horizon prediksi?
- Apa tindakan setelah prediksi?
- Berapa biaya kesalahan?
- Apa batas regulasinya?

Contoh rumusan yang baik:

Mengestimasi probabilitas debitur UMKM mengalami tunggakan lebih dari 90 hari dalam 12 bulan mendatang agar unit risiko dapat melakukan pemantauan dan intervensi lebih awal.

Rumusan tersebut lebih jelas dibandingkan "membuat model risiko kredit".

8.2 Penentuan Target

Target harus didefinisikan secara konsisten.

Contoh target:

- Default dalam 12 bulan.
- Fraud terkonfirmasi.
- Arus kas bersih negatif.
- Penurunan rating.
- Kerugian pasar melebihi limit.
- Financial distress dalam dua tahun.

8.3 Penentuan Unit Analisis

Unit dapat berupa:

- Nasabah.
- Rekening.
- Kredit.
- Transaksi.
- Perusahaan.
- Portofolio.
- Cabang.
- Hari perdagangan.

8.4 Penentuan Horizon

Horizon memengaruhi desain model.

Model fraud menggunakan horizon beberapa milidetik atau detik. Model likuiditas dapat menggunakan horizon harian atau bulanan. Model financial distress dapat menggunakan horizon satu sampai tiga tahun.

8.5 Pengumpulan dan Integrasi

Data dikumpulkan dari sumber relevan. Data lineage harus didokumentasikan agar setiap variabel dapat ditelusuri.

8.6 Eksplorasi

Eksplorasi bertujuan memahami:

- Distribusi.
- Hubungan.
- Tren.
- Musiman.
- Pencilan.
- Perubahan struktural.
- Ketidakseimbangan kelas.

8.7 Rekayasa Fitur

Fitur dapat berupa:

- Rasio.
- Tren.
- Perubahan.
- Frekuensi.
- Durasi.

- Konsentrasi.
- Volatilitas.
- Jarak terhadap limit.
- Variabel perilaku.
- Variabel jaringan.
- Representasi teks.

8.8 Pemodelan

Model dasar perlu dibangun sebelum model kompleks. Model dasar menjadi pembanding dan membantu menilai apakah kompleksitas memberikan nilai tambah.

8.9 Validasi

Validasi mencakup:

- Kinerja statistik.
- Kinerja ekonomi.
- Kalibrasi.
- Stabilitas.
- Interpretabilitas.
- Fairness.
- Robustness.
- Implementasi.

8.10 Deployment

Model dimasukkan ke dalam sistem operasional.

8.11 Monitoring

Monitoring mencakup:

- Perubahan data.
- Perubahan performa.
- Perubahan kalibrasi.
- Perubahan fairness.
- Kualitas data.
- Latensi.
- Kegagalan sistem.
- Override manusia.

8.12 Retirement

Model harus dihentikan apabila tidak lagi relevan, tidak stabil, atau digantikan model yang lebih baik.

9. Model Statistik Tradisional

Model statistik tradisional tetap penting dalam keuangan karena:

- Mudah dijelaskan.
- Memiliki teori yang matang.
- Relatif stabil.
- Mudah divalidasi.
- Cocok untuk kebutuhan regulasi tertentu.

9.1 Regresi Linier

Regresi linier digunakan untuk memprediksi nilai kontinu.

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \epsilon$$

Model ini mudah dipahami, tetapi mengasumsikan hubungan linier dan sensitif terhadap pencilan.

9.2 Regresi Logistik

Regresi logistik banyak digunakan untuk credit scoring.

$$P(Y=1|X) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)]}$$

Keunggulannya:

- Menghasilkan probabilitas.
- Koefisien dapat dijelaskan.
- Relatif mudah dikalibrasi.
- Mudah digunakan dalam scorecard.

Keterbatasannya:

- Hubungan perlu dispesifikasikan.
- Interaksi harus ditambahkan.
- Nonlinearitas kompleks sulit ditangkap.

9.3 Discriminant Analysis

Metode ini mengklasifikasikan observasi berdasarkan kombinasi variabel. Metode tersebut memiliki asumsi distribusi yang kadang sulit dipenuhi.

9.4 Model Survival

Survival analysis memprediksi waktu sampai suatu kejadian.

Dalam risiko kredit, model ini dapat memperkirakan:

- Waktu sampai gagal bayar.

- Waktu sampai pelunasan.
- Waktu sampai restrukturisasi.
- Waktu sampai churn.

9.5 Model Deret Waktu

Model seperti ARIMA, VAR, dan GARCH digunakan untuk:

- Return.
- Volatilitas.
- Suku bunga.
- Nilai tukar.
- Arus kas.
- Likuiditas.

9.6 Model Skor Keuangan

Model financial distress dapat menggunakan rasio:

- Likuiditas.
- Leverage.
- Profitabilitas.
- Aktivitas.
- Arus kas.

Model sederhana tetap bermanfaat sebagai baseline dan alat komunikasi.

10. Machine Learning dalam Keuangan

Machine learning mempelajari pola dari data dan dapat menangkap hubungan nonlinier serta interaksi kompleks.

10.1 Decision Tree

Pohon keputusan membagi data berdasarkan aturan.

Contoh:

- Apabila rasio utang tinggi.
- Apabila keterlambatan meningkat.
- Apabila saldo menurun.
- Maka risiko masuk kategori tertentu.

Keunggulannya adalah interpretabilitas. Kelemahannya adalah ketidakstabilan.

10.2 Random Forest

Random forest menggabungkan banyak pohon keputusan. Model ini mampu menangkap hubungan kompleks dan relatif tahan terhadap overfitting.

10.3 Gradient Boosting

Gradient boosting membangun model secara bertahap untuk memperbaiki kesalahan sebelumnya.

Algoritma populer:

- XGBoost.
- LightGBM.
- CatBoost.

Model ini sering berkinerja tinggi pada data tabular keuangan.

10.4 Support Vector Machine

SVM mencari batas pemisah optimal. Model ini dapat bekerja baik pada dimensi tinggi, tetapi interpretasinya lebih sulit.

10.5 K-Nearest Neighbors

KNN mengklasifikasikan berdasarkan observasi terdekat. Metode ini sederhana, tetapi sensitif terhadap skala dan mahal pada data besar.

10.6 Neural Networks

Jaringan saraf dapat mempelajari pola nonlinier yang kompleks.

Jaringan saraf digunakan dalam:

- Credit scoring.
- Forecasting.
- Fraud detection.
- NLP.
- Market microstructure.

10.7 Ensemble Learning

Ensemble menggabungkan beberapa model.

Tujuannya:

- Meningkatkan akurasi.
- Mengurangi varians.
- Memperbaiki stabilitas.
- Menyeimbangkan kelemahan model.

Studi benchmark credit scoring menunjukkan bahwa perbandingan model harus mencakup beberapa algoritma dan ukuran kinerja. Tidak ada algoritma yang selalu terbaik dalam semua data dan kondisi.

([ScienceDirect](#))

11. Risiko Kredit

Risiko kredit merupakan bidang penerapan analitik prediktif yang paling matang.

11.1 Komponen Kerugian Kredit

Kerugian kredit dapat dirumuskan:

$$[\\text{EL} = \\text{PD} \\times \\text{LGD} \\times \\text{EAD}]$$

di mana:

- **PD** adalah Probability of Default.
- **LGD** adalah Loss Given Default.
- **EAD** adalah Exposure at Default.
- **EL** adalah Expected Loss.

11.2 Probability of Default

PD memperkirakan kemungkinan debitur gagal bayar dalam horizon tertentu.

Variabel yang digunakan dapat mencakup:

- Riwayat pembayaran.
- Rasio utang.
- Pendapatan.
- Arus kas.
- Usia usaha.
- Leverage.
- Profitabilitas.
- Industri.
- Kondisi ekonomi.
- Perilaku transaksi.

11.3 Loss Given Default

LGD memperkirakan proporsi eksposur yang hilang setelah mempertimbangkan:

- Agunan.
- Recovery.
- Biaya penagihan.
- Waktu penyelesaian.
- Senioritas.
- Kondisi ekonomi.

11.4 Exposure at Default

EAD memperkirakan eksposur ketika default terjadi. Untuk fasilitas revolving, nasabah dapat menarik tambahan dana sebelum default.

