
PEMODELAN PROBABILITY OF DEFAULT PORTOFOLIO PEMBIAYAAN BERSAMA FINTECH LENDING DAN MULTI FINANCE: STUDI KASUS BANK ABC

Putri Metasari Pangaribuan

putri.metasari@alumni.ui.ac.id

Adi Vithara Purba

adi.vithara@hotmail.com

Fakultas Ekonomi dan Bisnis

Universitas Indonesia

ABSTRACT

This research is to find the appropriate credit scoring model technique to build the default model based on the joint financing schemes product between bank and fintech lending and multi finances that conducted by bank ABC. The credit scoring model to be compared using traditional approach, logistic regression against machine learning technique. This research is case study in Bank ABC's portfolio starting April 2019 up to December 2022 and will be classified into default or non-default so the model can predict the possibility of customer default during the period. The analysis conducted based on variables from application and transaction data that not breaching the confidentiality of personal data in Bank ABC. Furthermore, the analysis only applicable for joint financing schemes product to fintech lending and multi finances that already have operated in Indonesia only. The significant variable to model the probability of default for joint financing schemes between bank and fintech lending or multi finances are tenure, loan purpose, interest amount, job description, home city, age and declared income. Furthermore, the analysis showed that the traditional technique logistic regression had higher accuracy compared to machine learning using decision tree in the case study.

ABSTRAK

Dalam melakukan pembiayaan bersama antara bank dengan lembaga keuangan non-bank, bank terpapar salah satu risiko yaitu risiko kredit yang terjadi apabila penerima pinjaman gagal memenuhi kewajibannya. Penelitian ini bertujuan untuk mengetahui teknik pemodelan risiko kredit yang sesuai dengan cara membandingkan antara teknik tradisional regresi logistik dan machine learning pada produk pembiayaan bersama antara bank dan fintech lending/multi finance. Analisis dilakukan menggunakan kombinasi variabel aplikasi dan transaksi pada Bank ABC yang telah memiliki portfolio dengan beberapa fintech lending dan multi finance di Indonesia. Penelitian ini bersifat kuantitatif dengan periode pengamatan sejak April 2019 hingga Desember 2022 pada data panel yang diamati sejak aplikasi dibuka hingga menjadi default atau periode terakhir observasi pada Desember 2022. Validasi model dilakukan menggunakan metode in-sample dan out-sample untuk mengetahui keandalan kedua model yang digunakan. Variabel yang signifikan dalam memodelkan probability of default pada pembiayaan Bersama antara bank dan fintech lending atau multi finance adalah durasi pinjaman, tujuan pinjaman, nominal suku bunga, deskripsi pekerjaan, kota domisili, umur dan pendapatan yang dideklarasikan. Lebih lanjut, hasil penelitian menunjukkan bahwa menggunakan kombinasi variabel data aplikasi dan transaksi dengan teknik regresi logistik lebih baik dalam memodelkan default dibandingkan dengan teknik machine learning.

Kata kunci: pembiayaan bersama; regresi logistik; *machine learning*; risiko kredit; default; *fintech lending*; *multi finance*

PENDAHULUAN

Institusi keuangan memerlukan teknik pemodelan yang andal untuk mengetahui karakter nasabah dalam kemampuan bayar sehingga mengetahui apakah fasilitas pinjaman yang diberikan akan *default* atau tidak. Teknik *credit scoring* membantu dalam proses klasifikasi berdasarkan karakteristik masing-masing dan diterjemahkan ke dalam

bentuk skor atau nilai. Model *credit scoring* yang baik mampu dalam membedakan 'penerima pinjaman yang baik' dan 'penerima pinjaman yang buruk'. Biasanya model ini menggunakan data demografi dan data historis transaksi dengan data suplemen sehingga memungkinkan model menghasilkan informasi dengan tingkat akurasi yang baik.

Berdasarkan data OJK per Oktober 2022, kontribusi pembiayaan terbesar berasal dari industri perbankan untuk *fintech lending* maupun *multi finance* berasal dari pendanaan yang diterima dari bank dalam negeri. Lebih lanjut, pada studi kasus bank ABC, bank tidak memiliki kontrol atas pemberian limit kepada pengguna *fintech lending* maupun *multi finance* sehingga penelitian ini penting dilakukan karena adanya risiko kredit yang harus diantisipasi oleh bank akibat pemberian pinjaman dari bank kepada *fintech lending* maupun *multi finance*. OJK juga membuat aturan terkait mitigasi risiko gagal bayar akibat pemberian kredit ritel yang mempengaruhi industri keuangan khususnya perbankan dan *fintech lending* atau *multi finance* dengan pendekatan masing - masing.

Berdasarkan POJK 35/POJK.05/2018 tentang Penyelenggaraan Usaha Perusahaan Pembiayaan, terdapat skema pembiayaan yang dilakukan dalam proses pembiayaan antara bank dan perusahaan pembiayaan yaitu *joint financing* (pembiayaan bersama) dan *channeling* (pembiayaan penerusan). Risiko sepenuhnya berada pada pemilik dana yaitu bank pada skema pembiayaan *channeling*, sementara *joint financing* - pertanggungjawaban risiko yang timbul dari pembiayaan sesuai dengan proporsi dana yang dikeluarkan. Dalam melakukan kerjasama melalui *joint financing* atau *channeling*, perusahaan pembiayaan wajib memiliki sistem informasi dan teknologi yang memadai untuk memastikan kesesuaian data debitur yang dimiliki oleh perusahaan pembiayaan.

Persamaan dari kedua entitas penerima pembiayaan baik *fintech lending* atau perusahaan *multi finance* adalah kemampuan untuk menyalurkan pembiayaan kembali kepada Penerima Pinjaman berupa dana tunai tanpa agunan kepada segmen *underserved* atau *unbankable* sehingga produk pembiayaan diatur berdasarkan pilihan tenor dan pemberian limit tiap individu untuk pembelian barang dan dana tunai. Dari sisi perhitungan risiko kredit juga memiliki pendekatan yang berbeda dengan bank. Bank sebagai Pemberi Pinjaman atau pihak yang menyalurkan dana pembiayaan dengan skema *joint financing*, bank perlu memperhatikan dalam pengambilan keputusan investasi berdasarkan historis data pinjaman dalam rangka pengukuran risiko kredit sebagai kewajiban membentuk pencadangan modal untuk menutupi kerugian yang dibebankan dalam bentuk suku bunga apabila terjadi gagal bayar kepada Penerima Pinjaman.

Sementara pendekatan risiko kredit yang dilakukan *fintech lending* atau perusahaan *multifinance* menggunakan historis transaksi yang dilakukan pada platform atau aplikasi yang digunakan sebagai alat bayar non tunai untuk melihat kemampuan pembayaran dan karakter Penerima Pinjaman. Hal ini mendorong pentingnya menghitung besar kerugian gagal bayar yang mungkin terjadi terhadap dana yang diberikan oleh *fintech lending* maupun perusahaan *multi finance* sebagai penyelenggara. Informasi ini perlu terus diperhatikan sepanjang jangka waktu pinjaman berlangsung sehingga pengukuran risiko menggambarkan kondisi profil Penerima Pinjaman dengan tingkat akurasi yang dapat dipertanggungjawabkan.

