

Received : April

Accepted: Juni

Published : Juni

Pemanfaatan Data Smartphone Untuk Prediksi Tingkat Risiko Gagal Bayar Debitur

Helen^{1*}, Roy Kurniawan²

*1*Management Department, Universitas Bina Nusantara

*2*Information Systems Department, Universitas Bina Nusantara

**helen.helen906@binus.ac.id*

Abstract

The purpose of this study is to demonstrate that smartphone usage data can be used to make predictions and find the best classification method for Credit Scoring even if the sample dataset is small. The study used a classification algorithm to divide customers into paying and arrears customers using mobile data, and then compared the predicted results with the actual results. There are several related works that are publicly accessible where mobile data has been used for Credit Scoring, but all of them use much larger data samples. Small companies cannot use large data sets as used in previous studies. In this paper the author tries to argue that there is data smartphone has good predictive power even though the dataset is small. The author concludes that with a sample data consisting of smartphone data as many as 5,702 debtors can still predict credit risk well. The best classification method using random forest yielded a result of 0.68 AUC equivalent to a ratio of 0.36

Keywords : credit scoring, classification, smartphone data

Abstrak

Tujuan dari penelitian ini adalah untuk mendemostrasikan bahwa data penggunaan ponsel dapat digunakan untuk membuat prediksi dan menemukan metode klasifikasi terbaik untuk Credit Scoring meskipun sample datasetnya sedikit. Penelitian ini menggunakan algoritma klasifikasi untuk membagi pelanggan menjadi pelanggan yang membayar dan menunggak menggunakan data seluler, dan kemudian membandingkan hasil yang diprediksi dengan hasil aktual. Ada beberapa karya terkait yang dapat diakses publik di mana data seluler telah digunakan untuk Credit Scoring, tetapi semuanya menggunakan sample data yang jauh lebih besar. Perusahaan kecil tidak dapat menggunakan kumpulan data sebesar yang digunakan pada penelitian sebelumnya. Dalam makalah ini penulis mencoba untuk berpendapat bahwa ada data smartphone memiliki predictive power baik meskipun himpunan datanya kecil. Penulis menyimpulkan bahwa dengan sample data yang terdiri dari data smartphone sebanyak 5.702 debitur masih dapat memprediksi risiko kredit dengan baik. Metode klasifikasi terbaik menggunakan random forest memberi hasil 0,68 AUC setara dengan gini rasio 0,36

Kata kunci : credit scoring, model klasifikasi, smartphone data

1. Pendahuluan

Credit Scoring model sangat membantu dalam meningkatkan kecepatan dan konsistensi pada proses aplikasi pinjaman dan memungkinkan perusahaan pemberi pinjaman untuk mengotomatiskan proses pinjaman mereka [1]. Dalam hal ini, Credit Scoring secara signifikan mengurangi keterlibatan

manusia dalam evaluasi kredit dan mengurangi biaya pengiriman kredit [2]. Selain itu, dengan menggunakan credit scoring lembaga keuangan dapat mengukur risiko yang terkait dengan pemberian kredit kepada pemohon tertentu dalam jangka waktu yang lebih singkat. Salah satu penelitian yang dilakukan pada sebuah bank Kanada menemukan bahwa waktu yang dibutuhkan untuk memproses aplikasi

pinjaman konsumen dipersingkat dari sembilan hari menjadi tiga hari setelah Credit Scoring digunakan [3]. Dengan demikian, optimalisasi waktu pemrosesan pinjaman berarti bahwa waktu yang dihemat untuk pemrosesan dapat dipertimbangkan untuk mengatasi aspek yang lebih kompleks di perusahaan. Banasiak dan Kiely [4] menyimpulkan bahwa dengan bantuan skor kredit, lembaga keuangan mampu membuat keputusan yang lebih cepat, lebih baik dan lebih tinggi.

Ada lebih dari 2 miliar orang di dunia yang tidak memiliki rekening bank [5]. Hal ini membuat sulit untuk melakukan latihan evaluasi kredit untuk orang-orang ini. Dengan maraknya big data dan sumber data baru, berbagai alternatif data dapat digunakan untuk menjelaskan inklusi keuangan individu-individu yang tidak memiliki rekening bank ini. Misalnya, data penggunaan seluler adalah sumber data baru yang dapat digunakan. Data penggunaan ponsel dapat dianggap sebagai data alternatif yang prediktif untuk Credit Scoring.