11.5 Application Scoring

Application scoring digunakan ketika nasabah mengajukan kredit.

Model membantu:

- Menyetujui.
- Menolak.
- Meminta informasi tambahan.
- Menetapkan limit.
- Menentukan harga.

11.6 Behavioral Scoring

Behavioral scoring menggunakan perilaku setelah kredit berjalan.

Fitur dapat mencakup:

- Pola pembayaran.

- Penggunaan limit.
- Perubahan saldo.
- Interaksi dengan bank.
- Perubahan aktivitas usaha.

11.7 Collection Scoring

Collection scoring memprediksi:

- Kemungkinan membayar.
- Kemungkinan cure.
- Strategi penagihan terbaik.
- Kanal komunikasi.
- Waktu kontak.

11.8 Early Warning System

Early warning system digunakan untuk mengenali penurunan kualitas sebelum default.

Sinyal dapat mencakup:

- Pembayaran terlambat.
- Penurunan omzet.
- Peningkatan leverage.
- Berita negatif.
- Perubahan manajemen.
- Gangguan rantai pasok.
- Penurunan rating.

Penelitian menunjukkan bahwa machine learning dapat meningkatkan daya prediksi risiko kredit dengan menangkap pola nonlinear dan

mengintegrasikan data transaksi dengan informasi tradisional. Namun, peningkatan performa harus dibandingkan dengan kebutuhan interpretabilitas, kalibrasi, dan stabilitas. ([ScienceDirect](#))

12. Kasus Ilustratif Risiko Kredit UMKM

Sebuah bank memiliki portofolio kredit UMKM. Bank ingin mengembangkan model untuk memprediksi default dalam 12 bulan.

12.1 Data

Data yang digunakan:

- Usia usaha.
- Sektor.
- Omzet.
- Arus kas.
- Leverage.
- Riwayat tunggakan.
- Saldo rekening.
- Frekuensi transaksi.
- Konsentrasi pelanggan.
- Harga komoditas.
- Kondisi wilayah.

12.2 Target

Target adalah tunggakan lebih dari 90 hari dalam 12 bulan.

12.3 Rekayasa Fitur

Fitur yang dibentuk:

- Pertumbuhan omzet tiga bulan.
- Volatilitas arus kas.
- Rasio pembayaran utang.
- Penurunan saldo.
- Konsentrasi pemasukan.
- Frekuensi keterlambatan.
- Tren penggunaan limit.

12.4 Model

Model yang dibandingkan:

- Regresi logistik.
- Decision tree.
- Random forest.
- Gradient boosting.

12.5 Hasil

Gradient boosting memiliki AUC tertinggi. Namun, perbedaannya terhadap regresi logistik relatif kecil pada segmen tertentu.

Manajemen tidak langsung mengganti seluruh proses dengan gradient boosting. Bank menggunakan pendekatan bertingkat:

- Regresi logistik sebagai benchmark dan model kebijakan.
- Gradient boosting sebagai challenger.
- SHAP untuk penjelasan.
- Review manual untuk kasus tertentu.

12.6 Tindakan

Nasabah berisiko tinggi tidak otomatis ditolak. Tindakan dapat berupa:

- Pemantauan lebih sering.
- Penyesuaian limit.
- Konsultasi arus kas.
- Restrukturisasi awal.
- Penguatan agunan.
- Diversifikasi pembayaran.

Kasus ini menunjukkan bahwa prediksi perlu dihubungkan dengan strategi yang konstruktif.

13. Kalibrasi Probabilitas Kredit

Model klasifikasi sering dievaluasi berdasarkan kemampuan membedakan debitur baik dan buruk. Namun, kemampuan membedakan tidak sama dengan kemampuan menghasilkan probabilitas yang akurat.

Model dapat memberikan peringkat yang baik, tetapi probabilitasnya terlalu tinggi atau terlalu rendah.

Apabila model memberikan PD 10 persen kepada 1.000 debitur, model terkalibrasi apabila sekitar 100 debitur benar-benar default dalam horizon yang ditentukan.

Kalibrasi penting untuk:

- Cadangan.
- Harga kredit.
- Modal.
- Limit.
- Stress testing.
- Expected loss.

Metode kalibrasi meliputi:

- Platt scaling.
- Isotonic regression.
- Beta calibration.
- Calibration-in-the-large.
- Recalibration periodik.

Evaluasi kalibrasi menggunakan:

- Brier score.
- Calibration curve.
- Expected calibration error.
- Observed-to-expected ratio.

Penelitian kalibrasi credit scoring menunjukkan bahwa evaluasi model perlu mencakup kemampuan peringkat dan kualitas probabilitas. Model nonparametrik yang kuat dalam diskriminasi dapat memerlukan kalibrasi tambahan. ([ScienceDirect](#))

14. Risiko Pasar

Risiko pasar muncul dari perubahan nilai instrumen akibat perubahan faktor pasar.

14.1 Prediksi Return

Model dapat memprediksi return berdasarkan:

- Harga historis.
- Volume.
- Momentum.

- Fundamental.
- Faktor makro.
- Sentimen.
- Order flow.

Namun, pasar bersifat kompetitif. Pola yang menguntungkan dapat menghilang setelah diketahui dan digunakan.

Penelitian machine learning dalam asset pricing menunjukkan bahwa pohon dan jaringan saraf dapat menangkap interaksi nonlinier yang tidak tertangkap model linear. Namun, peningkatan statistik tidak selalu menghasilkan keuntungan setelah biaya transaksi, turnover, likuiditas, dan risiko implementasi diperhitungkan. ([OUP Academic](#))

14.2 Prediksi Volatilitas

Volatilitas penting untuk:

- VaR.
- Expected Shortfall.
- Margin.
- Hedging.
- Alokasi aset.
- Limit perdagangan.

Model dapat menggunakan:

- GARCH.
- Stochastic volatility.
- Random forest.
- Gradient boosting.
- Neural networks.

- LSTM.

14.3 Value-at-Risk

VaR mengestimasi kerugian maksimum pada tingkat keyakinan tertentu dalam kondisi yang diasumsikan.

Sebagai contoh, VaR satu hari sebesar Rp10 miliar pada tingkat 99 persen berarti bahwa berdasarkan model, kerugian diperkirakan tidak melebihi Rp10 miliar pada 99 persen hari.

VaR tidak menggambarkan rata-rata kerugian setelah ambang terlampaui.

14.4 Expected Shortfall

Expected Shortfall memperkirakan rata-rata kerugian pada ekor distribusi setelah ambang VaR terlampaui.

14.5 Prediksi Korelasi

Korelasi antar aset dapat berubah dalam krisis. Model prediktif dapat memperkirakan korelasi dinamis, tetapi ketidakstabilan hubungan harus diperhatikan.

14.6 Order Book Analytics

Deep learning dapat menganalisis struktur order book dan aliran transaksi. Penelitian pada data pasar berfrekuensi tinggi menunjukkan bahwa model deep learning dapat menangkap pola price formation lintas saham, tetapi penggunaan praktisnya tetap menghadapi risiko perubahan pasar, latensi, biaya, dan crowding. ([Rama](#))

15. Prediksi Pasar dan Batas Efisiensi

Prediksi harga pasar merupakan salah satu bidang yang paling menarik sekaligus paling berisiko disalahpahami.

Model dapat terlihat berhasil karena:

- Overfitting.
- Data leakage.
- Survivorship bias.
- Look-ahead bias.
- Pemilihan periode.
- Pengabaian biaya transaksi.
- Pengabaian delisting.
- Penggunaan data revisi.

15.1 Look-Ahead Bias

Model menggunakan informasi yang belum tersedia pada waktu keputusan.

15.2 Survivorship Bias

Data hanya mencakup perusahaan yang bertahan.

15.3 Backtest Overfitting

Banyak strategi diuji, tetapi hanya strategi terbaik dilaporkan.

15.4 Transaction Costs

Keuntungan sebelum biaya belum tentu menjadi keuntungan sesudah biaya.

15.5 Regime Change

Hubungan dalam pasar berubah akibat:

- Kebijakan moneter.
- Regulasi.
- Teknologi.
- Krisis.

- Perubahan pelaku.
- Perubahan likuiditas.