Dikatakan *default* apabila Penerima Pinjaman tidak mampu memenuhi kewajiban untuk membayar sesuai dengan ketentuan yang berlaku sehingga *probability of default* (PD) adalah peluang terjadinya Penerima Pinjaman gagal memenuhi kewajiban sesuai dengan waktu yang ditentukan. Penelitian terkait PD terhadap portofolio *joint financing* pada *fintech lending* maupun perusahaan pembiayaan bersama dilakukan menggunakan pendekatan teknik tradisional regresi logistik dikarenakan baik bank dan *fintech lending* atau *multi finance* sejenis memiliki pendekatan masing – masing dalam menghitung PD untuk produk pinjaman dana tunai tanpa agunan.

Untuk pemodelan klasik yang biasanya digunakan dalam mengukur risiko kredit gagal bayar adalah regresi logistik. Dengan menggunakan teknik tradisional diharapkan mampu memodelkan PD dengan akurasi yang layak terhadap perilaku Penerima Pinjaman melalui *fintech lending* maupun *multifinance* yang dibiayai oleh bank sesuai dengan

proporsi pembiayaan masing – masing bila dibandingkan dengan memodelkan teknik lain seperti *machine learning*. Selain teknik regresi logistik, teknik *machine learning* digunakan penerapannya pada pemodelan risiko kredit dengan berbagai pendekatan. Hal ini mendorong perbandingan keandalan model sehingga dapat ditentukan teknik pemodelan yang sesuai dengan karakteristik data.

Penelitian ini bertujuan untuk memodelkan *probability of default* pada portofolio bank yang memiliki skema *joint financing* (pembiayaan Bersama) dengan *fintech lending* dan *multi finance* di salah satu bank di Indonesia menggunakan kombinasi variabel pada aplikasi dan transaksi antara teknik tradisional regresi logistik dan *machine learning*.

KAJIAN LITERATUR

Industri Fintech Lending & Multi Finance

Layanan Pendanaan Bersama Berbasis Teknologi (LPBBTI) atau yang umumnya dikenal dengan istilah *fintech lending* atau *peer-to-peer lending*, merupakan usaha yang menyediakan wadah bagi pemberi dana bertemu penerima dana untuk memberikan pendanaan konvensional atau berdasarkan prinsip syariah secara langsung melalui sistem elektronik dengan menggunakan internet. Dalam layanan ini, Pemberi Dana atau Penerima Dana dapat berupa perorangan, Lembaga yang berbadan hukum, dan/atau badan usaha yang memberikan pendanaan. Untuk memberikan layanan ini kepada masyarakat, LPBBTI harus memiliki izin usaha dari Otoritas Jasa Keuangan.

Hadirnya *fintech lending* dan *multi finance* membuka kesempatan bagi segment *underbank* atau *unbanked* di Indonesia untuk memperoleh akses kredit instan sehingga memudahkan proses pembayaran melalui layanan digital. Hal ini menciptakan wadah inovasi untuk mempertemukan Pemberi Pinjaman dan Penerima Pinjaman dengan *fintech lending* dan *multi finance* sebagai Perantara di antaranya. Menurut POJK 35/POJK.05/2018 tentang Penyelenggaraan Usaha Perusahaan Pembiayaan, terdapat dua skema pembiayaan yang diatur oleh OJK yaitu *joint financing* dan *channeling*.

Beberapa contoh *fintech lending* yang terdaftar di OJK antara lain adalah Danamas (PT Pasar Dana Pinjaman), investree (PT Investree Radhika Jaya), amarnya (PT Amarnya Mikro Fintek), Kredit Pintar (PT Kredit Pintar Indonesia), Lentera Dana Nusantara (PT Lentera Dana Nusantara), KrediFazz (PT FinAccel Digital Indonesia). Perusahaan menawarkan limit fasilitas *Buy Now Pay Later* (BNPL) kepada nasabah untuk digunakan pada pembayaran transaksi e-commerce atau transaksi digital lainnya (Bisnis.com, 2022).

Risiko Kredit

Risiko kredit diartikan sebagai potensi yang ditimbulkan penerima pinjaman bank atau rekanan akan gagal memenuhi tanggungjawab berdasarkan ketentuan yang disepakati. Tujuan dari manajemen risiko kredit adalah untuk memaksimalkan tingkat pengembalian yang disesuaikan dengan risiko dengan melakukan peninjauan eksposur risiko kredit berada dalam parameter yang dapat diterima. Secara umum, pinjaman merupakan kontribusi terbesar yang mendatangkan risiko kredit di dalam industri perbankan, sehingga bank diharapkan memiliki kewaspadaan untuk mengidentifikasi, mengukur, memonitor hingga mengawasi risiko kredit maupun menentukan apakah bank memiliki kecukupan modal terhadap risiko tersebut. Salah satu prinsip dari proses monitoring risiko adalah bank memiliki kemampuan untuk mengukur risiko kredit yang terjadi berdasarkan informasi yang memadai termasuk komposisi portofolio kredit yang bisa memicu adanya konsentrasi risiko (BCBS, Bank for International Settlements, 2000).

Komite Basel mengatur parameter yang dapat digunakan untuk mengukur eksposur risiko kredit antara lain adalah

- Probability of Default* (PD): persentase rata – rata penerima pinjaman yang gagal bayar dalam bentuk rating dalam satu tahun.
- Loss Given Default* (LGD): estimasi dari jumlah pinjaman yang merupakan dampak gagal bayar dari penerima pinjaman.
- Exposure at Default* (EAD): persentase eksposur yang mungkin dialami bank jika penerima pinjaman gagal bayar.

$$EL = EAD * PD * LGD$$

dengan menggunakan PD, LGD, dan EAD dapat dihitung *expected loss* bank dan dapat diketahui kebutuhan modal yang diperlukan oleh bank dalam menutup risiko kredit (BCBS, Bank for International Settlement, 2010). Dalam menghitung komponen risiko seperti *Probability of Default*, bank diharapkan dapat melakukan estimasi dengan memodelkan berdasarkan dua pendekatan yaitu IRB (*Internal Rating Based*) atau SA (*Standardized Approach*) sehingga diketahui besar kecukupan modal yang dibutuhkan (BIS, 2006).

Risiko Kredit pada Fintech Lending/Multi Finance

Menurut Tang (2019), Pemberi Pinjaman kepada *fintech lending* bersifat substitusi terhadap peran bank dalam memberikan pinjaman. Dengan hadirnya teknologi, kemampuan operasional *fintech lending* dapat dipenuhi sehingga mampu berfokus pada pinjaman yang berukuran lebih kecil daripada pinjaman yang ditawarkan oleh bank. Selain itu, pinjaman pada *fintech lending* lebih berisiko dibandingkan dengan pinjaman yang diberikan oleh bank karena kecenderungan memberikan pinjam kepada Penerima Pinjaman yang dikategorikan *underserved segment*.

Pemberi Pinjaman pada *fintech lending* cenderung memberi pinjaman kepada individu yang *less creditworthy*, tetapi seiring waktu kualitas dari portofolio Penerima Pinjaman akan mengalami perbaikan dan syarat yang diberikan oleh Pemberi Pinjaman akan berbeda. Secara spesifik, besarnya pinjaman yang ditawarkan akan berbeda dan dikenakan suku bunga yang lebih tinggi. Pinjaman *fintech lending* memiliki kecenderungan untuk gagal bayar. Lebih lanjut, kurangnya informasi yang diterima mengenai Penerima Pinjaman akan mengakibatkan Pemberi Pinjaman memberikan pinjaman dengan historis penolakan dari perusahaan pembiayaan tradisional (Maggio & Yao, 2020). Oleh karena itu, bank sebagai Pemberi Pinjaman perlu menghitung risiko kredit yang terjadi akibat pinjaman yang diberikan melalui *fintech lending* atau *multi finance*.