Penggunaan data smartphone untuk menentukan kepribadian pengguna ponsel adalah topik hangat, yang banyak dibahas. Jumlah pengguna ponsel telah mencapai 6 miliar di seluruh dunia dan [6] penyedia layanan memungkinkan peningkatan akses ke log telepon kepada para peneliti dan mitra komersial. Jika dimanfaatkan secara tepat, data smartphone dapat memberikan insight yang berharga dan hemat biaya. Misalnya, pemasar dan perusahaan manufaktur telepon mungkin berusaha mengakses informasi disposisional tentang pelanggan mereka sehingga dapat merancang penawaran dan promosi yang disesuaikan. Dengan demikian, perusahaan mendapat manfaat dari menggunakan data yang dikumpulkan secara otomatis. Kemampuan untuk mengekstrak kepribadian dan variabel psikososial lainnya dari data smartphone merupakan penemuan yang baru di bidang ilmu sosial

Penggunaan ponsel untuk memprediksi kepribadian masyarakat adalah hasil dari kemajuan machine learning dan ilmu sosial

komputasi yang telah memungkinkan untuk menyimpulkan berbagai keadaan dan sifat psikologis berdasarkan bagaimana orang menggunakan ponsel mereka setiap hari. Misalnya, beberapa penelitian telah menunjukkan bahwa kepribadian orang dapat diprediksi berdasarkan pola bagaimana mereka menggunakan media sosial seperti Facebook atau Twitter [7]. Peneliti lain telah menggunakan informasi tentang penggunaan orang-orang dari berbagai aplikasi seluler seperti YouTube, Kalender Internet, permainan, dan sebagainya untuk membuat kesimpulan [8]. Meskipun pendekatan ini luar biasa, mereka membutuhkan akses ke informasi terkait seluruh jejaring sosial seseorang. Keterbatasan ini sangat melemahkan penggunaan metode klasifikasi semacam itu untuk penyelidikan skala besar

Kontribusi utama dari makalah ini adalah untuk menunjukkan bahwa data penggunaan ponsel, alih-alih pengumpulan data konvensional dan proses semimanual lainnya, dapat berhasil digunakan untuk memprediksi risiko kredit meskipun datasetnya kecil. Kontribusi masih perlu untuk didiskusikan secara ilmiah serta penerapan praktisi di industri. Dari sudut pandang teknis, kemampuan predictive data alternative sama kuatnya dengan data konvensional.

2. Tinjauan Pustaka

Credit Scoring adalah pemanfaatan statistical model untuk mengkuantifikasi tingkat resiko kegagalan bayar dari debitur. Pada dasarnya, *Credit Scoring* adalah proksi untuk mengestimasi tingkat risiko dan kepercayaan. pengembangan credit scoring yang subjektif pertama kali disediakan oleh biro kredit abad ke-19 [3]

Banyak tinjauan pustaka berfokus pada pengembangan, penerapan, dan evaluasi model prediktif yang digunakan dalam kredit [9]. Algoritma yang digunakan untuk membangun credit model bisa berupa parametric approach model seperti regresi logistic, maupun non parametric approach seperti support vector machine, neural network dan random forest

Model-model ini menentukan kelayakan kredit pemohon berdasarkan serangkaian variabel deskriptif. Credit Scoring untuk perusahaan menggunakan data dari laporan tentang posisi keuangan, rasio keuangan atau makro-ekonomi, sedangkan model risiko ritel menggunakan data yang diperoleh dalam formulir aplikasi seperti riwayat transaksi pelanggan [10]. Perbedaan antara variabel yang digunakan dalam corporate credit scoring dan ritel menunjukkan bahwa lebih banyak tantangan muncul pada credit scoring ritel daripada Credit Scoring perusahaan. Tulisan ini berfokus pada bisnis ritel.

