

Perbandingan Algoritma Machine Learning Untuk Prediksi Gagal Bayar Pinjaman Koperasi yang Optimal

Hilmi Aziz¹; Rianto²

1 Informatika, Universitas Siliwangi, Jl. Siliwangi No.24

² Jl. Raya, RT.4/RW.1, Meruya Sel., Kec. Kembangan, Jakarta, Daerah Khusus Ibukota Jakarta 11650

¹ 41520010181@student.mercubuana.ac.id, ² rianto@unsil.ac.id

Kata kunci:

accuracy, decision tree, default prediction, f1-score, k-nearest neighbors, knn, logistic regression, precision, random forest, recall

Abstract

Predicting loan repayment defaults is quite an important thing to do in a financial institution such as a Savings and Loans Cooperative. The aim is to minimize the occurrence of loan defaults by borrowers to cooperatives so that bankruptcy does not occur. In this study, the development of a predictive model was carried out using several popular machine learning algorithms, namely logistic regression, decision tree, random forest and k-nearest neighbors (KNN), then the four models were compared and evaluated in order to find out which model with the most effective algorithm. in predicting loan defaults in cooperatives. Program evaluation is carried out by metrics such as accuracy, precision, recall, and f1-score. The dataset itself is obtained from a loan list which includes attributes such as borrower profile, loan amount, number of installments, etc. This dataset is divided into training data and test data to train and evaluate the model. The results showed that the Random Forest algorithm model provided the best accuracy, reaching 89%, followed by the Decision Tree with the highest accuracy value, which reached 84%, and finally Logistic Regression and K-Nearest Neighbors with the same accuracy value, namely 81%. These four algorithms were chosen because they are well-known algorithms among other algorithms for financial predictions because of their ability to understand complex relationships, provide interpretable results, overcome overfitting problems, and consider the interrelationships between similar entities.

Pendahuluan

A. Latar Belakang

Model prediksi tersebut bisa dibuat dengan algoritma machine learning tradisional seperti Decision Tree, XGBoost, Logistic Regression, Random Forest, dan K-Nearest Neighbors. Ada juga beberapa algoritma Deep Learning seperti Recurrent Neural Network dan Convolutional Neural Network. Masing-masing algoritma tersebut memiliki kelebihan dan kekurangannya karena bisa membuat model prediksinya dengan karakter/caranya masing-masing. Namun untuk melakukan prediksi dengan gagal bayar ini, algoritma machine learning dirasa sudah cukup karena algoritma deep learning yang memiliki kemampuan mesin yang mumpuni dan kompleksitas pada suatu fitur sehingga membutuhkan data yang besar dan lebih kompleks lagi yang justru kurang cocok untuk dataset koperasi ini yang terhitung cukup kecil[2]. Hal ini juga ditemukan dalam penelitian analisis resiko kredit oleh Peter et al.(2018) yang menemukan bahwa model berbasis pohon lebih stabil daripada model berdasarkan jaringan saraf tiruan multilayer[3].