Regresi Logistik

Penelitian terbaru menunjukkan bahwa dengan meningkatkan kemampuan memodelkan skor kredit dapat mengurangi kerugian secara signifikan. Metode tradisional yang telah banyak digunakan untuk memprediksi skor kredit adalah regresi logistik atau linier. Dalam penelitian sebelumnya, ditemukan bahwa umur, tingkat pendidikan, alamat, dan status perkawinan dapat dijadikan fitur secara perilaku kredit untuk aktivitas kredit yang dimiliki. Untuk mengetahui apakah model statistik tradisional yang digunakan dapat memprediksi dengan baik, dilakukan uji hipotesis (Machado & Karray, 2022).

Machine Learning

Machine learning berdasarkan tipe data dapat dikategorikan menjadi lima bagian antara lain adalah *supervised learning*, *unsupervised learning*, *semi-supervised learning*, *reinforcement learning* dan *deep learning*. Pendekatan dengan tujuan memberi label berdasarkan informasi input yang sudah diformulasikan sebelumnya adalah termasuk tipe pendekatan *supervised learning* sehingga tujuan algoritma ini adalah memodelkan fungsi klasifikasi. Salah satu penerapan *supervised learning* adalah mengenali gambar dan huruf. Beberapa algoritma yang biasa digunakan dalam *supervised learning* adalah *Decision Tree* (DT), *Naïve Baiyes Classifier*, *Logistic Regression*, *K-Nearest Neighbour* (KNN), *Artificial Neural Network* dan *Support Vector Machines* (Dulhare, 2020).

Model Risiko Menggunakan Supervised Learning

Menurut Khandani et al. (2010), ketika menilai kemungkinan Penerima Pinjaman gagal bayar, hal ini dinilai sebagai masalah klasifikasi biner. Algoritma *machine learning* dapat digunakan untuk menentukan analisis risiko kredit daripada menggunakan model logistik tradisional karena *machine learning* dapat mengolah ukuran sampel yang lebih besar dan dengan tingkat kompleksitas yang lebih tinggi di dalam data transaksi dan karakteristik. Dengan menggunakan *machine learning*, bank sebagai Pemberi Pinjaman dapat menilai risiko gagal bayar berdasarkan kredit yang diberikan oleh *fintech lending*. Hal ini juga akan menjadi nilai tambah bagi bank sebagai Pemberi Pinjaman untuk dapat menentukan fungsi profit sehingga memudahkan dalam evaluasi performa model untuk melakukan prediksi (Liu et al., 2022).

Lebih lanjut mengenai *machine learning* khususnya *supervised learning* dalam memodelkan PD, teknik ini menjadi populer digunakan karena implementasi mudah dan hal penting yang perlu diperhatikan terkait asumsi yang digunakan dalam memodelkan. Salah satu teknik *supervised learning* adalah *decision tree*, teknik ini digunakan untuk membentuk node baru terus menerus hingga kriteria terpenuhi untuk berhenti yaitu kelompok – kelompok kecil yang bersifat homogen. Teknik ini dapat digunakan pada variable kontinu maupun diskrit atau kategorik.

Penelitian Terkait Sebelumnya

Penilaian kredit merupakan teknik yang digunakan secara umum untuk menolong bank dalam menentukan apakah calon debitur layak untuk diberikan pinjaman. Beberapa penelitian sebelumnya yang telah biasa digunakan untuk melakukan klasifikasi default: teknik tradisional regresi logistik (LR). Teknik machine learning yang umum telah digunakan untuk melakukan penilaian risiko sebelumnya antara lain adalah Support Vector Machine (SVM), Genetic Programming (GP), Artificial Neural Network (ANN) dan Decision Tree.

Penggunaan data mining untuk meningkatkan penilaian kelayakan kredit dengan menggunakan perbandingan antara model *credit scoring*, model regresi logistik dan model decision tree . Regresi logistik memiliki tingkat sensitivitas yang tertinggi dan *decision tree* yang terendah. Sementara tingkat kesalahan model untuk melakukan klasifikasi antara *credit scorecard*, regresi logistik dan *decision tree* adalah 27,9%, 28,8% dan 28,1%, masing-masing (Yap, et al, 2011). Teknik Random Forest (RF) unggul dibandingkan model regresi logistik standar dan bahkan menggunakan regresi logistik yang disesuaikan, di mana variabel dikodekan ulang dan dipilih menggunakan pengetahuan ahli tentang penilaian kredit (Kruppa, et.al, 2013). Besaran nilai pinjaman, tingkat pendidikan, dan tingkat penghasilan mempengaruhi dalam pinjaman kecil hingga menengah.

Dengan menggunakan in-sample dan out-sample ditunjukkan bahwa semakin lengkap informasi yang berkualitas dimiliki dalam memodelkan maka tingkat akurasi semakin tinggi dan membuktikan bahwa risiko kredit dari pinjaman p2p lending dapat diprediksi (Liu et al., 2018). Teknik pemodelan menggunakan decision tree-based heterogeneous ensemble menggunakan gabungan BGD, XGBoost dan Light GBM sebagai classifier menghasilkan model optimal dalam memprediksi probability of default nasabah P2P Lending (Zho, et. Al, 2019).

Risiko gagal bayar meningkat terkait variabel spread pinjaman, jangka waktu pinjaman dan usia pelanggan, tetapi menurun jika pelanggan memiliki lebih banyak kartu kredit. Nasabah yang menerima gaji dan pinjaman di lembaga perbankan yang sama memiliki lebih sedikit peluang gagal bayar daripada nasabah yang menerima gaji mereka di institusi lain. Teknik yang digunakan dalam penelitian ini adalah regresi logistik (Costa E Silva et al., 2020).

METODE PENELITIAN

Objek pada penelitian ini adalah data aplikasi dan data transaksi dari Penerima Pinjaman yang mendapatkan pembiayaan yang berasal dari portofolio salah satu bank di Indonesia. Sejak April 2019, bank ABC sudah bekerja sama dengan *fintech lending* dan perusahaan pembiayaan *multi finance* dengan kriteria yang telah ditetapkan sesuai dengan *risk appetite* bank dan ketentuan OJK atau BI yang berlaku untuk produk pembiayaan. Dikarenakan kriteria yang diterapkan oleh bank sebagai Pemberi Pinjaman telah disesuaikan dengan *risk appetite* bank ABC sehingga tidak seluruh informasi dan transaksi yang dilaksanakan *fintech lending* atau *multi finance* dimasukkan ke dalam portofolio bank. Selanjutnya *fintech lending* dan perusahaan *multi finance* akan disebut dengan Penyelenggara dan nasabah pada Penyelenggara akan disebut Penerima Pinjaman serta dalam hal ini Bank ABC disebut sebagai Pemberi Pinjaman.

Untuk menghindari pelanggaran kerahasiaan data konsumen, data hanya disajikan setelah tidak mengandung informasi sensitif yang bersifat rahasia konsumen sesuai aturan regulator sehingga informasi yang bersifat data public dan tetap menjamin struktur data tidak berubah sehingga mencegah pelanggaran informasi yang sensitif milik bank ABC.

Aplikasi memiliki *average ticket size* berkisar hingga 1 juta rupiah baik transaksi *fintech lending* maupun perusahaan *multifinance* dengan bentuk produk berupa dana tunai tanpa agunan atau pembiayaan transaksi *commerce* baik *online* atau *offline*.