Ada beberapa indikasi tentang penggunaan data penggunaan ponsel dalam dunia korporat credit scoring. Namun, hanya sangat sedikit makalah yang relevan [11] yang terbuka untuk komunitas penelitian yang lebih luas. Bjrkegren dan Grissen menggunakan data ponsel untuk memprediksi default dengan akurasi yang hampir mirip dengan metode Credit Scoring yang menggunakan riwayat keuangan. Pendekatan ini divalidasi menggunakan catatan panggilan yang dicocokkan dengan hasil pinjaman untuk sampel peminjam di negara Karibia. Pelamar dalam kuartil risiko tertinggi menurut ukuran penulis enam kali lebih mungkin gagal bayar daripada mereka yang berada di kuartil terendah. Mereka menggunakan dua algoritma yang berbeda, Random Forest dan Logistic regression. Hasil yang diperoleh dengan algoritma Random Forest adalah 0,710 AUC dan dengan regresi Logistik 0,760 AUC. Himpunan data mencakup informasi tentang 7.068 pelanggan dari negara Amerika Selatan [11]. Jose San Pedro et al. membangun MobiScore [12], sebuah metodologi yang digunakan untuk membangun model risiko keuangan pengguna menggunakan data yang dikumpulkan dari penggunaan seluler. MobiScore menggunakan data pada 60.000 orang yang diperoleh dari perusahaan telekomunikasi dan penyedia layanan keuangan di negara Amerika Latin. Mereka menggunakan peningkatan gradien, mesin vektor pendukung, dan model regresi linier untuk menyelesaikan masalah. Hasil

AUC dengan kombinasi yang berbeda adalah antara 64,1 dan 72,5 . Speakman et al. [13] mendemonstrasikan bagaimana menggunakan decision tree yang didorong untuk menciptakan skor kredit bagi populasi yang kurang memiliki rekening bank, memungkinkan mereka untuk mengakses fasilitas kredit yang sebelumnya ditolak karena tidak tersedianya data keuangan. Hasil penelitian mereka adalah penurunan 55% dalam tingkat default sementara simultaneously menawarkan peluang kredit kepada satu juta pelanggan yang diberi batas kredit 0 dalam model asli bank. Himpunan data berisi 295.926 contoh berlabel dengan lebih dari 30 fitur kategoris dan bernilai nyata. Hasil AUC dengan algoritma decision tree yang ditingkatkan adalah 0,764 dan dengan regresi logistik 0,74 [13]

3. Metode Penelitian

Penelitian ini diawali dengan tinjauan pustaka serta pengumpulan data sample yang dari perusahaan lembaga keuangan yang memiliki mobil apps sebagai channel akuisisi, Sample data yang digunakan sebesar 5.702 berupa data produk cash loan periode 2019 hingga 2021.

Sedangkan data smartphone didapat dengan permission dari borrower untuk mengakses device data.

3.1. Data

Sample data terdiri dari informasi tentang 5.702 debitur yang telah memperoleh cash loan, sehingga memungkinkan untuk mendapatkan target variable berdasarkan ketepatan pembayaran mereka.

Dengan menggunakan perilaku pembayaran mereka, penulis dapat memisahkan debitur good dan bad. Berdasarkan Analisa flow rate dan vintage diputuskan target variable adalah debitur yang telah mengalami keterlambatan pembayaran 60 hari atas angsuran mereka. Selain itu, dataset will mencakup sekitar 4.516 debitur yang membayar tanpa menunggak melebihi 60 hari. Sehingga perbandingan good dan bad debitur menjadi 79,2% dan 20,8%.

Pengguna ponsel Android dapat diminta untuk memberikan data berikut tentang perangkat mereka (lihat Tabel 1). Untuk penelitian ini penulis tidak menggunakan nomor telepon, teks isi kalender atau pesan teks (SMS).

Dari antara semua parameter yang bervariasi, 22 variabel dipilih untuk digunakan dalam percobaan yang diperlukan untuk penelitian (variabel ditunjukkan pada Tabel 2). Variabel dipilih dengan menggunakan tinjauan manual dan analisis statistik dependensi. Penulis memilih variabel yang kurang bergantung satu sama lain. Dengan menggunakan variabel, salah satunya adalah variabel kategorikal sedangkan yang lain bersifat numerikal. Dalam beberapa percobaan, penulis mendisketisasi beberapa variabel numerik menjadi binning sehingga jenis datanya berubah menjadi kategorikal

3.2. Pengukuran

Harris [14] menjelaskan bahwa dalam proses train model perlu dipisahkan sample untuk train, test dan evaluasi. Hal ini bertujuan untuk melakukan evaluasi dengan data yang berbeda dengan data training.