Penelitian LSTM pada saham menunjukkan hasil historis yang kuat pada sebagian periode, tetapi keuntungan berlebih dapat melemah setelah biaya dan pada periode yang lebih baru. Temuan tersebut mengingatkan bahwa performa backtest tidak menjamin performa masa depan.

([IDEAS/RePEc](#))

16. Risiko Likuiditas

Risiko likuiditas memiliki dimensi:

1. **Funding liquidity risk**, kemampuan memperoleh dana.
2. **Market liquidity risk**, kemampuan menjual aset tanpa penurunan harga besar.

16.1 Prediksi Arus Kas

Model dapat memprediksi:

- Penarikan dana.
- Pembayaran kredit.
- Pencairan fasilitas.
- Kebutuhan settlement.
- Margin call.
- Penerimaan bunga.

16.2 Perilaku Depositor

Fitur dapat meliputi:

- Konsentrasi depasan.
- Stabilitas saldo.

- Sensitivitas bunga.
- Jenis nasabah.
- Kanal digital.
- Riwayat penarikan.
- Kondisi pasar.

16.3 Intraday Liquidity

Sistem pembayaran membutuhkan likuiditas sepanjang hari. Prediksi intraday dapat membantu menentukan waktu dan jumlah pendanaan.

16.4 Liquidity Stress

Model perlu mempertimbangkan skenario:

- Penarikan besar.
- Penurunan nilai agunan.
- Penutupan pasar pendanaan.
- Kenaikan haircut.
- Gangguan sistem.
- Berita reputasi.

16.5 Tantangan

Data krisis relatif sedikit. Model yang dilatih pada periode normal dapat meremehkan risiko saat tekanan.

Kajian mengenai machine learning dalam risiko perbankan menunjukkan bahwa risiko likuiditas masih relatif kurang diteliti dibandingkan risiko kredit. Hal ini menunjukkan peluang, tetapi juga mengharuskan kehati-hatian dalam mengadopsi model kompleks. ([MDPI](#))

17. Risiko Operasional

Risiko operasional berkaitan dengan kegagalan proses, manusia, sistem, dan kejadian eksternal.

Analitik prediktif dapat digunakan untuk:

- Memprediksi gangguan sistem.
- Mendeteksi anomali proses.
- Mengidentifikasi kesalahan transaksi.
- Memperkirakan volume keluhan.
- Mendeteksi insider threat.
- Memprediksi kegagalan vendor.
- Memprioritaskan audit.

17.1 Data Risiko Operasional

Data dapat berasal dari:

- Loss event database.
- Audit.
- Keluhan.
- Log sistem.
- Ticket layanan.
- Data vendor.
- Data keamanan.
- Key risk indicators.

17.2 Masalah Frekuensi Rendah

Kerugian operasional besar relatif jarang. Data yang sedikit menyulitkan model.

Solusi:

- Scenario analysis.
- Expert judgment.
- Extreme value analysis.
- Simulation.
- Data eksternal.
- Model Bayesian.

17.3 Prediksi Kegagalan Sistem

Model dapat menggunakan:

- Kapasitas.
- Latensi.
- Error rate.
- Perubahan konfigurasi.
- Beban transaksi.
- Riwayat insiden.

17.4 Third-Party Risk

Ketergantungan pada cloud, platform, dan vendor meningkatkan risiko pihak ketiga. Model prediktif dapat memonitor:

- Service-level performance.
- Insiden.
- Konsentrasi.
- Kondisi keuangan vendor.
- Paparan siber.

18. Fraud Detection

Fraud detection merupakan masalah klasifikasi yang sangat tidak seimbang. Sebagian besar transaksi sah, sedangkan fraud merupakan bagian kecil.

18.1 Jenis Fraud

- Card fraud.
- Account takeover.
- Application fraud.
- Transfer fraud.
- Insurance fraud.
- Internal fraud.
- Merchant fraud.
- Identity fraud.

18.2 Fitur Fraud

- Nilai transaksi.
- Lokasi.
- Waktu.
- Merchant.
- Perangkat.
- IP address.
- Kecepatan transaksi.
- Perubahan perilaku.
- Hubungan penerima.
- Riwayat autentikasi.

18.3 Anomaly Detection

Anomaly detection mengenali perilaku yang berbeda dari pola normal.

Metode:

- Isolation forest.
- Autoencoder.
- One-class SVM.
- Clustering.
- Statistical rules.

18.4 Supervised Fraud Detection

Model belajar dari transaksi berlabel.

Tantangan:

- Label tidak lengkap.
- Fraud berubah.
- Kelas tidak seimbang.
- Fraudster beradaptasi.
- Kebutuhan respons cepat.

18.5 Graph Analytics

Fraud sering melibatkan jaringan rekening, perangkat, alamat, dan merchant.

Graph analytics membantu menemukan:

- Kelompok rekening terhubung.
- Circular transactions.
- Shared devices.

- Mule accounts.
- Hub mencurigakan.

18.6 Biaya Kesalahan

False negative menyebabkan fraud lolos.

False positive menyebabkan transaksi sah ditolak.

Keduanya memiliki biaya. Model harus dioptimalkan berdasarkan biaya, bukan akurasi umum.

Penelitian fraud detection menekankan pentingnya cost-sensitive learning karena nilai kerugian berbeda antar transaksi. Model yang mengurangi jumlah kesalahan belum tentu memberikan penghematan terbesar apabila tidak memperhitungkan nilai transaksi dan biaya pemeriksaan. ([ORBilu](#))

19. Ketidakseimbangan Kelas

Ketidakseimbangan kelas terjadi ketika kejadian risiko jauh lebih sedikit daripada kejadian normal.

Contohnya:

- Fraud 0,1 persen.
- Default 2 persen.
- Kegagalan sistem 0,01 persen.

Model yang selalu memprediksi "normal" dapat memiliki akurasi sangat tinggi tetapi gagal mendeteksi risiko.

Metode untuk menangani ketidakseimbangan:

- Oversampling.
- Undersampling.
- SMOTE.

- Class weighting.
- Cost-sensitive learning.
- Threshold adjustment.
- Anomaly detection.
- Ensemble khusus.

Namun, resampling harus dilakukan hanya pada data pelatihan. Melakukan resampling sebelum pembagian data dapat menyebabkan leakage.

Ukuran evaluasi yang lebih tepat:

- Precision.
- Recall.
- F1-score.
- Precision-recall AUC.
- Recall at fixed false-positive rate.
- Expected cost.
- Savings.
- Profit.

20. Anti-Money Laundering

Anti-money laundering bertujuan mengenali transaksi dan hubungan yang berpotensi digunakan untuk menyamarkan sumber dana ilegal.

20.1 Rule-Based System

Sistem tradisional menggunakan aturan:

- Nilai di atas ambang.
- Transaksi berulang.

- Negara berisiko.
- Pola structuring.
- Perubahan profil.

Aturan mudah dijelaskan, tetapi dapat menghasilkan banyak alarm palsu.

20.2 Machine Learning

Machine learning dapat:

- Memprioritaskan alert.
- Mengelompokkan pola.
- Mendeteksi jaringan.
- Menganalisis teks.
- Mengidentifikasi perubahan perilaku.

20.3 Graph-Based AML

Graph analytics penting karena pencucian uang melibatkan relasi.

Node dapat berupa:

- Nasabah.
- Rekening.
- Perusahaan.
- Perangkat.
- Alamat.
- Beneficial owner.

Edge dapat berupa:

- Transfer.
- Kepemilikan.

- Penggunaan perangkat.
- Hubungan bisnis.

20.4 Human Investigation

Model tidak dapat menggantikan penyelidikan. Sistem memberikan prioritas, sedangkan analis menilai konteks.

21. Financial Distress dan Kebangkrutan

Analitik prediktif dapat memperkirakan kemungkinan perusahaan mengalami:

- Kesulitan keuangan.
- Penurunan rating.
- Restrukturisasi.
- Pelanggaran covenant.
- Kebangkrutan.

21.1 Data Keuangan

- Profitabilitas.
- Leverage.
- Likuiditas.
- Arus kas.
- Efisiensi.
- Pertumbuhan.
- Struktur jatuh tempo.

21.2 Data Pasar

- Return.

- Volatilitas.
- Market capitalization.
- Credit spread.
- Trading volume.

21.3 Data Teks

- Laporan tahunan.
- Catatan auditor.
- Berita.
- Earnings call.
- Pengungkapan risiko.