Periode pengamatan yang akan dipakai dalam penelitian adalah aplikasi Penerima Pinjaman pada *fintech lending* atau *multi finance* yang memiliki satu fasilitas pembiayaan bersama dengan bank ABC maksimum sebanyak satu kali hingga Desember 2022. Status aplikasi tersebut diamati hingga menjadi *default* atau periode pengamatan berakhir. Terdapat variabel independen yang berasal dari data aplikasi merupakan data demografi yang dimiliki oleh Penerima Pinjaman pada saat melakukan pengajuan aplikasi melalui perusahaan pembiayaan maupun *fintech lending*. Data aplikasi yang digunakan sebagai variabel bebas dalam penelitian ini adalah:

- Usia, merupakan umur Penerima Pinjaman saat mengajukan transaksi.
- Pendapatan, merupakan deklarasi penghasilan Penerima Pinjaman per bulan yang telah disetahunkan.
- Pekerjaan, merupakan tipe pekerjaan Penerima Pinjaman dan berdasarkan informasi yang dideklarasikan oleh Penerima Pinjaman.
- Lama bekerja, merupakan jangka waktu Penerima Pinjaman bekerja di posisi pekerjaan yang sekarang.
- Kota Domisili, merupakan kota dimana rumah Penerima Pinjaman tinggal berdasarkan informasi yang dideklarasikan.
- Status Pendidikan, merupakan pendidikan terakhir yang ditempuh Penerima Pinjaman berdasarkan informasi yang dideklarasikan
- Deskripsi Pekerjaan, merupakan deskripsi pekerjaan mengacu pada klasifikasi internal yang bank.
- Jumlah Pihak Tertanggung, merupakan jumlah pihak yang ditanggung oleh Penerima Pinjaman saat aplikasi diajukan.
- Status Pernikahan, merupakan status pernikahan Penerima Pinjaman.
- Industri Bidang Pekerjaan, merupakan industri dimana Penerima Pinjaman diklasifikasikan secara internal oleh bank.

Variabel independen lainnya dibentuk dari data transaksi merupakan data transaksi yang dilakukan oleh Penerima Pinjaman pada saat melakukan pengajuan aplikasi melalui perusahaan pembiayaan maupun *fintech lending*. Data transaksi yang digunakan sebagai variabel bebas dalam penelitian ini adalah:

- Kredit Limit Partner, merupakan nominal plafon yang diberikan oleh *fintech lending* dan bersifat revolving seperti kartu kredit. Adapun Penerima Pinjaman pada perusahaan pembiayaan menggunakan layanan dalam bentuk pembiayaan barang sehingga limit yang diberikan yang diberikan tidak bersifat revolving dan akan disesuaikan dengan harga barang yang dibeli.
- Jenis aplikasi, merupakan tipe transaksi yang dilakukan Penerima Pinjaman dan dapat dikategorikan menjadi dua hal yaitu pinjaman tanpa agunan (pencairan limit untuk memperoleh dana tunai) dan pembiayaan barang (pencairan limit guna melunasi barang/produk).
- Tenor, merupakan jangka waktu pinjaman atau pembiayaan yang diberikan kepada Penerima Pinjaman.
- Nominal transaksi, merupakan nominal transaksi yang dicairkan kepada Penerima Pinjaman.
- Nominal suku bunga, merupakan nominal suku bunga yang dibebankan kepada Penerima Pinjaman untuk menanggung kemungkinan risiko kredit dari pembiayaan yang diberikan.
- Nominal cicilan, merupakan besaran cicilan yang dibebankan kepada Penerima Pinjaman untuk dibayarkan setiap bulan sepanjang tenor pinjaman berlaku.
- Status Pinjaman, merupakan variabel dependen yang menjadi acuan dikatakan *default* atau tidaknya suatu pinjaman.

- Tanggal Aplikasi, merupakan tanggal dimana aplikasi disetujui untuk diberikan pinjaman oleh bank melalui fintech lending atau multi finance.
- Tujuan Pinjaman, merupakan tujuan penggunaan dana pinjaman pada fintech lending atau multi finance sebagai media bayar atau aplikasi commerce.
- Jumlah Hari Tunggakan, merupakan jumlah hari sejak cicilan pinjaman tidak dibayarkan oleh Penerima Pinjaman pada tanggal jatuh tempo.

Untuk melakukan pengelompokkan dan memilih variabel secara otomatis dan interaktif digunakan *The Weight of Evidence* (WOE) dan pengukuran *Information Value* (IV). WOE memiliki definisi yaitu logaritma dari rasio dari proporsi 'baik' atas proporsi 'buruk' dalam atribut data. Apabila hasil logaritma bernilai negatif tinggi maka menandakan risiko tinggi, sebaliknya apabila bernilai positif tinggi bersesuaian dengan risiko rendah. Perhitungan WOE dapat ditulis sebagai berikut:

$$WOE_{attribute} = \log\left(\frac{p_{goodattribute}}{p_{badattribute}}\right)$$

$$\text{dimana } p_{goodattribute} = \frac{\#goodsattribute}{\#goods} \text{ dan } p_{badattribute} = \frac{\#badsattribute}{\#bads} \quad (1)$$

Setelah dilakukan perhitungan WOE, untuk penentuan variabel prediktif diperlukan pengukuran menggunakan IV dengan memberikan bobot pada variabel atribut. Bobot tersebut merupakan selisih antara proporsi 'baik' dan proporsi 'buruk' dalam atribut tertentu. Perhitungan IV dapat ditulis sebagai berikut:

$$Information\ Value = \sum p_{goodattribute} - p_{badattribute} * WeightofEvidence \quad (2)$$

Regresi logistik akan membentuk variabel ($\log(p/(1-p))$) yang merupakan hubungan linier antar variabel independen. Nilai variabel ini ditransformasikan menjadi probabilitas dengan fungsi logit. Model regresi logistik dapat ditulis dengan formula sebagai berikut :

$$\log\left[\frac{P(Y=1)}{1-P(Y=1)}\right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (3)$$

dimana $P(Y = 1)$ merupakan peluang terjadinya hal yang ingin diketahui.

Selanjutnya persamaan (1) dapat ditulis sebagai berikut :

$$P(Y = 1) = \frac{1}{1+e^{-z}} \text{ dimana } z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (4)$$

Dengan menggunakan regresi logistik dalam *credit scoring* adalah untuk menentukan peluang dari nasabah menjadi *default* atau *non-default* dimana ($Y = 1$) merupakan tujuan pemodelan (*default*).

Pada penelitian ini, *probability of default* (PD) untuk setiap aplikasi dimodelkan sebagai berikut :

$$\log\left(\frac{1-P(Y=1)}{1-P(Y=1|x_i)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i, i = 1, 2, \dots, 16 \quad (5)$$

$$\Leftrightarrow \log\left(\frac{1-P(Y=1)}{1-P(Y=1|x_i)}\right) = e^{-(\alpha + \sum_{i=1}^{16} \beta_i x_i)} \quad (6)$$

dimana x_1 adalah usia, x_2 adalah pendapatan, x_3 adalah pekerjaan, x_4 adalah lama bekerja, x_5 adalah limit, x_6 adalah outstanding, x_7 adalah jumlah hari tunggakan, x_8 adalah jenis transaksi, x_9 adalah nominal tunggakan, x_{10} adalah jumlah frekuensi transaksi, x_{11} adalah nominal transaksi, x_{12} adalah tingkat utilisasi limit, x_{13} adalah tenor, x_{14} adalah suku bunga.