Dalam analisis ini, metrik kinerja dan metrik evaluasi model utama diwakili oleh wilayah di bawah kurva ROC (Receiver Operating Characteristic) yang disebut AUC. Kurva ROC, sering diadopsi oleh AUC, menggambarkan dua komponen kinerja diferensial di mana sensitivitas yaitu jumlah relatif dari positif aktual yang diperkirakan positif dan spesifisitas yaitu proporsi negatif aktual yang diperkirakan negatif. Biasanya, angka AUC ditunjukkan seperti pada gambar di bawah ini di mana menggambarkan jumlah total peringkat kelayakan kredit pelanggan. [15].

Grup data	Deskripsi data
Device data	ID Perangkat
Device data	Versi OS (sistem operasi)
Device data	Versi SDK (kit pengembangan perangkat lunak)
Device data	Versi rilis
Device data	Merek
Device data	Menampilkan

Device data	Perangkat keras
Device data	Produsen
Device data	Serial
Device data	Pengguna
Jaringan	ID Jaringan
Jaringan	Operator
Jaringan	Pelanggan
Kalender	ID Kalender
Info panggilan	ID Penelepon
Info panggilan	Penerima (kontak / tidak diketahui)
Info panggilan	masuk/keluar/terlewat/tidak terjawab
Info panggilan	Jumlah panggilan
Info panggilan	Tanggal
Info panggilan	Durasi
Info kontak	ID Kontak
Info kontak	Nomor kontak
Aplikasi yang diinstal	ID Aplikasi
Aplikasi yang diinstal	Nama paket
Aplikasi yang diinstal	Label
Aplikasi yang diinstal	Nama versi
Aplikasi yang diinstal	Kode versi
Aplikasi yang diinstal	Tanggal pemasangan
SMS	SMS ID
SMS	jumlah
SMS	Panjang pesan
SMS	Tanggal SMS
SMS	SMS ID
Gambar	ID Gambar
Gambar	Tanggal gambar
Gambar	Lokasi gambar
Penyimpanan data	ID penyimpanan data
Penyimpanan data	Terakhir dimodifikasi

Grup data	Titik data terhitung	Tipe data
Info panggilan	Jumlah rata-rata panggilan per bulan.	Numerik
Info panggilan	Rata-rata jumlah panggilan masuk per bulan.	Numerik
Info panggilan	Jumlah rata-rata panggilan keluar per bulan.	Numerik
Info panggilan	Jumlah rata-rata panggilan tak terjawab per bulan.	Numerik
Info panggilan	Jumlah rata-rata panggilan yang tidak dijawab per bulan.	Numerik
Info panggilan	Durasi panggilan rata-rata.	Numerik
Info panggilan	Durasi panggilan keluar rata-rata.	Numerik

Info panggilan	Durasi panggilan masuk rata-rata.	Numerik
Info panggilan	Durasi panggilan keluar maksimum.	Numerik
Info panggilan	Durasi panggilan masuk maksimum.	Numerik
Gambar	Jumlah rata-rata gambar per bulan.	Numerik
Gambar	Jumlah rata-rata gambar yang dibuat di tempat yang berbeda per bulan.	Numerik
SMS	Jumlah rata-rata SMS per bulan.	Numerik
SMS	Jumlah rata-rata SMS yang masuk per bulan.	Numerik
SMS	Jumlah rata-rata SMS yang masuk per bulan dari kontak.	Numerik
SMS	Jumlah rata-rata SMS yang masuk per bulan dari jumlah yang tidak diketahui.	Numerik
SMS	Jumlah rata-rata SMS keluar per bulan.	Numerik
SMS	Jumlah rata-rata SMS keluar per bulan dari kontak.	Numerik
SMS	Jumlah rata-rata SMS keluar per bulan dari jumlah yang tidak diketahui.	Numerik
SMS	Jumlah rata-rata percakapan SMS per bulan.	Numerik
Kontak	Jumlah kontak.	Numerik
Alat	Versi SDK.	Kategorik

$$Sensitivity = \frac{true\ positive}{true\ positive + false\ negative} \quad (1)$$

$$Specificity = \frac{true\ negative}{true\ positive + true\ negative} \quad (2)$$

$$AUC = \frac{(S1 - sensitivity) + [(sensitivity + 1) + 0.5]}{sensitivity + specificity}$$