21.4 Data Rantai Pasok

- Konsentrasi pelanggan.
- Konsentrasi pemasok.
- Ketergantungan logistik.
- Stabilitas mitra.

21.5 Early Warning

Model distress tidak hanya digunakan untuk memutuskan kredit. Model dapat mendukung:

- Restrukturisasi.
- Penyesuaian covenant.
- Penguatan agunan.
- Dialog dengan debitur.
- Pemantauan industri.

22. Natural Language Processing dalam Risiko Keuangan

NLP memungkinkan model memanfaatkan informasi teks.

22.1 Sentiment Analysis

Sentimen dapat dianalisis dari:

- Berita.
- Laporan perusahaan.
- Media sosial.
- Transkrip manajemen.
- Laporan analis.

22.2 Topic Modeling

Topic modeling dapat menemukan tema:

- Likuiditas.
- Hukum.
- Persaingan.
- Rantai pasok.
- Regulasi.
- Lingkungan.
- Tata kelola.

22.3 Named Entity Recognition

NER mengenali:

- Nama perusahaan.
- Individu.

- Lokasi.
- Produk.
- Nilai.
- Tanggal.

22.4 Document Classification

Dokumen dapat diklasifikasikan menjadi:

- Risiko tinggi.
- Risiko rendah.
- Keluhan.
- Fraud.
- Kepatuhan.
- Jenis kontrak.

22.5 Large Language Models

LLM dapat membantu:

- Merangkum dokumen.
- Mengekstrak informasi.
- Menghasilkan penjelasan.
- Membantu query data.
- Mendukung investigasi.

Namun, LLM dapat menghasilkan informasi yang salah. Dalam keputusan keuangan, hasilnya harus diverifikasi terhadap sumber.

23. Stress Testing

Stress testing menilai ketahanan terhadap skenario ekstrem tetapi masuk akal.

23.1 Credit Stress Testing

Skenario dapat mencakup:

- Kontraksi ekonomi.
- Peningkatan pengangguran.
- Penurunan harga komoditas.
- Kenaikan suku bunga.
- Depresiasi nilai tukar.

Model menerjemahkan variabel makro menjadi:

- PD.
- LGD.
- NPL.
- Kerugian.
- Modal.

23.2 Market Stress Testing

Skenario dapat mencakup:

- Penurunan pasar.
- Lonjakan volatilitas.
- Pelebaran spread.
- Perubahan korelasi.
- Penurunan likuiditas.

23.3 Liquidity Stress Testing

Skenario:

- Penarikan simpanan.
- Penutupan pasar.
- Penurunan nilai aset.
- Peningkatan collateral call.

23.4 Machine Learning dalam Stress Testing

Machine learning dapat menangkap hubungan nonlinear antara kondisi makro dan kinerja bank. Namun, model kompleks menghadapi masalah:

- Data krisis sedikit.
- Extrapolation.
- Penjelasan sulit.
- Ketidakpastian skenario.
- Model drift.

Penelitian stress testing menunjukkan bahwa beberapa metode machine learning dapat meningkatkan forecasting variabel bank dibandingkan model linear. Akan tetapi, model tetap perlu diuji pada skenario di luar distribusi historis. ([ScienceDirect](#))

24. Evaluasi Model Prediktif

Evaluasi model harus disesuaikan dengan tujuan.

24.1 Confusion Matrix

Confusion matrix terdiri atas:

- True positive.
- True negative.

- False positive.
- False negative.

24.2 Accuracy

$$\left[\begin{array}{l} \text{Accuracy} = \\ \frac{TP+TN}{TP+TN+FP+FN} \end{array} \right]$$

Akurasi tidak cocok sebagai ukuran tunggal pada kelas tidak seimbang.

24.3 Precision

$$\left[\begin{array}{l} \text{Precision} = \\ \frac{TP}{TP+FP} \end{array} \right]$$

Precision menjawab: dari seluruh alarm, berapa yang benar?

24.4 Recall

$$\left[\begin{array}{l} \text{Recall} = \\ \frac{TP}{TP+FN} \end{array} \right]$$

Recall menjawab: dari seluruh kasus risiko, berapa yang berhasil ditemukan?

24.5 Specificity

$$\left[\begin{array}{l} \text{Specificity} = \\ \frac{TN}{TN+FP} \end{array} \right]$$

24.6 F1-Score

$$[$$
$$F1 =$$
$$2 \times$$
$$\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
$$]$$

24.7 ROC-AUC

ROC-AUC mengukur kemampuan model membedakan kelas pada berbagai ambang.

24.8 Precision-Recall AUC

PR-AUC lebih informatif ketika kelas positif sangat jarang.

24.9 KS Statistic

KS sering digunakan dalam credit scoring untuk mengukur pemisahan distribusi skor.

24.10 Gini

Gini berkaitan dengan AUC:

$$[$$
$$\text{Gini} = 2(\text{AUC}) - 1$$
$$]$$

24.11 Brier Score

Brier score mengukur kesalahan probabilitas:

$$[$$
$$\text{BS} =$$
$$\frac{1}{N}$$
$$\sum_{i=1}^N (p_i - y_i)^2$$
$$]$$

24.12 Mean Absolute Error

$$\left[\begin{aligned} &\text{MAE} = \\ &\frac{1}{N} \\ &\sum |y_i - \hat{y}_i| \end{aligned} \right]$$

24.13 Root Mean Squared Error

$$\left[\begin{aligned} &\text{RMSE} = \\ &\sqrt{\frac{1}{N} \\ &\sum (y_i - \hat{y}_i)^2} \\ &\} \\ &\} \end{aligned} \right]$$

24.14 Economic Value

Ukuran ekonomi dapat mencakup:

- Kerugian yang dihindari.
- Pendapatan.
- Biaya investigasi.
- Capital saved.
- Return-adjusted risk.
- Customer friction.

25. Pemilihan Threshold

Model klasifikasi menghasilkan skor atau probabilitas. Organisasi harus menentukan ambang keputusan.

Contoh:

- Skor di atas 0,8 ditolak.
- Skor 0,5 sampai 0,8 direview.
- Skor di bawah 0,5 disetujui.

Threshold optimal bergantung pada:

- Biaya false positive.
- Biaya false negative.
- Kapasitas review.
- Risk appetite.
- Regulasi.
- Dampak nasabah.

Dalam fraud detection, threshold tinggi mengurangi false positive tetapi meningkatkan fraud yang lolos.

Dalam kredit, threshold rendah dapat memperluas akses, tetapi meningkatkan kerugian.

Threshold tidak boleh dipilih hanya berdasarkan metrik statistik. Ia merupakan keputusan manajemen risiko.

26. Validasi yang Menghormati Waktu

Data keuangan memiliki urutan waktu. Pembagian data secara acak dapat menghasilkan estimasi performa yang terlalu optimistis.

Pendekatan yang lebih tepat:

- Train pada periode lama.
- Validate pada periode berikutnya.
- Test pada periode terbaru.

26.1 Rolling Validation

Model dilatih dan diuji menggunakan jendela waktu yang bergerak.

26.2 Expanding Window

Data pelatihan bertambah seiring waktu.

26.3 Out-of-Time Validation

Model diuji pada periode yang benar-benar terpisah.

26.4 Crisis Validation

Model diuji pada periode tekanan, apabila data tersedia.

Validasi harus meniru kondisi penggunaan nyata.

27. Overfitting

Overfitting terjadi ketika model mempelajari noise.

Gejala:

- Kinerja training sangat tinggi.
- Kinerja test menurun.
- Fitur terlalu banyak.
- Model tidak stabil.
- Penjelasan berubah drastis.

Pencegahan:

- Cross-validation.
- Regularization.
- Pruning.
- Early stopping.

- Feature selection.
- Simpler model.
- Independent test.
- Time-based validation.

Dalam keuangan, overfitting berbahaya karena data relatif bising dan rezim berubah.

28. Model Drift

Model drift dapat berbentuk:

28.1 Data Drift

Distribusi prediktor berubah.

28.2 Concept Drift

Hubungan antara prediktor dan target berubah.

28.3 Label Drift

Proporsi kejadian berubah.

28.4 Performance Drift

Kinerja menurun.