Metode kedua adalah menggunakan *machine learning* yang tergolong dalam *supervised learning* yaitu *decision tree* (DT). Model ini bertujuan membagi sekumpulan data observasi menjadi kelompok data homogen yang lebih kecil dengan variabel target tertentu. Umumnya variable target bersifat kategorik sehingga model *decision tree* digunakan untuk menghitung probabilitas untuk setiap target kategori dan juga dapat diterapkan pada data numerik. Terdapat dua node di dalam model DT yaitu *Decision node* dan *Leaf node*. *Decision node* digunakan untuk membuat keputusan dan memiliki satu cabang atau lebih, sementara *Leaf node* merupakan output dari *decision node* dan tidak memiliki cabang. Diperoleh keputusan *node* diperoleh berdasarkan atribut dataset yang ada.

Aturan klasifikasi yang akan digunakan dalam *decision tree* menggunakan dua puluh (20) variabel yang telah disebutkan sebelumnya menggunakan algoritma Classification and Regression Tree algorithm (CART).

Dengan menggunakan algoritma CART, dipilih *tree* yang terbaik berdasarkan *cross-validation* dan *minimum error rule* dimulai dari *root node (default)*, kemudian mencari atribut terbaik dari variabel yang tersedia pada. Selanjutnya membagi atribut menjadi sub-set yang bersifat homogen dan dilakukan secara rekursif sampai node tidak memiliki *leaf node*. Algoritma C4.5 merupakan salah satu metode yang umumnya dipakai untuk membangun model DT. Algoritma memulai *tree* dan mengevaluasi atribut yang menghasilkan partisi bersifat informatif.

Algoritma C4.5 bertujuan menghasilkan *decision tree* dengan menghitung dan membandingkan Gain Ratio, kemudian pada node yang terbentuk di level berikutnya, dan hal yang sama dilakukan secara rekursif. Tahapan dalam membuat DT menggunakan algoritma C4.5 yaitu :

1. Mempersiapkan data training dari data historis Bank ABC
2. Menentukan *root node* dengan menghitung nilai gain tertinggi dari masing-masing atribut atau berdasarkan nilai indeks entropy terendah. Terlebih dahulu index entropy harus dihitung dengan rumus:

$$Entropy(S) = \sum_{i=1}^n -p_i \log_2 p_i$$

dimana

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi s_i terhadap S

3. Setelah itu dihitung nilai *gain* dengan Gain Information melalui persamaan berikut :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|s_i|}{|S|} * Entropy(s_i)$$

dimana

S = himpunan kasus

A = Atribut

n = jumlah partisi Atribut A

$|s_i|$ = proporsi s_i terhadap S

$|S|$ = Jumlah kasus dalam S

4. Lakukan perhitungan yang sama secara rekursif pada entropy dan gain di setiap atribut yang ada hingga seluruh atribut terpartisi atau tidak memiliki *leaf node* (Dash, 2022).

Untuk memastikan model regresi logistik yang dihasilkan telah menggambarkan data dengan baik dan variabel prediktor yang sesuai diperlukan uji keandalan. Salah satu uji yang dapat dilakukan adalah likelihood ratio test, dengan tujuan membandingkan likelihood data antara model seluruhnya terhadap likelihood menggunakan model dengan variabel prediktor lebih sedikit. Kedua model tersebut divalidasi menggunakan uji *sensitivity*, *precision*, dan *f1-score* (Yap, et al, 2011).

Confusion Matrix

	+	-
+	TP	FP
-	FN	TN

$$Sensitivity = \left(\frac{TP}{TP + FN} \right)$$

$$Precision = \left(\frac{TP}{TP + FP} \right)$$

$$f1 - score = \left(\frac{2}{\frac{1}{Precision} + \frac{1}{Sensitivity}} \right)$$

Selain itu, validasi model dalam rangka perbandingan antara kedua model dapat dilihat menggunakan kurva Receiver Operating Characteristic (ROC). Hal ini bertujuan untuk melihat hasil klasifikasi model menggunakan dua parameter yaitu True Positive Rate (Recall) dan False Positive Rate (Yap et al., 2011).

HASIL DAN PEMBAHASAN

Berdasarkan berbagai variabel yang digunakan dalam penelitian, model dibangun menggunakan data sebanyak 1.496.251 yang terdiri dari aplikasi yang dibuka pada periode April 2019 hingga Desember 2022. Data tersebut diamati secara bulanan hingga menjadi default atau hingga periode pengamatan berakhir. data merupakan hasil kombinasi variabel aplikasi dan variabel transaksi pada pembiayaan bersama bank ABC.

Statistik Deskriptif

Variabel	count	mean	std	min	25%	50%	75%	max
Pendapatan	1,496,252	66,520,090	123,869,100	0	30,000,000	50,400,000	72,000,000	28,164,000,000
Lama_Bekerja	1,496,252	7	7	0	2	5	10	99
Status_Pendidikan	1,496,252	4	2	1	1	4	5	7
Jumlah_Pihak_Tertanggung	1,496,252	2	5	0	0	0	1	99
Status_Pernikahan	1,496,252	2	1	1	1	1	3	3
Industri_Bidang_Pekerjaan	1,496,252	2	1	0	1	3	3	8
Kredit_Limit_Partner	1,496,252	6,465,806	5,027,605	225,000	2,555,100	4,884,300	9,321,840	30,000,000
Tenor	1,496,252	6	3	1	3	6	7	18
Nominal_Transaksi	1,496,252	2,299,000	2,090,321	200,007	679,023	1,800,000	3,144,600	27,000,000
Nominal_Suku_Bunga	1,496,252	237,880	291,231	0	33,510	158,332	321,198	6,198,849
Nominal_Cicilan	1,496,252	415,129	379,937	19,131	169,969	321,798	524,045	9,164,970
Jumlah_Hari_Tunggakan	1,496,252	12	41	0	0	0	0	804

Tabel 1 Statistik Deskriptif Data 2019 - 2022

Berdasarkan matriks korelasi terhadap seluruh variabel independen dapat dilihat terdapat beberapa variabel yang memiliki korelasi mendekati 1 dimana x1 adalah Pendapatan, x2 adalah Lama_Bekerja, x3 adalah Status_Pendidikan, x4 adalah Jumlah_Pihak_Tertanggung, x5 adalah Status_Pernikahan, x6 adalah Industri_Bidang_Pekerjaan, x7 adalah Kredit_Limit_Partner, x8 adalah Tenor, x9 adalah Nominal_Transaksi, x10 adalah Nominal_Suku_Bunga, x11 adalah Nominal_Cicilan, x12 adalah Jumlah_Hari_Tunggakan, x13 adalah Status_Pinjaman yaitu Nominal Transaksi, Nominal Suku Bunga, Nominal Cicilan.