3.3. Eksperimen

Penulis melakukan tiga percobaan dengan 4 metode klasifikasi yang berbeda dan menggunakan AUC sebagai parameter kinerja. Seperti yang telah diilustrasikan oleh penulis sebelumnya, tidak ada aturan khusus untuk bekerja dengan data alternatif. Oleh karena itu, penulis melakukan empat percobaan

berdasarkan teknik pra-pemrosesan yang berbeda.

Dalam percobaan pertama penulis memasukkan semua variabel yang dihitung. Variabel SDK, yang kategoris, perlu dikodekan. Data SDK harus dikonversi menjadi angka agar dapat dibandingkan. Versi SDK terdiri dari enam nilai berbeda yang penulis hasilkan variabel tiruannya. Nilai parameter ini adalah 1 atau 0. Akibatnya, tidak ada informasi yang hilang dalam himpunan data. Langkah kedua dalam preprocessing data adalah menskalakan semua variabel untuk membuatnya dibandingkan satu sama lain.

Dalam percobaan kedua penulis menggunakan teknik pra-pemrosesan yang sama seperti pada percobaan pertama, tetapi penulis menambahkan eliminasi mundur. Penelitian [3] menyatakan bahwa model harus sesederhana mungkin sampai mencapai tingkat kinerja yang dapat diterima pada data pelatihan. Ini akan membantu menghindari pemasangan model yang berlebihan. Dengan backward elimination kita dapat membuang variabel dengan p-value (nilai probabilitas) >0,05 dan nilai p tertinggi. Setelah itu kita dapat menghitung ombinasi c barudari nilai p dan melanjutkan proses yang sama sampai kita memiliki satu set variabel, semua dengan p lebih rendah dari 0,05.

Dalam percobaan ketiga penulis menggunakan metode pra-pemrosesan yang sama seperti sebelumnya tetapi memodifikasi variabel. Penulis menggunakan teknik binning yang optimal untuk mengelompokkan variabel. Binning optimal adalah metode prediktor kategoris pra-pemrosesan di mana penulis menetapkan nilai untuk variabel dengan mengelompokkannya ke dalam bin yang optimal. Tujuannya adalah untuk mengurangi dampak noise statistik

Untuk memilih metode classifier untuk bagian eksperimental, penulis menggunakan 2 parameter:

Metode apa yang digunakan untuk train Credit Scoring sebelumnya?

Bagaimana metode train Credit Scoring sebelumnya menggunakan data seluler?

Menurut kedua parameter ini, penulis memilih untuk eksperimen penulis metode berikut: regresi logistik, random forest, SVM, dan neural network.

Saat membagi train dan test dataset ada kendala dalam menentukan ukuran set tes yang akan memberikan hasil yang signifikan secara statistik. Rasio yang umum diadopsi adalah 8:2 sesuai dengan prinsip Pareto.

Menurut penelitian isabelle Guyon dan formula yang dia temukan kita dapat menentukane contoh ukuran tes. Fraksi pola yang disediakan untuk set validasi harus berbanding terbalik dengan akar kuadrat dari jumlah parameter yang dapat disesuaikan bebas. Rasio himpunan validasi (v) dengan himpunan pelatihan (t) adalah v/t , dan s adalah $\ln(N/h-\max)$, di mana N adalah jumlah keluarga pengenalan, dan $h-\max$ adalah kompleksitas terbesar dari keluarga-keluarga tersebut. Setiap keluarga recognizer dicirikan oleh kompleksitasnya, yang mungkin atau mungkin tidak terkait dengan dimensi VC (dimensi nenkis VapnikChervo), panjang deskripsi, jumlah parameter yang dapat disesuaikan, atau ukuran kompleksitas lainnya [13]. Karena sample dataset kecil, penulis memilih tiga contoh ukuran pengujian yang berbeda untuk penelitian ini. Ukuran tes yang penulis pilih adalah 10%, 20% dan 30%.