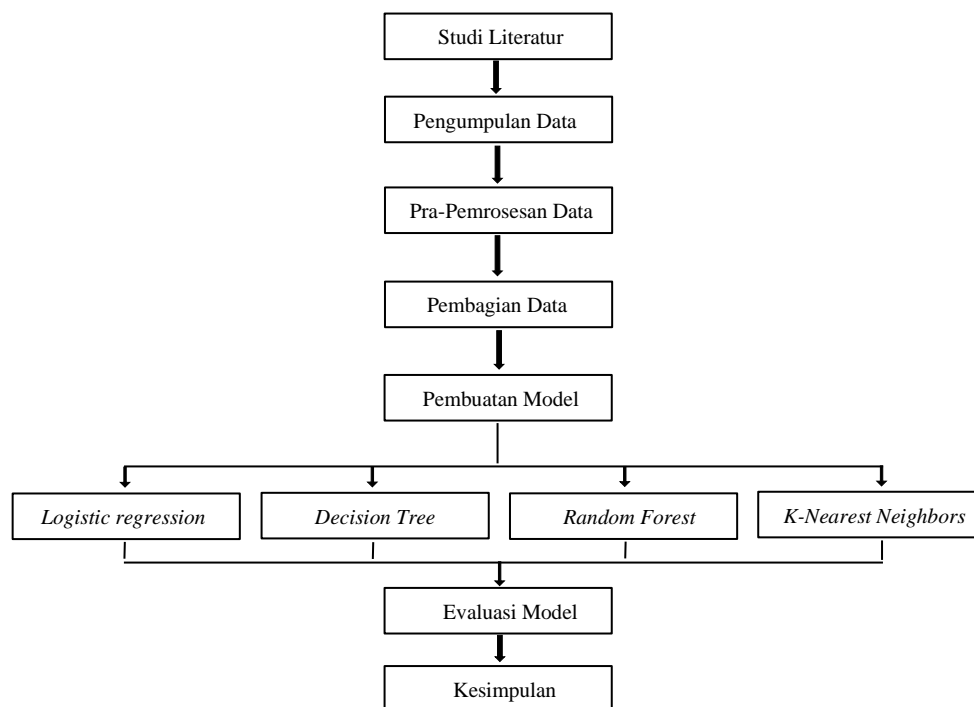
Kemampuan dari algoritma *machine learning* tradisional juga sudah cukup baik, seperti menurut sebuah studi yang dilakukan oleh Kadek et al.(2022), menggunakan beberapa algoritma pembelajaran mesin seperti *decision tree*, *random forest*, dan *logistic regression* dapat membantu memprediksi gagal bayar pinjaman mikro di asosiasi simpan pinjam di India. Kesimpulan dari hasil penelitian mereka menunjukkan bahwa algoritma *XGBoost* dapat memberikan kinerja prediksi yang cukup baik dibanding algoritma lainnya[4]. Hal ini juga terbukti dalam penelitian yang dilakukan oleh Lkhagvadorj et al.(2019) yang menunjukkan hasil dari model Deep Neural Networks dan model XGBoost mempunyai kinerja yang lebih baik dibandingkan dengan pendekatan pembelajaran mesin lainnya dalam hal AUC dan akurasi[5].

Namun hasil itu tidak mutlak di semua keadaan untuk memprediksi gagal bayar pinjaman karena perbedaan atribut yang digunakan untuk membuat model prediksi ataupun karena pengayaan data untuk mengisi data pada atribut data yang masih kosong. Dengan pertimbangan dataset yang digunakan dari koperasi ini sedikit/kecil dan tidak terlalu kompleks, maka pada penelitian ini cukup untuk tidak mengikutsertakan algoritma XRBoost yang kemampuannya cenderung bagus untuk menangani dataset dengan ukuran sedang hingga besar, terlebih sudah ada beberapa penelitian terkait yang sudah menunjukkan kelebihan algoritma ini dalam bidang prediksi. Untuk itu, algoritma yang digunakan pada penelitian ini hanya 4 algoritma saja yaitu *logistic regression*, *decision tree*, *random forest*, dan *k-nearest neighbors (KNN)*.

Tujuan dari penelitian ini adalah untuk mengembangkan model prediksi kegagalan bayar pinjaman menggunakan beberapa algoritma Machine Learning seperti *logistic regression*, *decision tree*, *random forest*, dan *k-nearest neighbors* untuk kemudian dibandingkan agar diketahui algoritma mana yang paling efektif untuk membuat model prediksi gagal bayar pinjaman di Koperasi Simpan Pinjam Pamipiran yang datasetnya terbilang kecil. Hasil penelitian ini diharapkan dapat membantu dalam meminimalisir risiko kebangkrutan dan meningkatkan kinerja keuangan Koperasi Simpan Pinjam Pamipiran. Dari penelitian ini juga diharapkan bisa menjadi sebuah referensi untuk penelitian-penelitian serupa di masa yang akan datang.