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13
x1	1.00	0.10	-0.11	0.08	0.15	0.00	0.11	0.06	0.23	0.18	0.18	-0.02	0.02
x2	0.10	1.00	0.01	-0.05	0.16	-0.04	-0.13	0.00	-0.03	0.00	-0.03	-0.06	0.06
x3	-0.11	0.01	1.00	-0.06	-0.07	0.20	-0.11	-0.07	-0.17	-0.13	-0.13	0.04	-0.04
x4	0.08	-0.05	-0.06	1.00	0.41	-0.02	0.31	0.08	0.14	0.08	0.11	0.04	-0.05
x5	0.15	0.16	-0.07	0.41	1.00	-0.15	-0.08	0.46	0.33	0.36	0.11	0.05	-0.04
x6	0.00	-0.04	0.20	-0.02	-0.15	1.00	0.07	-0.17	-0.13	-0.14	-0.04	0.00	0.00
x7	0.11	-0.13	-0.11	0.31	-0.08	0.07	1.00	-0.15	0.26	0.12	0.36	-0.04	0.03
x8	0.06	0.00	-0.07	0.08	0.46	-0.17	-0.15	1.00	0.47	0.69	-0.04	0.19	-0.19
x9	0.23	-0.03	-0.17	0.14	0.33	-0.13	0.26	0.47	1.00	0.88	0.71	0.08	-0.07
x10	0.18	0.00	-0.13	0.08	0.36	-0.14	0.12	0.69	0.88	1.00	0.43	0.12	-0.12
x11	0.18	-0.03	-0.13	0.11	0.11	-0.04	0.36	-0.04	0.71	0.43	1.00	-0.01	0.02
x12	-0.02	-0.06	0.04	0.04	0.05	0.00	-0.04	0.19	0.08	0.12	-0.01	1.00	-0.98
x13	0.02	0.06	-0.04	-0.05	-0.04	0.00	0.03	-0.19	-0.07	-0.12	0.02	-0.98	1.00

Tabel 2 Matriks Korelasi Variabel Numerik Data 2019 - 2022

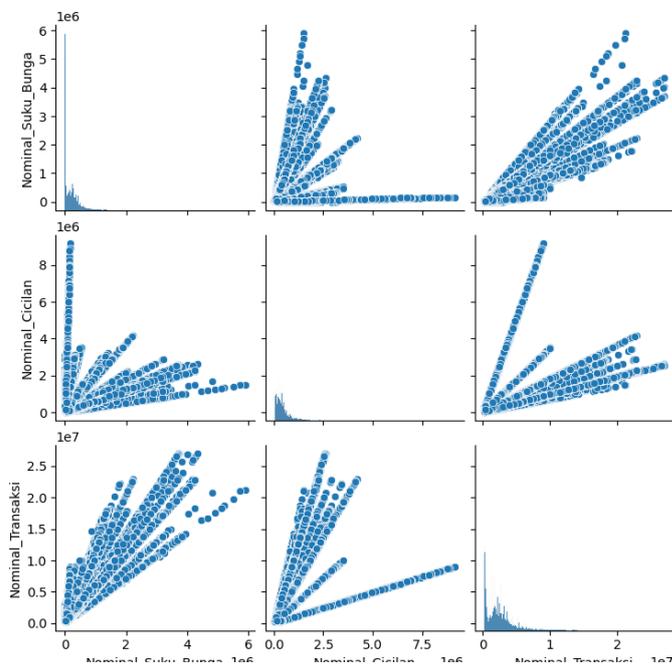
Perhitungan WOE dan IV

Dengan menggunakan formula VIF untuk memastikan kehadiran multikolinearitas diperoleh nilai VIF sebagai berikut:

Variabel	Nilai VIF
Nominal_Suku_Bunga	-1.52E-07
Nominal_Cicilan	-2.48E-08
Nominal_Transaksi	-4.42E-09

Tabel 3 Nilai VIF pada Tes Multikolinearitas

Berdasarkan nilai VIF tidak ditemukan multikolinearitas. Lebih lanjut dilakukan visual secara grafik untuk menggambarkan relasi antara ketiga variable dalam rangka memastikan hadirnya multikolinearitas. Gambar 1 menunjukkan variabel Nominal_Suku_Bunga memiliki hubungan linier dengan variabel Nominal_Transaksi sehingga salah satu variabel tidak



Gambar 1 Grafik Regresi Linier - Tes Lanjutan Multikolinearitas

diikutsertakan dalam memodelkan probability of default. Pada tahapan perhitungan WOE dan IV dilakukan untuk memilih variable signifikan. Sebelumnya dilakukan perubahan tipe data pada variabel Status Pinjaman yang sebelumnya merupakan variabel kategorik menjadi variabel biner (1 atau 0). Lebih lanjut,

Variabel	iv	iv_interpretation
Tujuan_Pinjaman	0.26	medium
Nominal_Suku_Bunga	0.23	medium
Deskripsi_Pekerjaan	0.15	medium
Status_Pendidikan	0.13	medium
Tenor	0.13	medium
Nominal_Transaksi	0.10	medium
Lama_Bekerja	0.05	weak
Kota_Domisili	0.03	weak
Jumlah_Pihak_Tertanggung	0.02	weak
Flag_Declared_Income	0.02	useless
Jenis_Aplikasi	0.02	useless
Pekerjaan	0.02	useless
Flag_Age	0.01	useless
Kredit_Limit_Partner	0.01	useless
Pendapatan	0.00	useless
Nominal_Cicilan	0.00	useless
Industri_Bidang_Pekerjaan	0.00	useless
Jumlah_Hari_Tunggakan	0.00	useless
Status_Pinjaman	0.00	useless
Status_Pernikahan	0.00	useless

Tabel 4 Nilai IV dan Interpretasi Nilai IV Periode 2019-2022

Variabel dengan nilai IV di antara 0.1 hingga 0.5 menjadikan variabel berikut sebagai atribut yang dapat digunakan untuk membangun model antara lain adalah Tenor, Tujuan Pinjaman, Nominal Suku Bunga, Deskripsi Pekerjaan, Status Pendidikan, Lama Bekerja, Kota Domisili, dan Jumlah Pihak Tertanggung.

Data In-Sample dan Out-Sample

Pada tahapan ini, menggunakan variabel yang telah dipilih sebelumnya, dilakukan pemisahan data menjadi in-sample (*data training*) dan out-sample (*data testing*) dengan komposisi masing – masing sebesar 70% dan 30%. Ukuran *data training* (*y train*) dan *data testing* (*y test*) dengan komposisi *default* masing – masing 74.152 nasabah dan 31.990 nasabah. Untuk melakukan pengecekan keandalan model terhadap data *in-sample* dan *out-sample* maka dilakukan perhitungan menggunakan *confusion matrix* adalah sebagai berikut:

		Predicted	
		Default	Paid
Actual	Default	170	31.820
	Paid	548	284.567

Tabel 5 Confusion Matrix - Regresi Logistik Periode 2019 – 2022

Koefisien untuk setiap variabel dapat dilihat pada Tabel 6.

Variabel	Koefisien
Deskripsi_Pekerjaan_n	0.074488
Lama_Bekerja	0.043653
Status_Pendidikan	0.012157
Tenor	0.010633
Jumlah_Pihak_Tertanggung	0.003199
Kota_Domisili_n	0.001008
Tujuan_Pinjaman_n	0.000243
Nominal_Suku_Bunga	-0.000001

Tabel 6 Koefisien Variabel pada Model Regresi Logistik (Data Training)

Variabel	Koefisien
Kota_Domisili_n	0.0031089
Tujuan_Pinjaman_n	0.0009304
Deskripsi_Pekerjaan_n	0.0001663
Lama_Bekerja	0.0000895
Tenor	0.0000331
Status_Pendidikan	0.0000316
Jumlah_Pihak_Tertanggung	0.0000115
Nominal_Suku_Bunga	0.0000003

Tabel 7 Koefisien Variabel pada Model Regresi Logistik (Data Testing)

Selanjutnya dilakukan perhitungan menggunakan indikator *sensitivity*, *precision* dan *f-1 score* dalam menentukan *default* dapat dilihat pada Tabel 7. Hal ini menunjukkan bahwa dengan akurasi 89.8% dalam memodelkan pembiayaan bersama secara umum tetapi memiliki tingkat *precision* dan *recall* yang rendah pada data training.