Pengujian kombinasi variabel apa pun terlebih dahulu menghasilkan semua variabel. Penulis kemudian memilih hanya variabel dengan $p < 0,05$ dan akhirnya mengikat variabel dengan $p < 0,05$. Interval variabel dapat ditentukan dalam berbagai cara. Untuk example, dengan menggunakan pengetahuan sebelumnya pada data. Batas-batas interval biasanya didefinisikan sebelumnya.

4. Hasil

Eksperimen dilakukan dengan menggunakan bahasa pemrograman Python dan lingkungan Spyder. Penulis juga menggunakan library Python numpy,

matplotlib, panda, dan scikit-learn untuk analisis statistik.

Tabel 3, 4 dan 5 menunjukkan representasi kinerja metode klasifikasi menggunakan data seluler. Hasil pada Tabel 3 menunjukkan kinerja pengklasifikasi dengan semua variabel. Hasil pada Tabel 4 menunjukkan kinerja pengklasifikasi dengan hanya variabel yang nilai p-nya lebih rendah dari 0,05. Tmampu 5 menunjukkan variabel yang diikat yang nilai p-nya lebih rendah dari 0,05. Tabel menyarankan model yang dibuat untuk prediksi kelayakan kredit seperti yang diilustrasikan oleh AUC pada himpunan data yang ditekan. Untuk menentukan pentingnya variasi kinerja between model kita dapat mengambil AUC sebagai parameter utama untuk melihat model mana yang memiliki kinerja terbaik. Tabel 3, 4 dan 5 juga dapat dibandingkan untuk akurasi pelatihan, akurasi tes, dan waktu pelatihan.

Variabel target yang dipilih adalah 0 untuk pelanggan yang berkinerjad 1 untuk pelanggan yang tidak berkinerja. Pelanggan yang tidak berkinerja dalam penelitian ini ditetapkan sebagai orang yang 90 hari atau lebih terlambat dalam membayar utang mereka. Menurut [3]aturan jatuh tempo 90 hari paling umum di negara Eropa dari mana data untuk penelitian ini dikumpulkan.

Metode Klasifikasi	Split train test	Train accuracy	Test accuracy	AUC
Regresi logistik	0.10 : 0.90	0.64	0.63	0.51
Regresi logistik	0.20 : 0.80	0.63	0.64	0.55
Regresi logistik	0.30 : 0.70	0.64	0.63	0.56
SVM	0.10 : 0.90	0.65	0.67	0.56
SVM	0.20 : 0.80	0.64	0.61	0.56
SVM	0.30 : 0.70	0.61	0.61	0.57
Neural network	0.10 : 0.90	0.71	0.64	0.57
Neural network	0.20 : 0.80	0.69	0.62	0.57
Neural network	0.30 : 0.70	0.71	0.66	0.55

Random forest	0.10 : 0.90	0.99	0.62	0.63
Random forest	0.20 : 0.80	0.99	0.64	0.60
Random forest	0.30 : 0.70	0.99	0.62	0.59

Random forest memberikan tingkat AUC rata-rata terbaik di seluruh eksperimen dengan ukuran pengujian yang berbeda. Random forest juga memberi peringkat AUC terbaik di 0,64 dengan semua variabel dan ukuran uji 10. Metode terbaik kedua adalah neural network dengan AUC tertinggi dan semua variabel menggunakan 10 untuk ukuran pengujian. Menurut pengetahuan penulis sebelumnya mengenai memilih ukuran pengujian untuk kumpulan data kecil yang terdiri dari 5.702 pelanggan, penulis dapat mengambil hasil yang paling stabil pada ukuran pengujian 40 . Dengan ukuran pengujian menjadi 40, penulis memperoleh hasil terbaik dari percobaan pertama dengan algoritma random forest AUC = 0,58, dan hasil yang sama dari algoritma neural network dalam percobaan kedua dengan hanya variabel di mana $p < 0,05$.

