

DURASI PENGGUNAAN NOMOR TELEPON SELULER SEBAGAI INDIKATOR RISIKO KREDIT: ANALISIS DATA EKSPLORATIF BERBASIS MACHINE LEARNING

Yosy Arisandy^{1*}, Yosza Bin Dasril², Shahrul Nizam Bin Salahudin³

Universitas Islam Negeri Fatmawati Sukarno Bengkulu, JL. Raden Fatah, Pagar Dewa Kec. Selebar
Bengkulu city, Bengkulu, Indonesia¹

Faculty of Electronics & Computer Technology & Engineering, Universiti Teknikal Malaysia
Melaka Durian Tunggal 76100 Melaka, Malaysia²

Universiti Tun Husein Onn Malaysia, Faculty of Technology Management and Business, 86400
Parit Raja, Batu Pahat, Johor, Malaysia³

Email: yosyarisandy@mail.uinfabengkulu.ac.id*, yosza@utem.edu.my, shahrulns@uthm.edu.my

ABSTRAK

Pertumbuhan pesat platform pinjaman digital tanpa agunan di Asia Tenggara menuntut sistem penilaian risiko kredit yang lebih akurat dan inklusif, khususnya bagi peminjam yang belum memiliki riwayat kredit formal. Pemanfaatan data alternatif, termasuk informasi penggunaan telepon seluler, semakin dipertimbangkan dalam pengembangan sistem credit scoring. Namun, kajian khusus yang menilai durasi penggunaan nomor telepon seluler sebagai indikator stabilitas perilaku dalam memprediksi risiko kredit masih terbatas. Penelitian ini bertujuan menganalisis peran durasi penggunaan nomor telepon seluler sebagai prediktor risiko kredit pinjaman tanpa agunan. Studi ini menggunakan data sekunder dari dataset lending di Kaggle yang mencakup 307.506 debitur di Asia Tenggara dengan 30 variabel terpilih. Analisis dilakukan menggunakan pendekatan Exploratory Data Analysis (EDA) melalui teknik univariat, bivariat, dan multivariat dengan bantuan Python pada platform Google Colaboratory. Hasil analisis menunjukkan bahwa peminjam dengan durasi penggunaan nomor telepon yang lebih pendek dan usia yang lebih muda cenderung memiliki probabilitas gagal bayar lebih tinggi dibandingkan peminjam dengan penggunaan nomor yang lebih lama. Meskipun hubungan linier relatif lemah, variabel ini menunjukkan potensi sebagai indikator stabilitas perilaku ketika dikombinasikan dengan karakteristik demografis dan status pekerjaan. Temuan ini menunjukkan bahwa indikator telekomunikasi sederhana dapat memperkaya pengembangan model credit scoring berbasis machine learning yang lebih inklusif dalam ekosistem keuangan digital.

Kata Kunci : Risiko kredit, Exploratory Data Analysis, Durasi nomor telepon, Machine learning, Pinjaman tanpa agunan.

ABSTRACT

The rapid growth of unsecured digital lending platforms in Southeast Asia has increased the need for more accurate and inclusive credit risk assessment systems, particularly for borrowers with limited or no formal credit history. In this context, alternative data, including mobile phone usage information, is increasingly considered in the development of credit scoring models. However, empirical studies examining the tenure of mobile phone number usage as an indicator of behavioral stability in predicting credit risk remain limited. This study analyzes the role of mobile phone number tenure as a predictor of credit risk in unsecured lending. The study uses secondary data from a lending dataset available on Kaggle, covering 307,506 borrowers in Southeast Asia with 30 selected variables. The analysis applies an Exploratory Data Analysis (EDA) approach using univariate, bivariate, and multivariate techniques implemented in Python on the Google Colaboratory platform. The results show that borrowers with shorter mobile phone number tenure and younger age tend to have a higher probability of default compared to those with longer phone number tenure. Although the linear relationship is relatively weak, this variable still provides meaningful predictive signals when combined with demographic characteristics and employment status. These findings suggest that simple telecommunications indicators can enhance machine learning-based credit scoring models in digital financial ecosystems.

Keywords : Credit risk; Exploratory Data Analysis; Mobile phone number tenure; Machine learning; Unsecured loans.

1. PENDAHULUAN

Di pasar berkembang Asia Tenggara, sebagian besar populasi dewasa masih berada dalam kondisi underbanked atau tidak memiliki riwayat kredit formal (credit-invisible), sehingga akses terhadap layanan kredit formal menjadi terbatas (Adam et al., 2025). Kesenjangan struktural dalam akses terhadap pinjaman beragunan ini beriringan dengan pertumbuhan pesat platform pinjaman digital tanpa agunan di Indonesia, yang secara agresif memasarkankan proses pendaftaran yang sederhana dan cepat (Nainggolan & Umaroh, 2023), bahkan seringkali mengklaim bahwa pemohon dapat memperoleh pinjaman hanya dengan menggunakan nomor telepon seluler (Putri Nurhidayati, 2025). Meskipun kemudahan akses tersebut diposisikan sebagai strategi inklusi keuangan, fenomena ini memunculkan pertanyaan mendasar: apakah penggunaan data berbasis telepon semata-mata merupakan strategi pemasaran untuk menarik peminjam, atau justru mencerminkan inovasi substantif dalam penilaian risiko kredit?

Penelitian terdahulu menunjukkan bahwa data perilaku yang bersumber dari perangkat seluler dapat meningkatkan kemampuan prediksi gagal bayar, khususnya pada peminjam dengan riwayat kredit terbatas (*thin-file borrowers*) (Agarwal et al., 2019; Björkegren & Grissen, 2019). Metode credit scoring tradisional yang sangat bergantung pada data historis keuangan dan catatan biro kredit seringkali gagal menangkap kelayakan kredit kelompok ini (Nwachukwu, 2024). Kondisi tersebut mendorong meningkatnya pemanfaatan data alternatif, terutama pola penggunaan telepon seluler, dalam pengembangan model evaluasi risiko kredit.

Meskipun data penggunaan telepon seluler telah menjadi instrumen penting dalam pemodelan risiko kredit (Björkegren & Grissen, 2019; Ots et al., 2020), masih terdapat celah penelitian yang signifikan. Berbagai studi sebelumnya lebih banyak mengeksplorasi sinyal perilaku berbasis metadata berfrekuensi tinggi, seperti pola panggilan dan penggunaan aplikasi (Auliya et al., 2021). Demikian pula, studi pada sistem peer-to-peer lending di Tiongkok lebih menekankan pola pergantian perangkat dan instalasi aplikasi (Liu et al.,

2019), tanpa mempertimbangkan stabilitas yang tercermin dari lamanya penggunaan nomor. Studi terbaru menunjukkan bahwa data penggunaan telepon seluler dapat dimanfaatkan sebagai sumber alternative data dalam meningkatkan akurasi prediksi risiko kredit. Penelitian tersebut memanfaatkan metadata penggunaan smartphone seperti riwayat panggilan (Razavi & Elbahnasawy, 2025), aktivitas pesan singkat, penggunaan aplikasi, atau pola aktivitas digital untuk membangun model klasifikasi kredit menggunakan algoritma machine learning (Chen et al., 2025).