Metode penelitian

Metode yang digunakan dalam penelitian ini adalah studi kepustakaan dan penelitian eksperimental. Pada penelitian ini dilakukan perbandingan beberapa model prediksi gagal bayar pinjaman menggunakan algoritma machine learning, yaitu *logistic regression*, *decision tree*, *random forest*, dan *k-nearest neighbors*. Proses penelitian terdiri dari beberapa tahap yang dijelaskan sebagai berikut:



Gambar 1 Tahapan Penelitian

Dari **Gambar 1** dapat diuraikan tahapan-tahapan penelitian yang dilakukan sebagai berikut:

A. Studi Literatur

Tahapan pertama penelitian dilakukan dengan membaca dan melakukan review penelitian terkait berupa artikel dari jurnal-jurnal yang terakreditasi untuk digunakan sebagai referensi yang berkualitas. Total artikel yang direview berjumlah 6 dengan hasil sebagai berikut:

Tabel 1 Review Artikel Penelitian Terkait

Penulis	Fokus	Dataset	Temuan
Kadek et al.[4]	Machine learning digunakan untuk membuat prediksi gagal bayar kredit menggantikan metode tradisional yang kurang akurat	Dataset berisi informasi 16.1715 pinjaman antara Januari dan Agustus 2021	Algoritma XGBoost mendapatkan hasil terbaik dengan akurasi 97%.
Lkhagvadorj et al.[5]	Model risiko kredit dengan machine learning bisa menggantikan atau membantu dengan berkolaborasi dengan model yang berbasis pakar domain keuangan yang masih mendominasi	Dataset diambil dari survei cross-sectional tiga tahunan, yang digunakan sebanyak 4245 data dengan 345 variabel	Algoritma <i>Deep Neural Network</i> mendapatkan hasil yang lebih menjanjikan di antara algoritma machine learning lainnya
Lin Zhua et al.[6]	Pengembangan model prediksi gagal bayar pinjaman dengan menggunakan beberapa pendekatan agar mendapat kinerja yang lebih optimal	Dataset berisi lebih dari 115.000 data pinjaman asli pengguna dengan 102 atribut	Random Forest memiliki kinerja terbaik, dengan akurasi 98% dibandingkan dengan mesin vektor pendukung dan regresi logistic.
Aida et al.[7]	Pengembangan model prediksi kuantitatif yang akurat sebagai sinyal peringatan dini untuk risiko gagal bayar	Dataset terdiri dari 924 catatan kredit perusahaan Tunisia dengan 24 indikatornya, yang diberikan oleh bank komersial Tunisia dari tahun 2003 hingga 2006	Hasil utama menunjukkan bahwa K-NN terbaik dengan k=3 untuk ketiga model, dan tingkat klasifikasi global terbaik berada di urutan 88,63% pada pengklasifikasian yang kedua
Laura et al.[8]	Mengembangkan model prediksi yang tepat dengan membandingkan 3 algoritma machine learning	Dataset terdiri dari data 1.015 perusahaan non-default dan 264 UKM default dengan 14 variabel	Algoritma Neural Networks mengungguli metode Regresi logistik dan pohon keputusan dengan memberikan tingkat deteksi tertinggi secara keseluruhan
Khandani et al.[9]	Penggunaan model dan algoritma daripada keputusan manusia untuk menghasilkan "skor" numerik untuk mendukung keputusan penilaian kelayakan kredit konsumen	Dataset yang digunakan berasal dari dataset eksklusif dari bank komersial yang cukup besar dari tahun 2005 sampai 2009.	Prediksi di luar sampel memiliki korelasi yang tinggi denganketerlambatan yang terjadi, dengan R2 regresi linear sebesar 85% untuk prediksi bulanan dalam jangka waktu 6 dan 12 bulan

B. Pengumpulan Data

Dataset yang digunakan berasal dari data historis pinjaman dari koperasi simpan pinjam Pamipiran pada satu bulan terakhir yang berjumlah 324 data pinjaman dengan 9 atribut. Data diambil dari database koperasi yang mencakup beberapa informasi mengenai profil peminjam, no transaksi, sejarah kredit, total pinjaman, banyak angsuran, dan status pembayaran pinjaman.