Sensitivity	1.00%
Precision	24.0%
F1-score	1.00%

Tabel 8 Nilai Sensitivity, Precision, F1-Score – Default – Regresi Logistik Periode 2019 – 2022 (Data Training)

Selanjutnya menggunakan model yang telah dibangun pada data training kemudian diterapkan pada data *testing* dan diperoleh hasil yang sama. Nilai indikator *sensitivity*, *precision* dan *f-1 score* dalam menentukan *default* dapat dilihat pada Tabel 7.

Sensitivity	1.00%
Precision	24.0%
F1-score	1.00%

Tabel 9 Nilai Sensitivity, Precision, F1-Score – Default – Regresi Logistik Periode 2019 – 2022 (Data Testing)

Selanjutnya, analisis model prediksi *default* yang digunakan untuk membangun model *probability of default* berdasarkan variabel aplikasi dan transaksi yaitu *decision tree* CART. *Confusion matrix* pada model *decision tree* dapat dilihat pada Tabel 9.

	Predicted		
	Default	Paid	
Actual	Default	6.050	25.940
	Paid	36.733	284.75

Tabel 10 Confusion Matrix – Decision Tree Periode 2019 – 2022

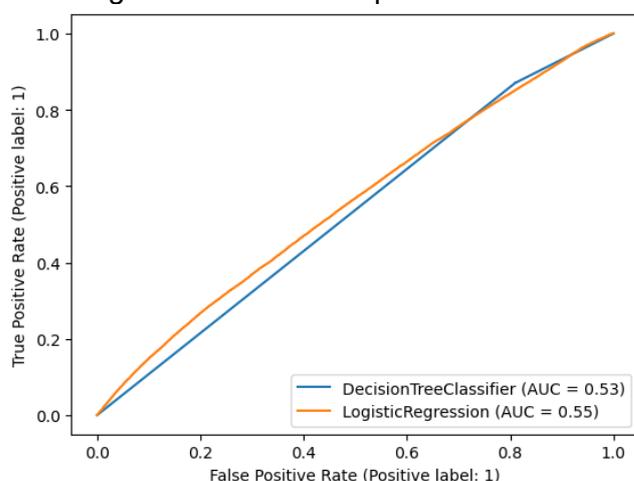
Selanjutnya menggunakan model yang telah dibangun pada data training kemudian diterapkan pada data *testing* dan diperoleh hasil yang sama. Berdasarkan nilai indikator *sensitivity*, *precision* dan *f-1 score* dalam menentukan *default* dapat dilihat pada Tabel 10, hasil akurasi model *decision tree* memiliki nilai *precision* lebih rendah dibandingkan dengan regresi logistik meskipun nilai *sensitivity* dan *f1-score* lebih tinggi.

Sensitivity	14.0%
Precision	0.19%
F1-score	16.0%

Tabel 11 Nilai Sensitivity, Precision, F1-Score – Default – Decision Tree Periode 2019 – 2022 (Data Training)

Nilai akurasi pada mode *decision tree* sebesar 80.2%, maka dapat disimpulkan akurasi model *probability of default* menggunakan *decision tree* dapat memodelkan data sebesar 80.2% dengan benar.

Selanjutnya, digunakan kurva ROC dengan dua indikator yaitu *True Positive* dan *True Negative* untuk menampilkan keandalan antara kedua model tersebut (Gambar).



Tabel 12 Kurva ROC Regresi Logistik vs Decision Tree Periode 2019 - 2022

Hal ini menunjukkan bahwa dengan menggunakan teknik tradisional regresi logistik dapat memodelkan lebih baik menggunakan variabel Tenor, Tujuan Pinjaman, Nominal Suku Bunga, Deskripsi Pekerjaan, Status Pendidikan, Lama Bekerja, Kota Domisili, Flag Age dan Kredit Limit Partner pada periode data 2019 hingga 2022 untuk aplikasi pembiayaan bersama antara bank dan *fintech lending/multi-finances*.

KESIMPULAN DAN SARAN

Model regresi logistik dapat digunakan untuk mengidentifikasi karakteristik dari aplikasi pembiayaan bersama pada *fintech lending/multi finance*. Penelitian dilakukan menggunakan data aplikasi berisi informasi yang dideklarasikan oleh Penerima Pinjaman berikut transaksi yang mendapatkan pembiayaan bersama untuk dapat menunjukkan pola transaksi yang berisiko bagi bank jika menjadi *default*. Identifikasi karakteristik pemodelan

risiko dilakukan untuk mengantisipasi peningkatan risiko pencadangan dan NPL yang signifikan seiring dengan pertumbuhan *fintech lending/multi finance* di Indonesia.

Penelitian ini menerapkan metode regresi logistik dalam memprediksi Penerima Pinjaman yang mendapat pembiayaan bersama antara bank dan *fintech lending/multi finance* dan dibandingkan hasil pemodelan tersebut dengan metode *machine learning* yaitu *decision tree* dalam memprediksi besarnya peluang *default* pada salah satu bank dengan portofolio pembiayaan bersama di Indonesia.

Berdasarkan analisis dan pembahasan pada penelitian ini, dapat disimpulkan beberapa hal sebagai berikut:

1. Variabel aplikasi yang signifikan memengaruhi *probability of default* pada pembiayaan bersama antara bank dan *fintech lending/multi finance* adalah Deskripsi Pekerjaan, Status Pendidikan, Lama Bekerja, Kota Domisili dan Jumlah Pihak Tertanggung.
2. Sementara Variabel aplikasi yang signifikan memengaruhi *probability of default* pada pembiayaan bersama antara bank dan *fintech lending/multi finance* adalah durasi pinjaman (tenor), tujuan pinjaman, dan nominal suku bunga yang dibebankan kepada Penerima Pinjaman.
3. Hasil validasi dengan metode in-sample dan out-sample pada periode 2019 hingga 2022 menunjukkan bahwa model yang dibangun tidak berlaku secara umum atau tidak mampu memprediksi *probability of default* pada waktu pengamatan yang berbeda menggunakan teknik tradisional regresi logistik. Hal tersebut dipengaruhi oleh adanya pergeseran jumlah observasi untuk beberapa variabel pada data yang digunakan untuk pembangunan model dan data yang digunakan untuk validasi model.

Hasil validasi dengan metode in sample dan out sample pada periode 2019 hingga 2022 menunjukkan bahwa model yang dibangun tidak berlaku secara umum atau tidak mampu memprediksi *probability of default* pada waktu pengamatan yang berbeda menggunakan *machine learning decision tree*. Hal tersebut dipengaruhi oleh adanya pergeseran jumlah observasi untuk beberapa variabel pada data yang digunakan untuk pembangunan model dan data yang digunakan untuk validasi model.