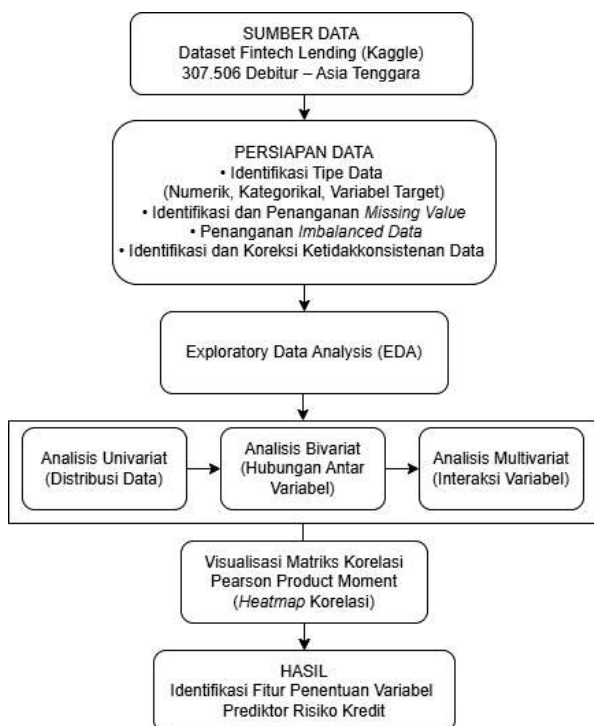
Di sisi lain, masih relatif terbatas penelitian yang mengkaji indikator identitas yang lebih sederhana namun persisten, seperti durasi penggunaan nomor telepon seluler (phone number tenure), sebagai sinyal stabilitas perilaku dalam penilaian risiko kredit. Padahal, variabel ini berpotensi mencerminkan kontinuitas identitas, stabilitas jaringan sosial, dan konsistensi perilaku finansial peminjam, yang secara teoritis dapat berhubungan dengan probabilitas gagal bayar. Namun, pendekatan tersebut umumnya memerlukan data perilaku digital yang kompleks dan berfrekuensi tinggi, yang tidak selalu tersedia atau mudah diakses oleh seluruh lembaga keuangan, terutama pada skala usaha yang lebih kecil. Oleh karena itu, penelitian ini berupaya mengisi kesenjangan tersebut dengan menguji durasi nomor telepon seluler sebagai prediktor sederhana namun stabil dalam konteks pinjaman tanpa agunan di Asia Tenggara, wilayah dengan penetrasi seluler tinggi dan pertumbuhan fintech yang pesat, namun akses kredit formal yang masih terbatas.

Penelitian ini menggunakan pendekatan Exploratory Data Analysis (EDA) untuk mengeksplorasi hubungan potensial antarvariabel dalam data pinjaman konsumen tanpa agunan. Pendekatan EDA memungkinkan visualisasi pola, anomali, serta hubungan yang sebelumnya tidak tampak secara eksplisit (Arisandy et al., 2023), termasuk korelasi antara durasi penggunaan nomor telepon dan risiko kredit. Melalui teknik visualisasi seperti *heatmap* dan diagram korelasi, EDA memberikan gambaran awal mengenai faktor-faktor yang memengaruhi risiko kredit, sebagai dasar untuk

pengembangan analisis lanjutan dalam kerangka pemodelan berbasis machine learning.

2. METODE

Penelitian ini menggunakan data sekunder yang diperoleh dari dataset fintech lending yang tersedia di Kaggle dan bersumber dari lembaga pemberi pinjaman di Eropa. Dataset `application.csv` mencakup 307.506 pemohon pinjaman dari kawasan Asia Tenggara dengan 30 variabel terpilih. Data tersebut berisi catatan anonim pemohon pinjaman yang meliputi variabel tradisional (usia, pendapatan, status pekerjaan), data perilaku (durasi penggunaan nomor telepon seluler, riwayat pembayaran), serta informasi pinjaman (jumlah pinjaman, jangka waktu, dan status gagal bayar). Variabel target menunjukkan status pinjaman, yaitu 1 untuk gagal bayar (default) dan 0 untuk pembayaran lancar (repayment). Variabel utama yang menjadi fokus penelitian adalah durasi penggunaan nomor telepon seluler (mobile phone tenure), yang diukur dalam satuan tahun dan merepresentasikan lamanya penggunaan nomor telepon, yang diasumsikan berkorelasi dengan stabilitas perilaku dan tanggung jawab finansial. Langkah-langkah penelitian terdapat pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

Proses analisis data menggunakan pendekatan Exploratory Data Analysis (EDA), yang memanfaatkan statistik deskriptif seperti rata-rata, median, dan modus, serta teknik visualisasi data (Titus, 2021). Analisis dilakukan melalui pendekatan univariat, bivariat, dan multivariat (Shrivastava, 2024) untuk mengidentifikasi pola, distribusi, dan hubungan antarvariabel. Visualisasi data dilakukan menggunakan Python dengan bantuan pustaka Seaborn untuk menghasilkan boxplot, histogram, dan grafik korelasi (Arisandy et al., 2023). Seluruh proses pengolahan dan analisis data dijalankan pada platform Google Colaboratory (Colab), yaitu lingkungan Jupyter Notebook berbasis komputasi awan yang terintegrasi dengan Google Drive.

3. HASIL DAN PEMBAHASAN

3.1. Importasi *Libraries* dan *Dataset*

Tahap awal analisis dilakukan dengan mengimpor dataset berformat CSV ke dalam lingkungan pemrograman Python. Proses ini menggunakan Pandas untuk membaca dan mengelola data tabular, NumPy untuk pengolahan data numerik berbasis vektor dan matriks, serta SciPy untuk komputasi ilmiah. Visualisasi data dilakukan menggunakan Matplotlib dan Seaborn yang memungkinkan penyajian grafik statistik secara informatif. Kombinasi pustaka tersebut mendukung proses eksplorasi data secara sistematis melalui analisis numerik dan visualisasi pola data.

3.2. Identifikasi Tipe Data

Identifikasi tipe data merupakan langkah penting dalam analisis karena setiap metode statistik hanya dapat diterapkan pada jenis data tertentu. Kesalahan dalam mengidentifikasi tipe data dapat menghasilkan interpretasi yang tidak akurat. Secara umum, dataset ini terdiri dari tiga kategori tipe data, yaitu *object*, *int64*, dan *float64*.

1) *Object* (Data Kategorikal),

Tipe *object* merepresentasikan variabel kategorikal, seperti jenis pinjaman, jenis kelamin, kepemilikan kendaraan, tingkat pendidikan, status keluarga, pekerjaan, hari pengajuan pinjaman, dan jenis perusahaan tempat klien bekerja. Variabel ini

dianalisis menggunakan distribusi frekuensi dan perbandingan proporsi.

2) *Int64* (Data Numerik Diskrit)

Tipe *int64* mencakup variabel numerik berbentuk bilangan bulat, seperti usia (hari), lama bekerja (hari), hari publikasi identitas, kepemilikan telepon rumah, penilaian wilayah oleh pemberi pinjaman, estimasi jam pengajuan pinjaman, serta jumlah angsuran. Variabel ini dianalisis menggunakan statistik deskriptif dan distribusi numerik.