Penyebab:

- Perubahan ekonomi.
- Kebijakan baru.
- Perubahan produk.
- Perubahan perilaku.
- Adaptasi fraudster.

- Perubahan sistem data.

Monitoring drift menggunakan:

- Population Stability Index.
- Kolmogorov–Smirnov.
- Jensen-Shannon divergence.
- Perubahan AUC.
- Perubahan calibration.
- Perubahan false-positive rate.

Model perlu direkalibrasi atau dilatih ulang ketika perubahan material terjadi.

29. Explainable Artificial Intelligence

Model keuangan sering memengaruhi hak, akses, harga, dan peluang individu. Karena itu, penjelasan penting.

29.1 Interpretabilitas Intrinsik

Model seperti regresi atau pohon kecil relatif dapat dipahami.

29.2 Penjelasan Post-Hoc

Model kompleks dijelaskan setelah prediksi.

Metode:

- SHAP.
- LIME.
- Partial dependence.
- Feature importance.
- Counterfactual explanation.

29.3 Penjelasan Global

Menjelaskan model secara keseluruhan.

29.4 Penjelasan Lokal

Menjelaskan satu keputusan.

29.5 SHAP

SHAP mengalokasikan kontribusi fitur terhadap prediksi.

Contoh:

- Riwayat tunggakan meningkatkan risiko.
- Arus kas stabil menurunkan risiko.
- Leverage meningkatkan risiko.

29.6 Counterfactual

Counterfactual menjelaskan perubahan yang dapat mengubah hasil.

Contoh:

Permohonan dapat memenuhi kriteria apabila rasio utang turun dan riwayat pembayaran membaik.

29.7 Keterbatasan XAI

Penjelasan post-hoc tidak selalu identik dengan mekanisme model.

Penjelasan dapat sensitif terhadap data dan parameter.

Penelitian explainable machine learning dalam credit risk menunjukkan bahwa metode XAI dapat meningkatkan transparansi, tetapi penjelasan harus disesuaikan dengan kebutuhan manajemen, validator, regulator, dan nasabah. ([Springer Link](#))

30. Fairness dan Bias

Model dapat menghasilkan dampak berbeda bagi kelompok.

30.1 Sumber Bias

- Sampling bias.
- Historical bias.
- Measurement bias.
- Label bias.
- Proxy variables.
- Model specification.
- Threshold policy.

30.2 Proxy

Variabel lokasi dapat menjadi proksi bagi kondisi sosial tertentu. Variabel perangkat dapat berkaitan dengan pendapatan.

30.3 Ukuran Fairness

- Demographic parity.
- Equal opportunity.
- Equalized odds.
- Predictive parity.
- Calibration by group.

Ukuran tersebut dapat bertentangan. Tidak ada satu ukuran fairness yang cocok untuk semua konteks.

30.4 Fairness dan Risiko

Menghapus variabel sensitif tidak selalu menghilangkan bias karena variabel lain dapat menjadi proksi.

30.5 Tindakan

- Audit data.

- Uji per kelompok.
- Review fitur.
- Fairness constraints.
- Threshold review.
- Human appeal.
- Monitoring.

Fairness perlu dipahami sebagai persoalan teknis, hukum, etis, dan kebijakan.

31. Privasi dan Pelindungan Data

Analitik prediktif keuangan menggunakan data pribadi yang sensitif.

Prinsip penting:

- Tujuan yang sah.
- Persetujuan atau dasar pemrosesan.
- Minimalisasi data.
- Pembatasan tujuan.
- Keamanan.
- Transparansi.
- Retensi terbatas.
- Hak subjek data.

Undang-Undang Nomor 27 Tahun 2022 mengatur jenis data pribadi, hak subjek data, kewajiban pengendali dan prosesor, pemrosesan, transfer, serta sanksi. Lembaga keuangan perlu memastikan bahwa pengembangan dan penggunaan model konsisten dengan kewajiban tersebut. ([Database Peraturan | JDIH BPK](#))

Teknik perlindungan dapat meliputi:

- Anonymization.
- Pseudonymization.
- Encryption.
- Access control.
- Differential privacy.
- Federated learning.
- Secure computation.

Namun, anonimisasi tidak selalu permanen. Penggabungan berbagai data dapat memungkinkan identifikasi ulang.

32. Model Risk Management

Model risk management mengelola risiko yang timbul dari penggunaan model.

32.1 Model Inventory

Organisasi harus mengetahui:

- Model apa yang digunakan.
- Pemilik model.
- Tujuan.
- Versi.
- Tingkat materialitas.
- Status validasi.
- Tanggal review.

32.2 Model Development

Dokumentasi harus mencakup:

- Tujuan.
- Data.
- Asumsi.
- Metode.
- Fitur.
- Keterbatasan.
- Kinerja.
- Penggunaan.

32.3 Independent Validation

Validator harus independen dari pengembang.

Validasi mencakup:

- Conceptual soundness.
- Process verification.
- Outcomes analysis.
- Benchmarking.
- Sensitivity.
- Stress testing.

32.4 Model Approval

Model material memerlukan persetujuan formal.

32.5 Change Control

Perubahan kode, fitur, threshold, dan data harus dikendalikan.

32.6 Usage Control

Model tidak boleh digunakan di luar tujuan tanpa penilaian.

32.7 Override

Keputusan manusia yang berbeda dari model perlu dicatat.

Analisis override dapat menunjukkan:

- Model kurang lengkap.
- Kebijakan tidak jelas.
- Pengguna tidak percaya.
- Bias manusia.
- Kondisi khusus.

33. Tata Kelola AI di Perbankan Indonesia

Penggunaan AI dalam perbankan Indonesia berkembang dalam konteks penguatan tata kelola dan transformasi digital.

OJK meluncurkan panduan Tata Kelola Kecerdasan Artifisial Perbankan Indonesia pada 29 April 2025. Panduan tersebut menekankan pengembangan dan penerapan AI secara bertanggung jawab dalam sektor perbankan. OJK juga telah menerbitkan panduan kode etik AI yang bertanggung jawab dan terpercaya bagi industri teknologi finansial. ([OJK Portal](#))

Prinsip tata kelola mencakup:

- Manajemen risiko.
- Akuntabilitas.
- Transparansi.
- Keamanan.
- Fairness.

- Human oversight.
- Perlindungan konsumen.
- Pengelolaan pihak ketiga.

Tata kelola bukan hambatan inovasi. Tata kelola menciptakan kondisi agar inovasi dapat digunakan secara berkelanjutan.

34. Human-in-the-Loop

Human-in-the-loop berarti manusia tetap terlibat dalam keputusan.

Bentuknya:

- Review kasus berisiko.
- Persetujuan keputusan material.
- Investigasi alert.
- Override.
- Penanganan banding.
- Monitoring.

Namun, human oversight dapat gagal apabila:

- Pengguna tidak memahami model.
- Pengguna terlalu percaya.
- Pengguna selalu menolak.
- Waktu review terlalu singkat.
- Penjelasan tidak memadai.

Human oversight harus didukung oleh:

- Pelatihan.

- Informasi.
 - Kewenangan.
 - Prosedur.
 - Akuntabilitas.
-

35. Arsitektur Teknologi

Sistem analitik prediktif memerlukan arsitektur yang kuat.

35.1 Sumber Data

Data berasal dari:

- Core banking.
- Payment.
- CRM.
- Market data.
- External data.
- Documents.
- Logs.

35.2 Data Ingestion

Data dimasukkan melalui:

- Batch.
- Streaming.
- API.
- File transfer.

35.3 Data Storage

- Data warehouse.
- Data lake.
- Lakehouse.
- Feature store.

35.4 Processing

- ETL.
- ELT.
- Distributed computing.
- Stream processing.

35.5 Modeling

- Notebook.
- ML platform.
- Statistical software.
- Automated machine learning.

35.6 Deployment

- Batch scoring.
- Real-time API.
- Embedded system.
- Dashboard.

35.7 Monitoring

- Performance.
- Drift.
- Latency.

- Fairness.
 - Security.
 - Availability.
-

36. MLOps

MLOps mengintegrasikan pengembangan model dengan operasi.

Komponen:

- Version control.
- Data versioning.
- Experiment tracking.
- Automated testing.
- Model registry.
- Deployment pipeline.
- Monitoring.
- Retraining.