Metode Klasifikasi	Split train test	Train accuracy	Test accuracy	AUC
Regresi logistik	0.10 : 0.90	0.67	0.63	0.55
Regresi logistik	0.20 : 0.80	0.77	0.57	0.54
Regresi logistik	0.30 : 0.70	0.77	0.54	0.51
SVM	0.10 : 0.90	0.62	0.66	0.54
SVM	0.20 : 0.80	0.61	0.58	0.57
SVM	0.30 : 0.70	0.61	0.58	0.58
Neural network	0.10 : 0.90	0.64	0.60	0.58
Neural network	0.20 : 0.80	0.64	0.61	0.56
Neural network	0.30 : 0.70	0.63	0.61	0.59
Random forest	0.10 : 0.90	0.97	0.67	0.57
Random forest	0.20 : 0.80	0.97	0.63	0.58

Random forest	0.30 : 0.70	0.97	0.61	0.59
----------------------	-------------	------	------	------

Metode Klasifikasi	Split train test	Train accuracy	Test accuracy	Auc
Regresi logistik	0.10 : 0.90	0.68	0.66	0.54
Regresi logistik	0.20 : 0.80	0.77	0.57	0.55
Regresi logistik	0.30 : 0.70	0.76	0.57	0.54
SVM	0.10 : 0.90	0.61	0.65	0.49
SVM	0.20 : 0.80	0.60	0.60	0.49
SVM	0.30 : 0.70	0.60	0.60	0.49
Neural network	0.10 : 0.90	0.62	0.64	0.55
Neural network	0.20 : 0.80	0.62	0.63	0.57
Neural network	0.30 : 0.70	0.63	0.60	0.58
Random forest	0.10 : 0.90	0.74	0.60	0.65
Random forest	0.20 : 0.80	0.76	0.58	0.67
Random forest	0.30 : 0.70	0.77	0.57	0.66

Hasil terlemah secara keseluruhan diperoleh dari algoritma SVM, yang menghasilkan hasil yang kurang baik, di mana AUC berada di bawah 0,50. Algoritma logistic regression menunjukkan hasil yang paling stabil di seluruh eksperimen dan ukuran pengujian, dengan AUC antara 0,54 dan 0,55.

Dibandingkan hasil dari credit model pada Tabel 6, jelas bahwa dalam penelitian ini hasil AUC lebih rendah daripada yang lain. mengingat ukuran sampel yang digunakan hanya 5.702 debitur dibandingkan dengan penelitian sebelumnya yang menggunakan sample jauh lebih besar.

Kerja	sample	Metode	Auc
Jose San Pedro, Davide Proserpio dan Nuria Oliver, 2015 [12]	295,926	decision tree	0.740
Skyler Speakman, Eric Mibuari, Isaac Markus, Felix Kwizera, 2017 [13]	60,000	Regresi linier	0.725
Bjrkegren dan Grissen, 2017 [11]	7,068	Random forest	0.710
Penelitian saat ini	2,503	Regresi logistik	0.56
Penelitian saat ini	2,503	SVM	0.58
Penelitian saat ini	2,503	Neural network	0.59

Penelitian saat ini	2,503	Random forest	0.67
---------------------	-------	---------------	------

Penelitian penulis tentang penggunaan data smartphone akan memungkinkan organisasi lain di sektor keuangan menggunakan data smartphone untuk Credit Scoring mereka. Pada penelitian sebelumnya menunjukkan bahwa data smartphone hanya berguna dengan sample dataset yang besar, penulis berpendapat bahwa dengan sample kecil pun bisa dijalankan.

5. Kesimpulan dan Saran

Tantangan utama bagi lending business selama beberapa tahun terakhir adalah inklusi keuangan di mana segment konsumen terbesar adalah masyarakat yang tidak memiliki akses terhadap perbankan. Ada lebih dari 2 miliar orang di dunia termasuk kategori unbankable dan jumlah pengguna ponsel telah mencapai 6 miliar di seluruh dunia.

Penelitian ini memiliki dua kontribusi teoritis yang penting. Pertama, berdasarkan penggunaan data smartphone untuk pengembangan model Credit Scoring, kita dapat melihat bahwa semua metode yang diuji dengan semua variabel menghasilkan hasil yang lebih baik daripada dalam studi acak.