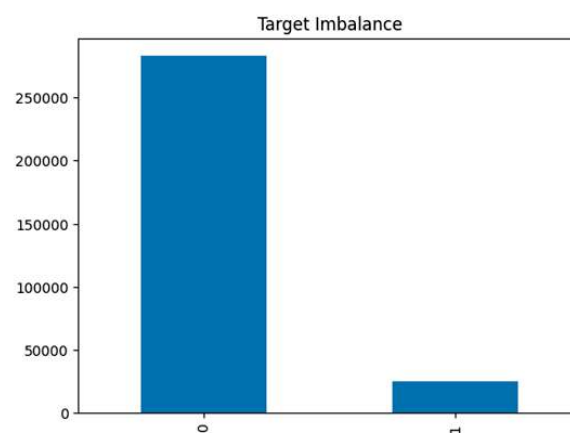
3) *Float64* (Data Numerik Kontinu)

Tipe *float64* menunjukkan variabel numerik kontinu dengan nilai desimal, seperti pendapatan klien, jumlah kredit, harga barang yang dibiayai, lama registrasi, frekuensi gagal bayar di lingkungan sekitar, lama penggunaan nomor telepon, permintaan informasi biro kredit, serta berbagai rasio keuangan seperti rasio kredit terhadap pendapatan dan rasio angsuran terhadap pendapatan.

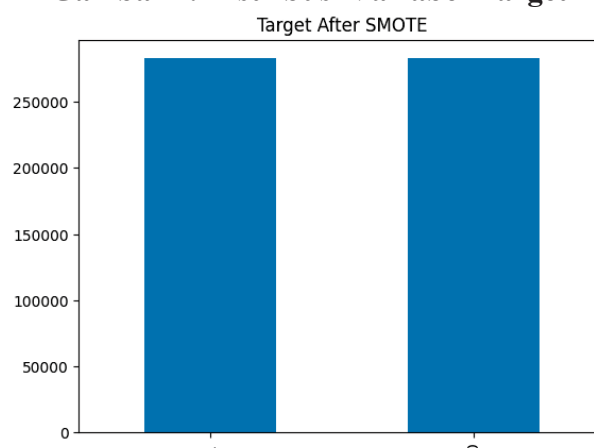
3.3. Handling Imbalance Data

Dataset dunia nyata umumnya menghadapi permasalahan class imbalance, yaitu kondisi ketika distribusi jumlah observasi antar kelas tidak seimbang (Gambar 2). Ketidakseimbangan ini dapat menyebabkan algoritma pembelajaran mesin cenderung bias terhadap kelas mayoritas. Pada dataset penelitian ini, jumlah pinjaman yang berhasil dilunasi tepat waktu jauh lebih besar dibandingkan pinjaman yang mengalami gagal bayar.

Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE). Teknik ini menghasilkan sampel sintetis pada kelas minoritas dengan memanfaatkan karakteristik data yang sudah ada sehingga distribusi kelas menjadi lebih seimbang. Melalui pendekatan ini, dihasilkan 282.682 sampel sintetis dari 24.824 data kelas minoritas, sehingga tercapai rasio 1:1 antara kelas mayoritas dan minoritas (Gambar 3). Proses penyeimbangan ini memungkinkan model pembelajaran mesin mengidentifikasi pola gagal bayar secara lebih proporsional.



Gambar 2. Distribusi Variabel Target



Gambar 3. Distribusi Variabel Target Setelah SMOTE

3.4. Identifikasi dan Penanganan *Missing Value*

Langkah selanjutnya adalah mengidentifikasi jumlah dan proporsi nilai hilang pada setiap kolom. Dataset yang digunakan terdiri dari 30 variabel, di mana sembilan di antaranya, seperti *AMT_ANNUITY*, memiliki nilai hilang. Dalam pengembangan model machine learning, penanganan nilai hilang menjadi tahap yang krusial dan umumnya dilakukan melalui teknik imputasi.

Meskipun alternatif lain adalah menghapus kolom dengan proporsi nilai hilang yang tinggi, pendekatan tersebut berisiko menghilangkan informasi yang berpotensi penting bagi model. Nilai prediktif suatu variabel tidak selalu dapat ditentukan hanya berdasarkan persentase missing value sebelum dilakukan analisis lebih lanjut. Oleh karena itu, penelitian ini mempertahankan seluruh variabel dan menerapkan strategi imputasi yang sesuai, sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Variable dengan missing value

<i>Variable</i>	<i>Missing Value</i>	<i>Duplicates</i>
SK_ID_CURR	0	0
TARGET	0	0
NAME_CONTRACT_TYPE	0	0
CODE_GENDER	0	0
FLAG_OWN_CAR	0	0
AMT_INCOME_TOTAL	0	0
AMT_CREDIT	0	0
AMT_ANNUITY	11	0
AMT_GOODS_PRICE	277	0
NAME_EDUCATION_TYPE	0	0
NAME_FAMILY_STATUS	0	0
REGION_POPULATION_RELATIVE	0	0
DAYS_BIRTH	0	0
DAYS_EMPLOYED	0	0
DAYS_REGISTRATION	0	0
DAYS_ID_PUBLISH	0	0
OWN_CAR_AGE	202929	0
FLAG_WORK_PHONE	0	0
OCCUPATION_TYPE	96391	0
REGION_RATING_CLIENT	0	0
REGION_RATING_CLIENT_W_CITY	0	0
WEEDAY_APPR_PROCESS_START	0	0
HOUR_APPR_PROCESS_START	0	0
ORGANIZATION_TYPE	0	0
OBS_30_CNT_SOCIAL_CIRCLE	1021	0
DEF_30_CNT_SOCIAL_CIRCLE	1021	0
DAYS_LAST_PHONE_CHANGE	1	0
AMT_REQ_CREDIT_BUREAU_QRT	41519	0
AMT_REQ_CREDIT_BUREAU_YEAR	41519	0

Dataset ini mengandung nilai hilang (*missing values*) yang dapat muncul akibat berbagai faktor sebagaimana telah dijelaskan

sebelumnya. Identifikasi nilai hilang dilakukan menggunakan fungsi *isnull()* pada pustaka Pandas, yang memungkinkan pendeteksian jumlah dan distribusi nilai yang tidak terisi pada setiap variabel.

Tabel 2 menyajikan hasil akhir dataset setelah dilakukan proses penanganan nilai hilang pada masing-masing variabel. Proses ini memastikan bahwa data yang digunakan dalam analisis selanjutnya telah memenuhi kelengkapan yang diperlukan untuk pengembangan model dan evaluasi risiko kredit secara lebih akurat.

Tabel 1. Dataset setelah dilakukan penanganan dengan missing value

<i>Variable</i>	<i>Missing Value</i>	<i>Duplicates</i>
SK_ID_CURR	0	0
TARGET	0	0
NAME_CONTRACT_TYPE	0	0
CODE_GENDER	0	0
FLAG_OWN_CAR	0	0
AMT_INCOME_TOTAL	0	0
AMT_CREDIT	0	0
AMT_ANNUITY	0	0
AMT_GOODS_PRICE	0	0
NAME_EDUCATION_TYPE	0	0
NAME_FAMILY_STATUS	0	0
REGION_POPULATION_RELATIVE	0	0
DAYS_BIRTH	0	0
DAYS_EMPLOYED	0	0
DAYS_REGISTRATION	0	0
DAYS_ID_PUBLISH	0	0
OWN_CAR_AGE	0	0
FLAG_WORK_PHONE	0	0
OCCUPATION_TYPE	0	0
REGION_RATING_CLIENT	0	0
REGION_RATING_CLIENT_W_CITY	0	0
WEEDAY_APPR_PROCESS_START	0	0
HOUR_APPR_PROCESS_START	0	0

<i>Variable</i>	<i>Missing Value</i>	<i>Duplicates</i>
<i>ORGANIZATION_T YPE</i>	0	0
<i>OBS_30_CNT_SOCI AL_CIRCLE</i>	0	0
<i>DEF_30_CNT_SOCI AL_CIRCLE</i>	0	0
<i>DAYS_LAST_PHON E_CHANGE</i>	0	0
<i>AMT_REQ_CREDIT BUREU_QRT</i>	0	0
<i>AMT_REQ_CREDIT BUREU_YEAR</i>	0	0

3.5. Identifikasi dan Koreksi Ketidakkonsistenan Data

Dalam penelitian ini, metode data generalization diterapkan untuk mengoreksi ketidakkonsistenan data. Metode ini dilakukan dengan menyederhanakan atau menyesuaikan data ke dalam kategori yang lebih tinggi berdasarkan struktur hierarkis yang relevan.