36.1 Reproducibility

Model harus dapat direproduksi dengan:

- Data yang sama.
- Kode yang sama.
- Parameter yang sama.
- Lingkungan yang sama.

36.2 Testing

Pengujian meliputi:

- Unit test.
- Data quality test.
- Integration test.
- Performance test.
- Security test.
- Fairness test.

36.3 Champion-Challenger

Champion adalah model yang digunakan.

Challenger adalah alternatif yang diuji.

Pendekatan ini memungkinkan pembaruan terkontrol.

37. Cybersecurity dan Adversarial Risk

Model dapat menjadi target serangan.

37.1 Data Poisoning

Penyerang memasukkan data untuk merusak model.

37.2 Evasion

Penyerang mengubah pola agar lolos deteksi.

37.3 Model Extraction

Penyerang mencoba meniru model.

37.4 Membership Inference

Penyerang mencoba mengetahui apakah data tertentu digunakan.

37.5 Prompt Injection

Pada sistem LLM, input dapat dirancang untuk mengubah perilaku sistem.

Mitigasi:

- Secure pipeline.
- Access control.
- Monitoring.
- Adversarial testing.
- Input validation.
- Red teaming.
- Incident response.

Fraudster beradaptasi terhadap sistem. Fraud detection karena itu merupakan permainan dinamis.

38. Generative AI dalam Risiko Keuangan

Generative AI dapat mendukung:

- Analisis dokumen.
- Ringkasan laporan.
- Penjelasan model.
- Penyusunan laporan risiko.
- Query bahasa alami.
- Dukungan investigasi.
- Pembuatan skenario.

Risiko:

- Hallucination.
- Kebocoran informasi.

- Bias.
- Prompt injection.
- Ketergantungan vendor.
- Hak cipta.
- Auditability.

Generative AI tidak seharusnya menjadi satu-satunya sumber dalam keputusan risiko. Output perlu diverifikasi.

Basel Committee mengakui potensi AI/ML untuk meningkatkan efisiensi dan manajemen risiko, tetapi menekankan tantangan tata kelola, kualitas data, explainability, dan ketergantungan pihak ketiga. ([Bank for International Settlements](#))

39. Integrasi Prediksi dengan Keputusan

Model hanya bernilai apabila dihubungkan dengan keputusan.

39.1 Credit Decision

Prediksi dapat memengaruhi:

- Approval.
- Limit.
- Pricing.
- Collateral.
- Monitoring.

39.2 Fraud Decision

Prediksi dapat menghasilkan:

- Approve.

- Reject.
- Step-up authentication.
- Manual review.
- Temporary hold.

39.3 Liquidity Decision

Prediksi dapat menghasilkan:

- Penambahan buffer.
- Penyesuaian maturity.
- Penjualan aset.
- Pendanaan.

39.4 Market Decision

Prediksi dapat memengaruhi:

- Hedging.
- Limit.
- Position size.
- Capital allocation.

39.5 Actionability

Prediksi harus memiliki tindakan yang dapat dilakukan.

Apabila organisasi tidak memiliki tindakan, model hanya menjadi laporan.

40. Risk Appetite dan Analitik Prediktif

Risk appetite adalah tingkat risiko yang bersedia diterima organisasi.

Model prediktif membantu menerjemahkan risk appetite menjadi:

- Threshold.
- Limit.
- Alert.
- Pricing.
- Capital.
- Escalation.

Namun, model tidak menentukan risk appetite. Risk appetite merupakan keputusan strategis pimpinan.

Analitik dapat menunjukkan konsekuensi alternatif.

Contoh:

- Threshold A menghasilkan pertumbuhan lebih tinggi tetapi default meningkat.
- Threshold B mengurangi kerugian tetapi menolak lebih banyak nasabah.
- Threshold C meningkatkan review manual.

Manajemen memilih trade-off sesuai strategi dan kemampuan modal.

41. Nilai Ekonomi Model

Akurasi bukan tujuan akhir.

Nilai ekonomi dapat dihitung:

[
Net\ Benefit =
Losses\ Avoided +
Revenue\ Gained -
Implementation\ Cost -

Operational\ Cost -
Customer\ Cost
]

41.1 Losses Avoided

Kerugian yang berhasil dicegah.

41.2 Revenue Gained

Pendapatan dari keputusan yang lebih tepat.

41.3 Implementation Cost

Biaya data, teknologi, dan tenaga ahli.

41.4 Customer Cost

False positive dapat mengganggu pelanggan.

41.5 Regulatory Cost

Model yang tidak patuh dapat menimbulkan sanksi dan reputasi buruk.

42. Organisasi dan Kompetensi

Implementasi membutuhkan tim multidisiplin:

- Risk manager.
- Data scientist.
- Data engineer.
- Model validator.
- IT security.
- Legal.
- Compliance.
- Business unit.

- Internal audit.

42.1 Three Lines Model

Lini pertama:

- Pemilik dan pengguna model.

Lini kedua:

- Risk management.
- Model validation.
- Compliance.

Lini ketiga:

- Internal audit.

42.2 Kompetensi Manajer

Manajer perlu memahami:

- Probabilitas.
- Ketidakpastian.
- Metrik.
- Bias.
- Drift.
- Explainability.
- Trade-off.

Manajer tidak harus menulis kode, tetapi harus mampu mempertanyakan model.

43. Pertanyaan Kritis untuk Manajemen

Sebelum menyetujui model, manajemen perlu menanyakan:

1. Masalah apa yang diselesaikan?
2. Apa target prediksinya?
3. Kapan prediksi dibuat?
4. Data apa yang digunakan?
5. Apakah data representatif?
6. Apakah terdapat leakage?
7. Model apa yang dibandingkan?
8. Apakah performa diuji out-of-time?
9. Apakah probabilitas terkalibrasi?
10. Apa biaya kesalahan?
11. Apakah model adil?
12. Dapatkah keputusan dijelaskan?
13. Bagaimana model dipantau?
14. Siapa yang bertanggung jawab?
15. Apa rencana apabila model gagal?

44. Tantangan Implementasi

44.1 Legacy Systems

Data tersebar dalam sistem lama.

44.2 Data Silos

Unit tidak berbagi data.

44.3 Kekurangan Talenta

Kompetensi Data Science dan risiko sulit ditemukan.

44.4 Model Tidak Dipercaya

Pengguna menolak model yang tidak dipahami.

44.5 Pilot Trap

Model berhenti sebagai eksperimen.

44.6 Regulatory Uncertainty

Aturan berkembang.

44.7 Vendor Dependency

Organisasi bergantung pada model pihak ketiga.

44.8 Culture

Keputusan masih bergantung pada hierarki dan intuisi.

45. Kerangka Implementasi

Tahap 1: Strategi

- Tentukan tujuan.
- Pilih kasus bernilai tinggi.
- Selaraskan dengan risk appetite.

Tahap 2: Kesiapan

- Audit data.
- Audit teknologi.
- Audit kompetensi.
- Audit tata kelola.

Tahap 3: Pilot

- Pilih masalah terbatas.
- Bangun baseline.
- Uji manfaat.

Tahap 4: Validation

- Validasi independen.
- Uji fairness.
- Uji stress.
- Uji keamanan.

Tahap 5: Deployment

- Integrasikan proses.
- Latih pengguna.
- Tetapkan eskalasi.

Tahap 6: Monitoring

- Pantau data.
- Pantau kinerja.
- Pantau dampak.

Tahap 7: Scale

- Perluas secara bertahap.
- Gunakan reusable platform.
- Standarkan kontrol.

46. Kasus Ilustratif Fraud Detection

Sebuah bank digital memproses jutaan transaksi.

Model awal memiliki akurasi 99,8 persen. Namun, fraud tetap banyak lolos karena kelas fraud sangat kecil.

Tim mengganti ukuran utama menjadi:

- Recall fraud.
- Precision.
- Cost saved.
- False decline rate.

Model baru menggunakan:

- Perangkat.
- Kecepatan.
- Lokasi.
- Jaringan penerima.
- Pola historis.

Bank menerapkan keputusan bertingkat:

- Risiko rendah: disetujui.
- Risiko menengah: autentikasi tambahan.
- Risiko tinggi: review atau blokir.

Hasilnya tidak hanya diukur dari fraud yang dihentikan. Bank juga mengukur jumlah transaksi sah yang terganggu.