Berdasarkan hasil, analisis dan kesimpulan penelitian tersebut, maka dapat disarankan beberapa hal sebagai berikut:

1. Bagi bank
 - a. Database yang berisi informasi terkait data aplikasi serta data transaksi debitur dapat diolah untuk membantu proses pengambilan keputusan atau membuat peraturan bahkan pertimbangan strategi bisnis bagi manajemen bank. Oleh karena itu, kualitas data sangat penting untuk ditingkatkan dengan cara melakukan verifikasi data calon demografi Penerima Pinjaman pada saat proses onboarding sebelum menerima pembiayaan bersama untuk menghindari terjadinya kesalahan data. Peningkatan kualitas data transaksi dapat dilakukan dengan cara melakukan pengecekan setiap variabel demografi mandatori untuk pembuatan model secara reguler sehingga proses pengolahan data dapat lebih cepat dan menghasilkan hasil yang lebih optimal.
 - b. Teknik regresi logistik dapat digunakan untuk melakukan profiling karakteristik Penerima Pinjaman dari pembiayaan bersama antara bank dan *fintech lending/multifinance*. Profiling tersebut dapat digunakan untuk mencegah risiko yang dapat terjadi di waktu mendatang serta dapat diteliti lebih lanjut untuk peningkatan portfolio melalui cross-selling dengan produk bank yang lain
3. Bagi akademisi dan penelitian berikutnya:
 - a. Pembangunan model risiko kredit pembiayaan bersama antara bank dan *fintech lending/multi finance* dapat diperluas dengan menggunakan data dari beberapa bank di Indonesia sehingga hasil penelitian dapat berlaku secara umum.
 - b. Pembangunan model risiko kredit pembiayaan bersama antara bank dan *fintech lending/multi finance* dapat dilakukan dengan menambahkan berbagai variabel aplikasi agar pola transaksi Penerima Pinjaman dapat diketahui dengan lebih seksama.

c. Baik teknik regresi logistik dan decision tree dapat digunakan untuk memodelkan risiko kredit pembiayaan bersama antara bank dan fintech lending/multi finance yang selanjutnya dapat digunakan untuk mengestimasi LGD.

DAFTAR PUSTAKA

- APPI. (2021). Retrieved from https://www.appi.id/download/statistic/Buku_Statistik_Lembaga_Pembiayaan_2021.pdf
- BCBS. (2000, September). Bank for International Settlements. Retrieved from Basel Committee Publication: <https://www.bis.org/publ/bcbs75.pdf>
- BCBS. (2010, September). Bank for International Settlement. Retrieved from <https://www.bis.org/fsi/awp2010.pdf>
- Bl. (2012). Retrieved from <https://www.ojk.go.id/Files/box/keuangan-berkelanjutan/pbi-nomor-14-15-pbi-2012.pdf>
- BIS. (2006). International Convergence of Capital Measurement and Capital Standards . Retrieved from <https://www.bis.org/publ/bcbs128.pdf>
- Bisnis.com. (2022, November 24). Ini Rencana OJK Setelah Kredit Macet Paylater Mendekati 8 Persen. Retrieved from <https://bisnis.tempo.co/read/1660763/ini-rencana-ojk-setelah-kredit-macet-paylater-mendekati-8-persen>
- Costa E Silva, E., Lopes, I. C., Correia, A., & Faria, S. (2020). A logistic regression model for consumer default risk. *Journal of Applied Statistics*, 47(13–15), 2879–2894. <https://doi.org/10.1080/02664763.2020.1759030>
- Dash, S. (2022, November 02). Decision Trees Explained — Entropy, Information Gain, Gini Index, CCP Pruning. Retrieved from Medium: <https://towardsdatascience.com/decision-trees-explained-entropy-information-gain-gini-index-ccp-pruning-4d78070db36c>
- Datacamp. (2022). Retrieved from <https://app.datacamp.com/learn/courses/ensemble-methods-in-python>
- Dulhare, U. N. (2020). *Machine learning and big data : Concepts, algorithms, tools and applications*. John Wiley & Sons, Incorporated.
- EY. (2017). Retrieved from https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/banking-and-capital-markets/ey-unleashing-the-potential-of-fintech-in-banking.pdf
- GitHub. (2020). Retrieved from <https://github.com/klaudia-nazarko/iv-and-woe-python>
- 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>
- Kyriakides, G., & Margaritis, K. G. (2019). *Hands-on ensemble learning with python : Build highly optimized ensemble machine learning models using scikit-learn and keras*. Packt Publishing, Limited.
- Liu, Y., Yang, M., Wang, Y., Li, Y., Xiong, T., & Li, A. (2022). Applying machine learning algorithms to predict default probability in the online credit market: Evidence from China. *International Review of Financial Analysis*, 79, 101971. <https://doi.org/10.1016/j.irfa.2021.101971>
- Liu, Y., Zhou, Q., Zhao, X., & Wang, Y. (2018). Can Listing Information Indicate Borrower Credit Risk in Online Peer-to-Peer Lending? *Emerging Markets Finance and Trade*, 54(13), 2982–2994. <https://doi.org/10.1080/1540496X.2018.1427061>
- Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Systems with Applications*, 200, 116889. <https://doi.org/10.1016/j.eswa.2022.116889>
- Maggio, M. D., & Yao, V. (2020). Fintech Borrowers: Lax Screening or Cream-Skimming? *Review of Financial Studies*, 4565–4618.
- Mathew. (2015). Retrieved from <https://www.r-bloggers.com/2015/08/evaluating-logistic-regression-models/>

- OJK. (2016). Retrieved from <https://www.ojk.go.id/id/regulasi/Documents/Pages/POJK-tentang-Layanan-Pinjam-Meminjam-Uang-Berbasis-Teknologi-Informasi/SAL%20-%20PENJELASAN%20POJK%20Fintech.pdf>
- OJK. (2016). Retrieved from <https://www.ojk.go.id/id/regulasi/otoritas-jasa-keuangan/peraturan-ojk/Documents/Pages/POJK-Nomor-77-POJK.01-2016/SAL%20-%20POJK%20Fintech.pdf>
- OJK. (2016). Retrieved from <https://www.ojk.go.id/id/regulasi/Pages/POJK-tentang-Penerapan-Manajemen-Resiko-bagi-Bank-Umum.aspx>
- OJK. (2017). Retrieved from <https://www.ojk.go.id/id/kanal/iknb/pages/lembaga-pembiayaan.aspx>
- OJK. (2017). Retrieved from <https://www.ojk.go.id/id/kanal/perbankan/pages/Bank-Umum.aspx>
- OJK. (2018). Penyelenggaraan Usaha Perusahaan Pembiayaan. Retrieved from <https://ojk.go.id/id/regulasi/Documents/Pages/Penyelenggaraan-Usaha-Perusahaan-Pembiayaan/pojk%2035-2018.pdf>
- OJK. (2019). Retrieved from <https://www.ojk.go.id/id/regulasi/Documents/Pages/Penilaian-Kualitas-Aset-Bank-Umum/pojk%2040-2019.pdf>
- OJK. (2022, June 7). Retrieved from Statistik Fintech: <https://www.ojk.go.id/id/kanal/iknb/data-dan-statistik/fintech/Pages/Statistik-Fintech-Lending-Periode-April-2022.aspx>
- Otoritas Jasa Keuangan. (2022, May 18). Retrieved from <https://www.ojk.go.id/id/kanal/iknb/financial-technology/default.aspx>
- Scikit Learn. (2022). Retrieved from [https://scikit-learn.org/stable/modules/tree.html#:~:text=Decision%20Trees%20\(DTs\)%20are%20a,as%20a%20piecewise%20constant%20approximation.](https://scikit-learn.org/stable/modules/tree.html#:~:text=Decision%20Trees%20(DTs)%20are%20a,as%20a%20piecewise%20constant%20approximation.)
- Tang, H. (2019). Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? The Review of Financial Studies, 1900–1938.
- Velayudhan, A. (2020). Retrieved from <https://www.linkedin.com/pulse/type-i-ii-errors-credit-scoring-need-clear-velayudhan-frm-dipifr/>
- Yap, B. W., Ong, S. H., & Husain, N. H. M. (2011). Using data mining to improve assessment of credit worthiness via credit scoring models. *Expert Systems with Applications*, 38(10), 13274–13283. <https://doi.org/10.1016/j.eswa.2011.04.147>
- Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning*. Springer.
- Zhou, L., Fujita, H., & Ding, H. (2021). Credit risk modeling on data with two timestamps in peer-to-peer lending by gradient boosting. *Applied Soft Computing*, 107672.