Pada variabel kode jenis kelamin (*Gender*), ditemukan empat data dengan label XNA pada Gambar 4. Untuk menjaga konsistensi dan standarisasi data, label tersebut digantikan dengan kategori yang paling dominan, yaitu "F" (*Female*). Proses ini bertujuan untuk memastikan keseragaman kategori sebelum dilakukan analisis lebih lanjut dalam pemodelan risiko kredit.

```
app_train['CODE_GENDER'].value_counts()
F      202448
M      105059
XNA         4
Name: CODE_GENDER, dtype: int64

# Correct inconsistencies
app_train['CODE_GENDER'] = app_train['CODE_GENDER'].replace('XNA', 'F')

app_train['CODE_GENDER'].value_counts()
F      202452
M      105059
Name: CODE_GENDER, dtype: int64
```

Gambar 4. Proses Koreksi Ketidakkonsistenan pada Kode Jenis Kelamin

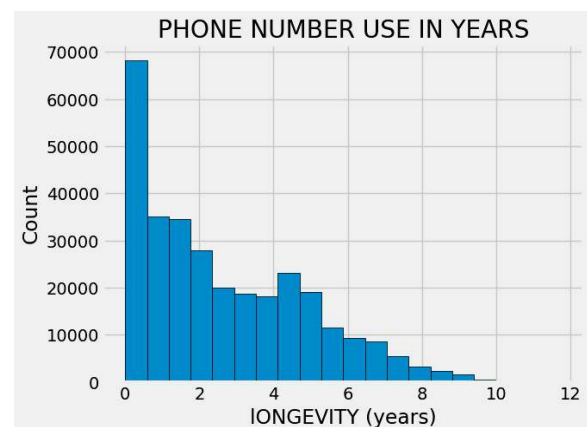
3.6. Exploratory Data Analysis (EDA)

Pendekatan Exploratory Data Analysis (EDA) memungkinkan peneliti untuk mengidentifikasi dan menganalisis informasi-informasi kunci yang terdapat dalam dataset secara sistematis sebelum tahap pemodelan lebih lanjut.

3.6.1 Uni variate analysis

Analisis univariat merupakan salah satu metode paling mendasar dalam deskripsi data. Tujuan utama analisis univariat adalah untuk menggambarkan karakteristik suatu variabel secara independen, tanpa mengeksplorasi hubungan dengan variabel lainnya.

Gambar 5 menampilkan histogram yang menunjukkan frekuensi (jumlah) klien berdasarkan durasi penggunaan nomor telepon dalam satuan tahun. Sumbu horizontal (x-axis) merepresentasikan lamanya penggunaan nomor telepon, yang dihitung dengan membagi nilai negatif variabel Days client changed phone number dalam tahun. Sumbu vertikal (y-axis) menunjukkan jumlah klien dalam setiap interval durasi penggunaan nomor telepon.



Gambar 5. Penggunaan nomer telepon dalam Tahun (Histogram)

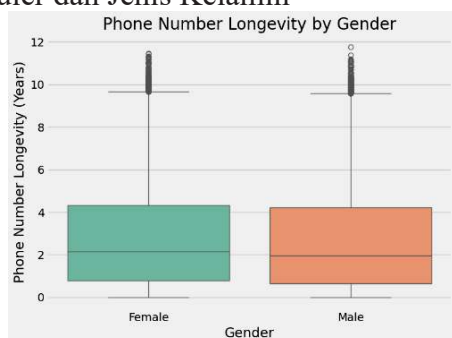
Distribusi durasi penggunaan nomor telepon seluruh klien ditampilkan pada Gambar 5. Grafik tersebut menunjukkan bahwa dataset mencakup rentang durasi penggunaan nomor telepon antara 1 hingga 10 tahun. Histogram memperlihatkan puncak pada beberapa interval tertentu, yang menandakan adanya konsentrasi jumlah klien yang lebih tinggi pada kelompok lama penggunaan nomor telepon.

Durasi penggunaan nomor telepon yang paling tinggi ditemukan adalah kurang dari 2 tahun. Temuan ini menunjukkan bahwa sebagian besar pemohon memiliki nomor telepon yang relatif baru, yang dapat mengindikasikan tingkat stabilitas identitas digital yang lebih rendah. Pola distribusi ini menjadi penting dalam konteks evaluasi risiko kredit, karena stabilitas penggunaan nomor telepon berpotensi berkaitan dengan konsistensi perilaku finansial peminjam.

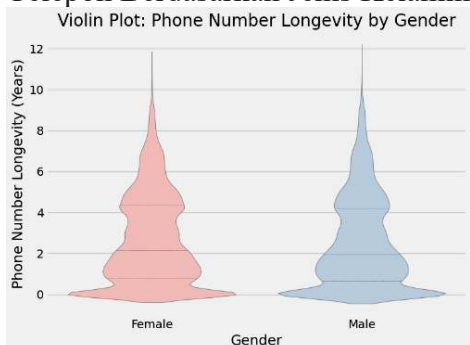
3.6.2 Bivariate analysis

Analisis bivariat merupakan salah satu metode statistik deskriptif yang digunakan untuk menganalisis hubungan antara dua variabel secara simultan. Pendekatan ini bertujuan untuk mengidentifikasi pola keterkaitan atau potensi korelasi antara kedua variabel yang dianalisis.

1) Durasi Penggunaan Nomor Telepon Seluler dan Jenis Kelamin



Gambar 6. Boxplot Durasi Penggunaan Nomor Telepon Berdasarkan Jenis Kelamin

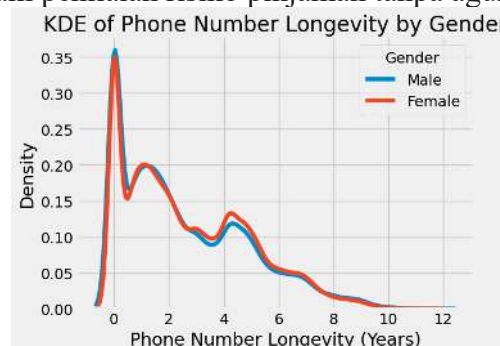


Gambar 7. Violin Plot Durasi Penggunaan Nomor Telepon Berdasarkan Jenis Kelamin

Gambar 6 dan Gambar 7 membandingkan durasi penggunaan nomor telepon berdasarkan jenis kelamin melalui boxplot dan violin plot. Meskipun nilai median durasi penggunaan nomor telepon tampak relatif serupa antara laki-laki dan perempuan, terdapat perbedaan pada rentang interkuartil (yang direpresentasikan oleh kotak pada boxplot) maupun pada bentuk distribusi keseluruhan yang terlihat pada violin plot. Perbedaan ini mengindikasikan adanya variasi dalam distribusi durasi penggunaan nomor telepon antara klien laki-laki dan perempuan. Perbedaan ini mengindikasikan variasi pola penggunaan nomor telepon antar kelompok gender.