47. Kasus Ilustratif Likuiditas

Sebuah bank mengalami volatilitas penarikan dana yang meningkat.

Model memprediksi arus kas harian menggunakan:

- Saldo historis.

- Jatuh tempo.
- Suku bunga.
- Hari kalender.
- Perilaku deposan.
- Sentimen.

Model memberikan:

- Point forecast.
- Prediction interval.
- Stress estimate.

Treasury menggunakan hasil untuk menentukan buffer.

Manajemen tidak hanya menggunakan nilai rata-rata. Mereka memperhatikan ekor distribusi karena risiko likuiditas berkaitan dengan kondisi ekstrem.

48. Kasus Ilustratif Stress Testing

Sebuah bank ingin menilai dampak resesi.

Skenario:

- Pertumbuhan turun.
- Pengangguran naik.
- Suku bunga tinggi.
- Nilai tukar melemah.
- Harga properti turun.

Model memperkirakan:

- PD meningkat.

- LGD meningkat.
- NPL meningkat.
- Pendapatan bunga turun.
- Kerugian pasar meningkat.

Manajemen kemudian mengevaluasi:

- Modal.
- Likuiditas.
- Dividen.
- Lending strategy.
- Recovery plan.

Machine learning dapat membantu pemetaan nonlinear, tetapi keputusan tetap menggunakan beberapa model dan skenario.

49. Keterbatasan Analitik Prediktif

49.1 Masa Lalu Tidak Selalu Mewakili Masa Depan

Perubahan struktural dapat membuat pola historis tidak relevan.

49.2 Rare Events

Krisis besar jarang terjadi.

49.3 Reflexivity

Prediksi dapat mengubah perilaku.

49.4 Feedback Loop

Keputusan model menghasilkan data baru.

Apabila model menolak kelompok tertentu, organisasi tidak memperoleh data kinerja kelompok tersebut. Model kemudian terus menganggap kelompok tersebut tidak pasti.

49.5 Black Swan

Kejadian sangat langka sulit diprediksi.

49.6 Measurement Error

Konsep seperti risiko dan kemampuan bayar tidak dapat diukur sempurna.

49.7 Ethical Limits

Tidak semua data yang tersedia layak digunakan.

50. Masa Depan Analitik Prediktif Keuangan

50.1 Multimodal Models

Model akan menggabungkan:

- Tabel.
- Teks.
- Gambar.
- Suara.
- Jaringan.

50.2 Graph Machine Learning

Graph neural networks akan berkembang untuk fraud, AML, dan contagion.

50.3 Causal Machine Learning

Organisasi akan bergerak dari "siapa yang berisiko" menuju "tindakan apa yang efektif".

50.4 Uncertainty Quantification

Model akan lebih menekankan interval dan ketidakpastian.

50.5 Federated Learning

Institusi dapat melatih model tanpa memusatkan data mentah.

50.6 Real-Time Risk

Monitoring bergerak menuju waktu nyata.

50.7 AI Governance

Penggunaan AI akan semakin terhubung dengan tata kelola model, keamanan, fairness, dan akuntabilitas.

50.8 Human-AI Collaboration

Manusia dan AI akan berbagi tugas:

- Mesin memproses pola.
- Manusia menilai konteks.
- Mesin memberikan peringatan.
- Manusia memutuskan dalam kasus kompleks.
- Mesin memonitor.
- Manusia bertanggung jawab.

51. Implikasi bagi Pendidikan Manajemen

Pendidikan manajemen keuangan perlu mengintegrasikan:

- Statistika.
- Manajemen risiko.
- Data Science.
- Machine learning.

- Etika.
- Regulasi.
- Komunikasi.

Mahasiswa tidak cukup hanya belajar menghitung rasio atau VaR. Mereka perlu memahami:

- Bagaimana data dibentuk.
- Bagaimana model belajar.
- Bagaimana model gagal.
- Bagaimana hasil digunakan.
- Bagaimana dampak sosial dikelola.

Pendekatan pembelajaran dapat menggunakan:

- Studi kasus.
- Dataset.
- Simulasi.
- Dashboard.
- Model sederhana.
- Diskusi etika.
- Proyek integratif.

52. Refleksi Manajerial

Analitik prediktif dapat menciptakan ilusi kepastian. Skor, probabilitas, dan grafik dapat terlihat objektif. Namun, setiap model mengandung asumsi.

Manajemen perlu menjaga tiga keseimbangan.

52.1 Akurasi dan Interpretabilitas

Model kompleks dapat lebih akurat, tetapi lebih sulit dijelaskan.

52.2 Efisiensi dan Fairness

Otomatisasi dapat menurunkan biaya, tetapi dapat memperbesar dampak bias.

52.3 Inovasi dan Pengendalian

Kecepatan inovasi perlu diseimbangkan dengan validasi dan tata kelola.

Model bukan hakim yang netral. Model merupakan instrumen yang dirancang manusia, menggunakan data dari sistem sosial, dan digunakan dalam organisasi yang memiliki tujuan tertentu.

Karena itu, pertanyaan penting bukan hanya:

Seberapa akurat model ini?

Pertanyaan yang lebih lengkap adalah:

Apakah model ini akurat, stabil, terkalibrasi, adil, aman, dapat dijelaskan, dan bermanfaat bagi keputusan?

53. Kesimpulan

Analitik prediktif telah menjadi instrumen penting dalam manajemen risiko keuangan. Dengan mengintegrasikan statistika, ekonometrika, machine learning, Data Science, dan pengetahuan domain, lembaga keuangan dapat memperkirakan risiko lebih awal dan lebih granular.

Penerapannya mencakup:

- Credit scoring.
- Probability of default.
- Loss given default.
- Early warning.

- Market forecasting.
- Volatility.
- Liquidity forecasting.
- Fraud detection.
- AML.
- Financial distress.
- Stress testing.
- Operational risk.

Machine learning menawarkan kemampuan menangkap nonlinearitas, interaksi, data besar, dan data tidak terstruktur. Namun, kompleksitas tidak selalu lebih baik. Model sederhana yang stabil, terkalibrasi, dan dapat digunakan dapat lebih bernilai daripada model kompleks yang hanya unggul sedikit.

Evaluasi harus melampaui akurasi. Lembaga perlu menilai:

- Discrimination.
- Calibration.
- Stability.
- Fairness.
- Explainability.
- Economic value.
- Robustness.
- Operational readiness.

Risiko model perlu dikelola melalui:

- Inventory.

- Dokumentasi.
- Validasi independen.
- Change control.
- Monitoring.
- Human oversight.
- Audit.

Analitik prediktif tidak menghilangkan ketidakpastian. Ia membantu organisasi memahami dan mengelola ketidakpastian secara lebih sistematis.

Keberhasilan tidak ditentukan oleh jumlah data atau kecanggihan algoritma. Keberhasilan ditentukan oleh kemampuan organisasi menghubungkan prediksi dengan tindakan yang tepat, proporsional, dan bertanggung jawab.

Dengan tata kelola yang kuat, analitik prediktif dapat mengubah manajemen risiko dari fungsi pelaporan menjadi kemampuan strategis. Organisasi tidak hanya mencatat kerugian, tetapi mengenali sinyal. Organisasi tidak hanya merespons krisis, tetapi membangun kesiapan. Organisasi tidak hanya menghindari risiko, tetapi memilih risiko secara lebih cerdas.

Pada akhirnya, Data Science dalam keuangan harus melayani tujuan yang lebih luas: menjaga kepercayaan, melindungi nasabah, memperkuat ketahanan lembaga, dan mendukung sistem keuangan yang sehat.

Glosarium

Accuracy

Proporsi seluruh prediksi yang benar.

Algorithm

Prosedur komputasi untuk memproses data atau menghasilkan prediksi.

AML

Anti-Money Laundering, proses pencegahan dan deteksi pencucian uang.

Anomaly Detection

Metode untuk mendeteksi observasi yang berbeda dari pola normal.

AUC

Area Under the Curve, ukuran kemampuan model membedakan kelas.

Backtesting

Pengujian model atau strategi menggunakan data historis.

Bias

Penyimpangan sistematis dalam data, model, atau keputusan.

Brier Score

Ukuran kesalahan probabilitas prediksi.

Calibration

Kesesuaian antara probabilitas prediksi dan frekuensi aktual.