Kedua, penulis secara empiris menyimpulkan bahwa metode terbaik untuk credit scoring berdasarkan data seluler adalah metode klasifikasi random forests dengan AUC 0,67.

Makalah ini didasarkan pada sintesis penelitian berpendapat bahwa smartphone data dapat memberikan hasil positif untuk Credit Scoring meskipun dengan sample kecil. Temuan penulis juga mengungkapkan bahwa model terbaik dalam hal untuk mentrain model Credit Scoring adalah metode random forest.

Jika perusahaan pembiayaan ingin memiliki data yang lebih akurat tentang pelanggan yang lebih cenderung membayar kembali pinjaman mereka, mereka perlu mencari sumber data alternatif seperti data ponsel. Ini akan memberikan keuntungan besar bagi lembaga keuangan di negara-negara dunia ketiga di mana kebanyakan orang tidak memiliki riwayat

bank, satu-satunya data yang mereka miliki adalah data ponsel mereka.

Penulis berharap penelitian ini membuka diskusi lebih lanjut mengenai pemahaman yang lebih akurat tentang bagaimana kita dapat menggunakan data smartphone untuk menghasilkan prediksi dan memberi nilai tambah. Penelitian ini tidak hanya relevan untuk perusahaan sektor keuangan tetapi juga di bidang asuransi atau pencegahan penipuan, di mana data mobil phone dapat membantu membuat prediksi.

Ada banyak peluang untuk meneliti lebih jauh mengenai pemanfaatan data smartphone ini. Prediksi dari data smartphone ini juga dapat digunakan di berbagai macam industri, tidak hanya di bidang keuangan. Sangat mungkin bisa digunakan untuk memprediksi risiko asuransi atau penipuan.

6. Daftar Pustaka

- [1] J. Rimmer, "Contemporary changes in credit scoring. *Credit Control*, vol. 26, no. 4, pp. 56-60, 2005.
- [2] C. Wendel and M. Harvey, "Credit scoring: Best practices and approaches," *Commercial Lending Review*, vol. 18, no. 1, p. 4, 2003.
- [3] R. Anderson, *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*, Oxford University Press, 2007.
- [4] M. Banasiak and G. Kiely, "Predictive collection score technology," *Business Credit*, vol. 102, no. 2, pp. 18-34, 2000.
- [5] C. Hodgson, "The world's 2 billion unbanked, in 6 charts," 2017.
- [6] W. Lance, "2011 ends with almost 6 billion mobile phone subscriptions (2012)," 2012.
- [7] M. D. Back, J. M. Stopfer, S. Vazire and Gaddis, "Facebook profiles reflect actual

- personality, not self-idealization," *Psychological science*, vol. 21, no. 3, pp. 372-374, 2010.
- [8] G. Chittaranjan, J. Blom and D. Gatica Perez, "Mining large-scale smartphone data for personality studies," *Personal and Ubiquitous Computing*, vol. 17, no. 3, pp. 433-450, 2013.
- [9] J. N. Crook, D. B. Edelman and L. C. Thomas, "Recent developments in consumer credit risk assessment. 183(3), 1447– 1465 (2007)," *European Journal of Operational Research*, vol. 183, no. 3, pp. 1447-1465, 2007.
- [10] L. C. Thomas, "A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. 16(2), 149–172 (2000)," *International journal of forecasting*, vol. 16, no. 2, pp. 149-172, 2000.
- [11] D. Bjrkegren and D. Grissen, The development of credit scoring quality measures for consumer credit applications. Netmob (2013), Netmob, 2013.
- [12] J. S. P. Pedro, D. Proserpio and N. Oliver, "Mobiscore: towards universal credit scoring from mobile phone data," *International Conference on User Modeling, Adaptation, and Personalization*, pp. 195-207, 2015.
- [13] S. Speakman, E. Mibuari, I. Markus and F. Kwizera, Mobile phone-based credit scoring. NetMob Book of Abstracts, NetMob Book of Abstracts, 2017.
- [14] T. Harris, "Quantitative credit risk assessment using support vector machines: Broad versus narrow default definitions," *Expert Systems with Applications*, vol. 40, no. 11, pp. 4404-4413, 2013.