Dalam konteks risiko kredit, variasi tersebut dapat mencerminkan perbedaan stabilitas identitas digital. Durasi penggunaan

nomor telepon yang lebih stabil berpotensi menunjukkan konsistensi perilaku dan risiko kredit yang lebih rendah, sedangkan pergantian nomor yang lebih sering dapat menjadi indikator awal ketidakstabilan yang relevan dalam penilaian risiko pinjaman tanpa agunan.



Gambar 8. Kernel Density Estimate (KDE) Durasi Penggunaan Nomor Telepon Berdasarkan Jenis Kelamin

Grafik KDE memberikan gambaran mengenai bentuk dan konsentrasi distribusi lama penggunaan nomor telepon pada masing-masing kelompok gender. Perbedaan posisi atau tinggi puncak kurva menunjukkan variasi konsentrasi nilai pada setiap kelompok. Puncak yang lebih tajam menandakan data terkonsentrasi pada rentang tertentu, sedangkan kurva yang lebih datar menunjukkan distribusi yang lebih menyebar. Tingkat tumpang tindih (overlap) antar kurva juga menjadi indikator penting: overlap yang besar mengindikasikan distribusi yang relatif serupa antara laki-laki dan perempuan, sementara kurva yang lebih terpisah menunjukkan adanya perbedaan pola penggunaan nomor telepon.

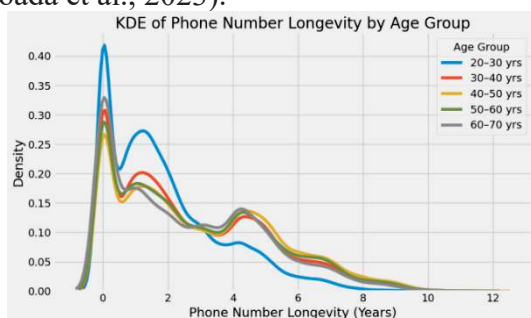
Secara keseluruhan, box plot, violin plot, dan KDE plot memberikan perspektif yang saling melengkapi dalam memahami hubungan antara gender dan lama penggunaan nomor telepon, masing-masing menampilkan tendensi sentral, kepadatan distribusi, serta bentuk probabilitas data.

Pada Gambar 8. tahun pertama, klien laki-laki tampak sedikit lebih sering melakukan pergantian nomor telepon. Namun demikian, perbedaan ini relatif kecil dan tidak menunjukkan signifikansi yang kuat secara statistik. Secara umum, distribusi data menunjukkan bahwa perempuan memiliki kecenderungan perubahan nomor telepon yang lebih sering pada beberapa periode tertentu.

Temuan ini memberikan indikasi awal bahwa karakteristik demografis seperti gender berpotensi memengaruhi stabilitas penggunaan nomor telepon. Stabilitas nomor telepon dapat mencerminkan tingkat mobilitas atau stabilitas sosial-ekonomi klien, yang dalam beberapa penelitian sering dikaitkan dengan perilaku keuangan dan risiko kredit (Muslim et al., 2025). Oleh karena itu, variabel ini berpotensi menjadi salah satu fitur yang relevan dalam pengembangan model prediksi risiko kredit berbasis machine learning pada tahap analisis selanjutnya.

2) Lama Penggunaan Nomor Telepon dan Usia Klien

Untuk memperoleh gambaran yang lebih jelas mengenai distribusi usia klien, data usia diklasifikasikan ke dalam beberapa kategori generasi menggunakan algoritma bins. Kategori tersebut meliputi Gen Z (18–25 tahun), Milenial (26–35 tahun), Gen X (36–45 tahun), Baby Boomer (46–55 tahun), serta kelompok usia di atas 56 tahun sebagaimana diklasifikasikan dalam penelitian sebelumnya (Boada et al., 2023).



Gambar 9. Durasi Penggunaan Nomor Telepon Seluler Berdasarkan Kelompok Usia

Dengan membandingkan puncak dan bentuk dari garis-garis berwarna yang berbeda pada Gambar 9, dapat diamati bagaimana distribusi durasi penggunaan nomor telepon bervariasi pada setiap kelompok usia. Hasil distribusi menunjukkan perbedaan yang cukup signifikan antar kelompok usia. Kelompok usia Gen Z cenderung memiliki puncak distribusi pada nilai durasi yang lebih rendah, sedangkan kelompok usia yang lebih tua menunjukkan puncak distribusi pada nilai durasi yang lebih tinggi.

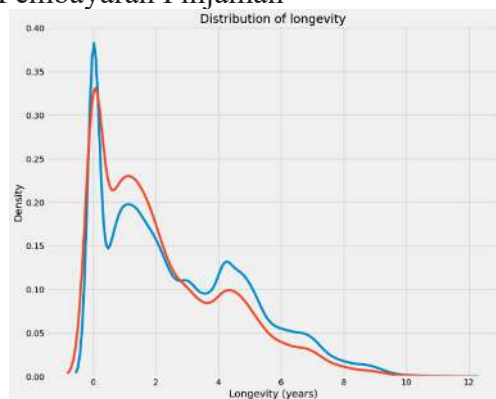
Perbandingan rinci terhadap puncak dan distribusi durasi penggunaan nomor telepon pada berbagai kelompok usia menunjukkan pola yang berbeda yang mencerminkan variasi

dalam penggunaan dan retensi nomor telepon seluler dari waktu ke waktu.

Hasil analisis menunjukkan bahwa distribusi durasi penggunaan nomor telepon berbeda secara signifikan antar generasi, yang terlihat dari perbedaan puncak distribusi pada masing-masing kelompok usia. Secara khusus, kelompok Generasi Z menunjukkan puncak pada durasi yang lebih rendah, yang mengindikasikan kecenderungan mempertahankan nomor telepon dalam waktu yang relatif lebih singkat. Hal ini kemungkinan dipengaruhi oleh pesatnya perkembangan teknologi serta kecenderungan generasi muda untuk lebih sering mengganti operator atau layanan kartu SIM (Rahma et al., 2024; Shorey et al., 2024). Pola penggunaan telepon seluler pada kelompok ini juga berkaitan dengan preferensi terhadap teknologi baru, perubahan preferensi komunikasi, serta tingginya aktivitas yang berkaitan dengan media sosial dan pengelolaan akun digital.

Sebaliknya, kelompok usia yang lebih tua menunjukkan puncak distribusi pada durasi yang lebih tinggi, yang menandakan bahwa mereka cenderung mempertahankan nomor telepon dalam jangka waktu yang lebih lama. Kondisi ini dapat dikaitkan dengan kebutuhan komunikasi yang lebih stabil, seperti penggunaan kontrak jangka panjang, konsistensi dalam jaringan sosial maupun profesional, serta tingkat perubahan teknologi yang relatif lebih rendah (Vulpe & Ilinca, 2020). Perbedaan pola ini menunjukkan bahwa faktor generasi berpengaruh terhadap perilaku retensi nomor telepon. Secara umum, semakin bertambah usia seseorang, semakin besar kecenderungannya untuk mempertahankan nomor telepon dalam periode yang lebih panjang. Temuan ini memberikan indikasi bahwa faktor usia, adopsi teknologi, dan perilaku komunikasi dapat menjadi variabel penting dalam memahami durasi penggunaan nomor telepon seluler.