Champion Model

Model utama yang digunakan dalam operasi.

Challenger Model

Model alternatif yang dibandingkan dengan champion.

Class Imbalance

Kondisi ketika jumlah observasi antar kelas sangat berbeda.

Concept Drift

Perubahan hubungan antara prediktor dan target sepanjang waktu.

Confusion Matrix

Tabel yang membandingkan kelas prediksi dengan kelas aktual.

Counterfactual Explanation

Penjelasan mengenai perubahan yang diperlukan untuk mengubah keputusan model.

Credit Scoring

Proses menilai risiko kredit menggunakan skor.

Data Drift

Perubahan distribusi data masukan.

Data Leakage

Masuknya informasi yang tidak semestinya ke proses pelatihan model.

Data Lineage

Catatan asal dan perjalanan data.

Data Science

Bidang interdisipliner untuk menghasilkan wawasan dan keputusan dari data.

Decision Tree

Model prediksi berbasis aturan bercabang.

EAD

Exposure at Default, eksposur ketika default terjadi.

Early Warning System

Sistem untuk mengidentifikasi sinyal risiko sebelum kejadian.

Ensemble Learning

Penggabungan beberapa model.

Expected Loss

Kerugian kredit yang diperkirakan.

Expected Shortfall

Rata-rata kerugian pada bagian ekor distribusi.

Explainable AI

Pendekatan untuk membuat keputusan AI dapat dipahami.

Fairness

Prinsip untuk menghindari dampak yang tidak adil.

False Negative

Kasus positif yang diprediksi negatif.

False Positive

Kasus negatif yang diprediksi positif.

Feature

Variabel yang digunakan model.

Feature Engineering

Proses membentuk variabel prediktor.

Fraud Detection

Proses mendeteksi transaksi atau tindakan penipuan.

Gradient Boosting

Metode ensemble yang memperbaiki kesalahan secara bertahap.

Human-in-the-Loop

Sistem yang mempertahankan keterlibatan manusia.

KS Statistic

Ukuran pemisahan distribusi skor dua kelompok.

LGD

Loss Given Default, proporsi kerugian ketika default terjadi.

Liquidity Risk

Risiko ketidakmampuan memenuhi kewajiban.

Machine Learning

Metode yang memungkinkan sistem mempelajari pola dari data.

Market Risk

Risiko kerugian akibat perubahan faktor pasar.

Model Drift

Penurunan relevansi atau kinerja model akibat perubahan.

Model Risk

Risiko kerugian karena kesalahan atau penyalahgunaan model.

MLOps

Praktik pengembangan dan pengoperasian machine learning.

Natural Language Processing

Teknologi untuk mengolah bahasa manusia.

Neural Network

Model komputasi berlapis untuk mempelajari pola nonlinear.

Operational Risk

Risiko akibat kegagalan proses, manusia, sistem, atau kejadian eksternal.

Out-of-Time Validation

Pengujian model pada periode waktu yang berbeda.

Overfitting

Kondisi ketika model terlalu menyesuaikan data pelatihan.

PD

Probability of Default, probabilitas gagal bayar.

Precision

Proporsi prediksi positif yang benar.

Predictive Analytics

Penggunaan data dan model untuk memperkirakan kejadian.

PR-AUC

Area under the precision-recall curve.

Random Forest

Ensemble dari banyak pohon keputusan.

Recall

Proporsi kasus positif yang berhasil ditemukan.

Risk Appetite

Tingkat risiko yang bersedia diterima organisasi.

ROC-AUC

Ukuran kemampuan model membedakan kelas pada berbagai threshold.

SHAP

Metode penjelasan berdasarkan kontribusi fitur.

Stress Testing

Pengujian ketahanan pada skenario tekanan.

Support Vector Machine

Algoritma yang mencari batas pemisah kelas.

Threshold

Ambang untuk mengubah skor menjadi keputusan.

Time Series

Data yang memiliki urutan waktu.

Value-at-Risk

Estimasi batas kerugian pada tingkat keyakinan tertentu.

Daftar Pustaka

- Agarwal, R., & Dhar, V. (2014). Editorial—Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, 25(3), 443–448. <https://doi.org/10.1287/isre.2014.0546>
- Bahnsen, A. C., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51, 134–142. <https://doi.org/10.1016/j.eswa.2015.12.030>
- Bahnsen, A. C., Stojanovic, A., Aouada, D., & Ottersten, B. (2013). Cost-sensitive credit card fraud detection using Bayes minimum risk. In *2013 12th International Conference on Machine Learning and Applications* (pp. 333–338). IEEE. <https://doi.org/10.1109/ICMLA.2013.68>
- Bequé, A., Coussement, K., Gayler, R., & Lessmann, S. (2017). Approaches for credit scorecard calibration: An empirical analysis. *Knowledge-Based Systems*, 134, 213–227. <https://doi.org/10.1016/j.knosys.2017.07.034>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199–231. <https://doi.org/10.1214/ss/1009213726>
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics*, 57, 203–216. <https://doi.org/10.1007/s10614-020-10042-0>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>
- Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 233–240). ACM. <https://doi.org/10.1145/1143844.1143874>

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. <https://doi.org/10.1016/j.ejor.2017.11.054>

Fritz-Morgenthal, S., Hein, B., & Papenbrock, J. (2022). Financial risk management and explainable, trustworthy, responsible AI. *Frontiers in Artificial Intelligence*, 5, 779799. <https://doi.org/10.3389/frai.2022.779799>

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273. <https://doi.org/10.1093/rfs/hhaa009>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>

Höppner, S., Baesens, B., Verbeke, W., & Verdonck, T. (2022). Instance-dependent cost-sensitive learning for detecting transfer fraud. *European Journal of Operational Research*, 297(1), 291–300. <https://doi.org/10.1016/j.ejor.2021.05.022>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management*, 48(4), 1009–1029. <https://doi.org/10.1111/fima.12295>

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11), 2767–2787. <https://doi.org/10.1016/j.jbankfin.2010.06.001>

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.

<https://doi.org/10.1007/978-1-4614-6849-3>

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. *Risks*, 7(1), Article 29.

<https://doi.org/10.3390/risks7010029>

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124–136. <https://doi.org/10.1016/j.ejor.2015.05.030>

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, 30, 4765–4774.

Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.).

Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. MIT Press.

Nami, S., & Shajari, M. (2018). Cost-sensitive payment card fraud detection based on dynamic random forest and k-nearest neighbors. *Expert Systems with Applications*, 110, 381–392. <https://doi.org/10.1016/j.eswa.2018.06.011>

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559–569. <https://doi.org/10.1016/j.dss.2010.08.006>

O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Provost, F., & Fawcett, T. (2013). *Data science for business*. O’Reilly Media.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM*

SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). ACM. <https://doi.org/10.1145/2939672.2939778>

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-STS330>

Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*, 19(9), 1449–1459. <https://doi.org/10.1080/14697688.2019.1622295>

Thomas, L. C., Crook, J. N., & Edelman, D. B. (2017). *Credit scoring and its applications* (2nd ed.). SIAM. <https://doi.org/10.1137/1.9781611974560>

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2015). APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decision Support Systems*, 75, 38–48. <https://doi.org/10.1016/j.dss.2015.04.013>

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. <https://doi.org/10.1111/jbl.12010>

West, D. (2000). Neural network credit scoring models. *Computers & Operations Research*, 27(11–12), 1131–1152. [https://doi.org/10.1016/S0305-0548\(99\)00149-5](https://doi.org/10.1016/S0305-0548(99)00149-5)

Yeh, I.-C., & Lien, C.-H. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2), 2473–2480. <https://doi.org/10.1016/j.eswa.2007.12.020>

Rudy C Tarumingkeng: **ANALITIK PREDIKTIF DALAM KEUANGAN - MENGELOLA RISIKO DENGAN DATA SCIENCE**

Naskah ini siap digunakan sebagai artikel akademik, bab buku **Data Science for Financial Management**, atau materi perkuliahan mengenai manajemen risiko berbasis AI dan analitik data.

Copilot for this article: ChatGPT, Access date: 14 July 2026. Author's account ([Rudy C Tarumingkeng](#)). <https://chatgpt.com/c/6a55aedb-5ee0-83ec-8a3c-cf2be8efd52e>