3) Durasi Penggunaan Nomor Telepon dan Pembayaran Pinjaman



Gambar 10. Distribusi Durasi Penggunaan Nomor Telepon Berdasarkan Status Pembayaran Pinjaman

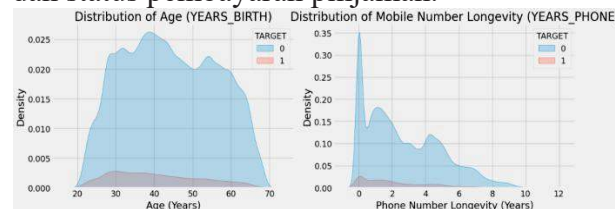
Gambar 10 menampilkan distribusi durasi penggunaan nomor telepon (dalam tahun) untuk dua kelompok klien: yang berhasil melunasi pinjaman dan yang mengalami gagal bayar. Kurva merah (default) menunjukkan puncak lebih tinggi pada durasi penggunaan nomor telepon yang lebih pendek. Hal ini mengindikasikan bahwa klien yang baru atau lebih sering mengganti nomor telepon cenderung memiliki probabilitas gagal bayar yang lebih tinggi.

Sebaliknya, kurva biru (repaid) terlihat lebih menyebar ke arah nilai durasi yang lebih tinggi, yang menunjukkan bahwa klien yang mempertahankan nomor telepon yang sama dalam jangka waktu lebih lama cenderung memiliki tingkat stabilitas yang lebih baik dan kemungkinan gagal bayar yang lebih rendah. Dengan demikian, variabel durasi penggunaan nomor telepon berpotensi menjadi indikator penting dalam model penilaian risiko kredit, karena dapat merefleksikan stabilitas perilaku dan konsistensi identitas digital peminjam.

Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa data perilaku berbasis telepon seluler dapat meningkatkan kemampuan prediksi risiko kredit (Arkhipova & Karminsky, 2023), khususnya pada peminjam dengan riwayat kredit terbatas. Data alternatif seperti pola penggunaan perangkat dan stabilitas identitas digital terbukti dapat melengkapi metode credit scoring tradisional yang sering kali tidak mampu menangkap profil risiko peminjam tanpa riwayat kredit formal.

3.6.3 Multivariate Analysis

Analisis ini bertujuan untuk memahami bagaimana berbagai variabel dalam dataset saling berinteraksi serta mengidentifikasi hubungan antara variabel yang memiliki lebih dari dua kemungkinan tingkat. Gambar 10 menunjukkan korelasi antara durasi penggunaan nomor telepon pelanggan, usia, dan status pembayaran pinjaman.



Gambar 11. Distribusi Durasi Penggunaan Nomor Telepon, Usia Klien, dan Status Pembayaran Pinjaman

Berdasarkan dataset yang dianalisis, Gambar 11 menunjukkan distribusi usia klien dalam beberapa rentang tahun. Dua kurva pada grafik tersebut, kurva biru untuk target = 0 (pinjaman dibayar) dan kurva abu-abu untuk target = 1 (gagal bayar), merepresentasikan kepadatan distribusi klien pada berbagai kelompok usia untuk masing-masing status pinjaman. Kurva abu-abu (gagal bayar) tampak lebih tinggi pada kelompok usia yang lebih muda, sekitar 20–35 tahun, sementara kurva biru (pinjaman dibayar) relatif lebih tinggi pada kelompok usia yang lebih tua.

Hasil Exploratory Data Analysis (EDA) pada dataset ini menunjukkan bahwa klien yang lebih muda, terutama yang memiliki durasi penggunaan nomor telepon yang lebih pendek, cenderung memiliki kemungkinan gagal bayar yang lebih tinggi dibandingkan klien yang lebih tua. Dengan demikian, usia dapat menjadi salah satu faktor penting dalam memprediksi kemampuan pembayaran pinjaman. Temuan ini juga sejalan dengan beberapa penelitian sebelumnya yang menunjukkan bahwa karakteristik demografis, termasuk usia, memiliki hubungan dengan risiko kredit dan perilaku pembayaran pinjaman (Lin & Wang, 2025; Maivalinda et al., 2023).

3.7 Matrix Visualization

Visualisasi Matriks (Matrix Visualization/MV) merupakan teknik yang digunakan untuk memvisualisasikan dan mengelompokkan data simbolik menggunakan

data konseptual berbasis interval, yang merupakan salah satu jenis data simbolik yang paling umum digunakan baik dalam penelitian maupun praktik (Kao et al., 2014; Prasetya et al., 2025).

Tabel 3. Interpretasi Koefisien Korelasi (Hinkle et al., 2003; Mukaka, 2012)

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to -1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	Almost no correlation

Tabel 3 memberikan interpretasi terhadap nilai koefisien yang diperoleh dari analisis korelasi Pearson sederhana, yang juga dikenal sebagai Pearson Product Moment. Nilai koefisien korelasi (r) berada pada rentang -1 hingga 1. Nilai yang semakin mendekati 1 atau -1 menunjukkan bahwa hubungan antara dua variabel semakin kuat, sedangkan nilai yang semakin mendekati 0 menunjukkan bahwa hubungan antarvariabel semakin lemah atau hampir tidak ada.

```

Most Positive Correlations:
CREDIT_TERM                0.012705
ANNUITY_INCOME_PERCENT     0.014281
AMT_REQ_CREDIT_BUREAU_YEAR 0.018166
OWN_CAR_AGE                 0.020870
FLAG_WORK_PHONE            0.028534
DEF_30_CNT_SOCIAL_CIRCLE   0.032224
DAYS_REGISTRATION          0.041981
DAYS_EMPLOYED_PERCENT      0.042204
DAYS_ID_PUBLISH            0.051462
NAME_EDUCATION_TYPE        0.054689
CODE_GENDER                 0.054726
DAYS_LAST_PHONE_CHANGE     0.055209
REGION_RATING_CLIENT_W_CITY 0.060894
DAYS_EMPLOYED              0.074951
TARGET                      1.000000
Name: TARGET, dtype: float64

Most Negative Correlations:
DAYS_BIRTH                  -0.078232
DAYS_EMPLOYED_ANOM         -0.045985
AMT_GOODS_PRICE            -0.039622
REGION_POPULATION_RELATIVE -0.037220
NAME_CONTRACT_TYPE         -0.030894
ORGANIZATION_TYPE         -0.030756
AMT_CREDIT                 -0.030365
HOUR_APPR_PROCESS_START    -0.024169
AMT_INCOME_TOTAL          -0.021998
FLAG_OWN_CAR               -0.021838
OCCUPATION_TYPE            -0.017441
AMT_ANNUITY                -0.012800
CREDIT_INCOME_PERCENT     -0.007716
NAME_FAMILY_STATUS        -0.004125
AMT_REQ_CREDIT_BUREAU_QRT -0.001839
Name: TARGET, dtype: float64
    
```

Gambar 12. Korelasi antara Variabel dan Status Pembayaran Pinjaman (Target)

Tabel 4. Nilai Korelasi Antarvariabel

index	Var1	Var2	Correlation
340	DAYS_EMPLOYED	DAYS_BIRTH	0.62
880	DAYS_EMPLOYED_PERCENT	DAYS_BIRTH	0.6
610	YPE	DAYS_BIRTH	0.4

370	DAYS_REGISTRATION	DAYS_BIRTH	0.34
490	OCCUPATION_TYPE	DAYS_BIRTH	0.33
400	DAYS_ID_PUBLISH	DAYS_BIRTH	0.28
790	CREDIT_INCOME_PERCENT	DAYS_BIRTH	0.14
308	DAYS_BIRTH	NAME_FAMILY_STATUS	0.14
307	DAYS_BIRTH	NAME_EDUCATION_TYPE	0.13
460	FLAG_WORK_PHONE	DAYS_BIRTH	0.13
850	CREDIT_TERM	DAYS_BIRTH	0.12
301	DAYS_BIRTH	CODE_GENDER	0.11
700	DAYS_LAST_PHONE_CHANGE	DAYS_BIRTH	0.1
306	DAYS_BIRTH	AMT_GOODS_PRICE	0.09
304	DAYS_BIRTH	AMT_CREDIT	0.09
300	DAYS_BIRTH	NAME_CONTRACT_TYPE	0.08
580	HOUR_APPR_PROCESS_START	DAYS_BIRTH	0.07
302	CESS_START	DAYS_BIRTH	0.07
820	FLAG_OWN_CAR	DAYS_BIRTH	0.07
760	AR	DAYS_BIRTH	0.06
	AMT_REQ_CREDIT_BUREAU_YEAR	DAYS_BIRTH	0.06

Gambar 12 menunjukkan bahwa beberapa variabel penting, yaitu lama bekerja (DAYS_EMPLOYED), penilaian pemberi pinjaman terhadap wilayah tempat tinggal klien (REGION_RATING_CLIENT_W_CITY), serta durasi penggunaan nomor telepon (DAYS_LAST_PHONE_CHANGE), memiliki hubungan positif dengan kemungkinan pelunasan pinjaman. Hasil visualisasi menunjukkan bahwa klien dengan masa kerja yang lebih lama, stabilitas penggunaan nomor telepon yang lebih tinggi, serta yang berasal dari wilayah dengan penilaian kredit yang lebih baik cenderung memiliki probabilitas yang lebih tinggi untuk melunasi pinjaman mereka. Selain itu, Tabel 4 menunjukkan bahwa korelasi antara lama bekerja dan usia klien sebesar $r = 0,62$, yang termasuk dalam kategori korelasi sedang, sehingga menunjukkan bahwa stabilitas pekerjaan cenderung meningkat seiring dengan bertambahnya usia.

Namun demikian, dari perspektif pemodelan risiko kredit, hubungan antarvariabel dalam dataset ini tidak sepenuhnya bersifat linear. Meskipun korelasi Pearson antara durasi penggunaan nomor telepon dan status gagal bayar terlihat relatif

lemah, hal tersebut tidak serta-merta menunjukkan bahwa variabel tersebut tidak relevan dalam pemodelan. Dalam praktik credit risk modeling, variabel dengan korelasi marginal yang kecil sering kali tetap memiliki nilai prediktif yang signifikan ketika dikombinasikan dengan variabel lain yang lebih kuat secara struktural, seperti usia dan stabilitas pekerjaan.

Temuan ini menunjukkan adanya mekanisme hubungan tidak langsung (mediating relationship) antara beberapa variabel utama. Usia klien memiliki hubungan positif dengan stabilitas pekerjaan, dan stabilitas pekerjaan tersebut berkontribusi pada penurunan risiko gagal bayar. Pada saat yang sama, usia juga menunjukkan hubungan negatif dengan frekuensi pergantian nomor telepon, yang menunjukkan bahwa klien yang lebih muda cenderung lebih sering mengganti nomor telepon dibandingkan klien yang lebih tua. Dengan demikian, durasi penggunaan nomor telepon dapat diinterpretasikan sebagai indikator perilaku yang mencerminkan stabilitas ekonomi dan sosial, bukan semata-mata sebagai faktor langsung yang memengaruhi kemampuan pembayaran pinjaman.

Lebih lanjut, pola ini mengindikasikan bahwa lama penggunaan nomor telepon dapat mencerminkan unsur kontinuitas identitas digital, stabilitas jaringan sosial, serta komitmen terhadap layanan komunikasi jangka panjang. Peminjam yang lebih muda, khususnya dari kelompok Generasi Z, cenderung menghadapi tingkat volatilitas yang lebih tinggi dalam pasar tenaga kerja, termasuk kontrak kerja sementara, mobilitas pekerjaan yang lebih tinggi, serta fase transisi kehidupan yang lebih dinamis. Kondisi ini tidak hanya memengaruhi stabilitas pekerjaan tetapi juga meningkatkan mobilitas perilaku, termasuk dalam hal pergantian nomor telepon. Kombinasi faktor-faktor tersebut pada akhirnya dapat meningkatkan probabilitas gagal bayar pada kelompok usia yang lebih muda.

Implikasi dari temuan ini adalah bahwa durasi penggunaan nomor telepon mungkin tidak menunjukkan hubungan yang kuat secara langsung terhadap gagal bayar, namun variabel ini berpotensi memberikan nilai prediktif yang signifikan dalam model machine learning yang

bersifat non-linear atau berbasis interaksi variabel. Oleh karena itu, integrasi indikator perilaku alternatif, seperti stabilitas penggunaan nomor telepon, menjadi penting dalam pengembangan model credit scoring multivariat, terutama untuk meningkatkan kemampuan model dalam mengidentifikasi risiko kredit pada peminjam dengan riwayat kredit terbatas.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil Exploratory Data Analysis (EDA) ditemukan bahwa peminjam dengan usia lebih muda dan durasi penggunaan nomor telepon yang lebih pendek cenderung memiliki probabilitas gagal bayar yang lebih tinggi dibandingkan peminjam yang lebih tua dengan durasi penggunaan nomor telepon yang lebih lama. Pola ini menunjukkan bahwa durasi penggunaan nomor telepon dapat mencerminkan tingkat stabilitas identitas dan kontinuitas perilaku peminjam. Meskipun analisis korelasi menunjukkan bahwa hubungan linier antara durasi penggunaan nomor telepon dan status pembayaran pinjaman relatif lemah, variabel ini tetap memberikan informasi yang relevan dalam konteks analisis risiko kredit. Hasil eksplorasi menunjukkan bahwa variabel tersebut menjadi lebih informatif ketika dianalisis bersama karakteristik lain, khususnya usia dan stabilitas pekerjaan, yang secara bersama-sama membentuk pola risiko yang lebih jelas. Secara keseluruhan, temuan penelitian ini menunjukkan bahwa durasi penggunaan nomor telepon seluler memiliki potensi sebagai indikator data alternatif dalam analisis risiko kredit, terutama dalam ekosistem pinjaman digital yang sering menghadapi keterbatasan riwayat kredit formal. Dengan demikian, indikator telekomunikasi yang sederhana namun persisten dapat memberikan kontribusi tambahan dalam pengembangan sistem credit scoring berbasis data dan analisis machine learning yang lebih inklusif.

4.2. Saran

Penelitian selanjutnya disarankan untuk mengkaji apakah data yang berasal dari kartu identitas resmi pemerintah (ID card) memiliki hubungan dengan risiko gagal bayar pada pinjaman tanpa agunan. Saat ini, banyak

platform digital yang menyetujui pengajuan pinjaman hanya berdasarkan informasi identitas, sehingga masih belum jelas apakah praktik tersebut lebih didorong oleh strategi pemasaran atau benar-benar memiliki dasar analitis yang kuat dalam penilaian risiko kredit. Atribut yang terhubung dengan identitas resmi kemungkinan mengandung sinyal demografis atau geografis yang mencerminkan tingkat stabilitas individu, yang pada akhirnya dapat berkontribusi dalam memprediksi kemampuan pembayaran pinjaman. Oleh karena itu, pengujian empiris menggunakan kerangka machine learning yang dapat diinterpretasikan (interpretable machine learning) menjadi penting untuk mengevaluasi apakah variabel-variabel tersebut memberikan nilai prediktif tambahan dalam model penilaian kredit. Selain itu, penelitian di masa depan juga perlu memastikan bahwa penggunaan data identitas tersebut tetap memperhatikan prinsip keadilan, transparansi, dan tata kelola data yang bertanggung jawab, khususnya dalam pengembangan sistem credit scoring berbasis kecerdasan buatan.

5. DAFTAR PUSTAKA

- Adam, L., Sarana, J., Suyatno, B., Soekarni, M., Suryanto, J., Ermawati, T., Saptia, Y., Adityawati, S., Mychelisda, E., & Pamungkas, Y. (2025). Driving Financial Inclusion in Indonesia with Innovative Credit Scoring. *Journal of Risk and Financial Management*, *18*(8), 442.
- Agarwal, S., Alok, S., Ghosh, P., & Gupta, S. (2019). Financial Inclusion and Alternate Credit Scoring: Role of Big Data and Machine Learning in Fintech. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3507827>
- Arisandy, Y., Dasril, Y. Bin, Salahudin, S. N. Bin, Muslim, M. A., Adnan, A., & Wen, G. K. (2023). Buy Now Pay Later Services on Generation Z: Exploratory Data Analysis Using Machine Learning. *J. Theor. Appl. Inf. Technol*, *15*(11), 4194–4204.
- Arkhipova, N., & Karminsky, A. (2023). Demographic characteristics as determinants of retail customers' credit behavior. Evidence from Russian regions. *Procedia Computer Science*, *221*, 1091–1098.
- Auliya, S. F., Nugroho, L. E., & Setiawan, N. A. (2021). *A review on smartphone usage data for user identification and user profiling*. *6*(1), 25–34. <https://doi.org/10.21924/CST.6.1.2021.363>
- Björkegren, D., & Grissen, D. (2019). *Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment* (Vol. 34). World Bank, Washington, DC. <https://doi.org/10.1596/1813-9450-9074>
- Chen, A. J. Y., Even-Tov, O., Kang, J. K., & Wittenberg-Moerman, R. (2025). Digital Lending and Financial Well-Being: Through the Lens of Mobile Phone Data. *The Accounting Review*, *100*(4), 135–159. <https://doi.org/10.2308/TAR-2024-0046>
- Kao, C. H., Nakano, J., Shieh, S. H., Tien, Y. J., Wu, H. M., Yang, C. K., & Chen, C. H. (2014). Exploratory data analysis of interval-valued symbolic data with matrix visualization. *Computational Statistics and Data Analysis*, *79*, 14–29. <https://doi.org/10.1016/j.csda.2014.04.012>
- Lin, L., & Wang, Y. (2025). Shap stability in credit risk management: A case study in credit card default model. *Risks*, *13*(12), 238.
- Liu, Q., Zou, L., Yang, X., & Tang, J. (2019). Survival or die: a survival analysis on peer-to-peer lending platforms in China. *Accounting and Finance*, *59*, 2105–2131. <https://doi.org/10.1111/ACFI.12513>
- Maivalinda, M., Sulistianingsih, H., & Riski, T. R. (2023). Mengukur Prilaku Pinjaman Online Melalui Literasi Keuangan Digital, Preferensi Risiko dan Faktor Demografi Sebagai Variabel Moderasi. *Jurnal Ekonomi Dan Bisnis Dharma Andalas*, *25*(2), 561–572.
- Muslim, B. L. B. B., Hamdani, M., Anugrahani, I. S., Minarsi, A., Rianty, E., Purnamaningrum, T. K., Pracoyo, A., & Juansa, A. (2025). *Literasi Keuangan: Wawasan, Perilaku dan Strategi Mengelola Keuangan*. PT. Star Digital Publishing, Yogyakarta-Indonesia.
- Nainggolan, D. S., & Umaroh, R. (2023). Analisis Efektifitas Pinjaman Bank Dan Peer-To-Peer Lending. *Journal of Economic, Business and Engineering*

- (*JEBE*), 4(2), 252–260.
<https://doi.org/10.32500/jebe.v4i2.4443>
- Nwachukwu, G. (2024). Enhancing credit risk management through revalidation and accuracy in financial data: The impact of credit history assessment on procedural financing. *International Journal of Research Publication and Reviews*, 5(11), 631–644.
- Ots, H., Liiv, I., & Tur, D. (2020). Mobile Phone Usage Data for Credit Scoring. *ArXiv: Computers and Society*.
<https://arxiv.org/abs/2002.12616>
- Prasetya, D. A., Choiruddin, A., Rakshit, S., Sari, A. P., Wara, S. S. M., Pratama, A. R., & Nasrudin, M. (2025). *Machine Learning Praktis Dengan Python: Dari Data Mining hingga Neural Network*. Thalibul Ilmi Publishing & Education.
- Putri Nurhidayati. (2025). *Cukup Nomor HP Aktif, Easycash Bisa Kasih Pinjaman Rp 50 Juta, Segini Cicilannya*. Rbtv.Disway.Id.
<https://rbtv.disway.id/bisnis-investasi/read/96625/cukup-nomor-hp-aktif-easycash-bisa-kasih-pinjaman-rp-50-juta-segini-cicilannya>
- Rahma, K., Indallaila, Fatimah, E., Mubarak, S., & Cinta, N. (2024). *Analisis Pengguna Ponsel Terhadap Perilaku Generasi Z dalam Kehidupan Sehari-hari*.
<https://doi.org/10.30997/karimahtauhid.v3>
- i2.11858
- Razavi, R., & Elbahnasawy, N. G. (2025). Unlocking credit access: Using non-CDR mobile data to enhance credit scoring for financial inclusion. *Finance Research Letters*, 73, 106682.
<https://doi.org/https://doi.org/10.1016/j.frl.2024.106682>
- Shorey, S., Vyugina, D., Waechter, N., & Dolev, N. (2024). *Communication Preferences and Behaviors* (pp. 31–42). Emerald (MCB UP).
<https://doi.org/10.1108/978-1-83797-092-620241004>
- Shrivastava, A. (2024). Attrition Unveiled: Analyzing Trends and Strategies in Employee Turnover. *Journal of Management and Service Science (JMSS)*, 4(2), 1–9.
<https://doi.org/10.54060/a2zjournals.jmss.72>
- Titus, M. A. (2021). *Using Descriptive Statistics and Graphs* (pp. 79–102). Springer, Cham.
https://doi.org/10.1007/978-3-030-60831-6_6
- Vulpe, S.-N., & Ilinca, C. (2020). AGE AND USAGE OF TECHNOLOGY. A STRCUTURAL EQUATION MODEL BASED ON THE THEORY OF PLANNED BEHAVIOR. *Revista Universitară de Sociologie*, 16(2).