C. Pra-Pemrosesan Data

Data yang telah dikumpulkan kemudian menjalani tahap pra-pemrosesan untuk mempersiapkannya agar siap untuk analisis dan dibuatkan modelnya. Langkah-langkah pre-processing data meliputi :

1. Pembersihan Data: data yang dianggap kurang penting untuk model prediksi dihapus dari dataset, seperti no transaksi, nama peminjam, dan tanggal meminjam.
2. Pemilihan Fitur/Atribut: atribut-atribut yang digunakan untuk prediksi pada penelitian ini yaitu umur, status_bekerja, nilai_kredit, banyak_angsuran, total_pinjaman, sisa_pinjaman, dan gagal_bayar (status pembayaran pinjaman).
3. Pemrosesan Missing Values: menangani nilai yang hilang dalam data, baik dengan mengisi nilai yang hilang atau menggunakan teknik imputasi.
4. d. Pemrosesan Outliers: Mengatasi adanya data outlier yang dapat memengaruhi hasil analisis.
5. e. Transformasi Data: Melakukan transformasi data jika diperlukan, seperti normalisasi atau standarisasi.

Setelah melewati tahap pra-pemrosesan data kemudian didapatlah data yang siap digunakan model untuk melakukan prediksi dengan jumlah 324 records dan 7 atribut, untuk detailnya sebagai berikut:

```
data_pinjaman.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 324 entries, 0 to 323
Data columns (total 7 columns):
#   Column              Non-Null Count  Dtype
---  ---             
0   umur                 324 non-null    int64
1   status_bekerja      324 non-null    int64
2   nilai_kredit        324 non-null    int64
3   banyak_angsuran    324 non-null    int64
4   total_pinjaman     324 non-null    int64
5   sisa_pinjaman      324 non-null    int64
6   gagal_bayar         324 non-null    int64
dtypes: int64(7)
memory usage: 17.8 KB
```

Gambar 2 Informasi Dataset

D. Pembagian Data

Data yang telah dipreproses kemudian dibagi menjadi dua subset yaitu data pelatihan (training data) sebesar 20% dan data pengujian (testing data) 80% sisanya. Pembagian ini dilakukan untuk menguji kinerja model yang dikembangkan pada data yang belum pernah dilihat sebelumnya.

E. Pembuatan Model

Algoritma machine learning yang digunakan pada penelitian ini yaitu *logistic regression*, *decision tree*, *random forest*, dan *k-nearest neighbors*. Model dibuat dengan menggunakan bahasa pemrograman python pada aplikasi *Jupyter Notebook* dengan memanfaatkan library-library yang sudah ada seperti *pandas*, *sklearn* dan lainnya. Kemudian keempat algoritma tersebut lalu diimplementasikan pada model untuk melakukan prediksi prediksi gagal bayar pinjaman dengan tahapan sebagai berikut:

- 1) Membaca Data

Tahapan pertama adalah membaca dataset dari file CSV yang telah diperoleh menggunakan library pandas. Data tersebut merupakan dataset yang berisi informasi mengenai pinjaman dan status gagal bayar.

2) Pembagian Dataset

Setelah memuat data, dataset perlu dibagi menjadi data training dan data testing seperti yang dijelaskan pada poin pembagian data.

3) Inisialisasi Model

Pada proses inilah keempat algoritma tersebut diinisialisasi dengan menggunakan kelas-kelas yang sudah disediakan oleh library *sklearn* pada python.

- *Logistic Regression*

Model diinisialisasi dengan menggunakan kelas 'LogisticRegression' yaitu kelas yang disediakan oleh library scikit-learn (*sklearn.linear_model*) untuk implementasi algoritma regresi logistik dalam tugas klasifikasi, kita dapat memodelkan probabilitas bahwa suatu sampel data masuk ke dalam kelas positif (1) atau kelas negatif (0). Regresi logistik menggunakan fungsi logistik (sigmoid) untuk memetakan input linier ke dalam probabilitas kelas.

- *Decision Tree*

Model diinisialisasi dengan menggunakan kelas 'DecisionTreeClassifier' yang di dalamnya struktur *decision tree* dibangun berdasarkan pemilihan fitur yang paling informatif untuk membagi data menjadi kelompok-kelompok yang semakin homogen dalam kelas target. Setiap node pada *decision tree* merepresentasikan keputusan berdasarkan fitur-fitur yang dievaluasi, sedangkan cabang-cabang dari node tersebut mewakili kemungkinan nilai fitur tersebut. Proses ini dilakukan secara rekursif hingga mencapai kondisi berhenti seperti kedalaman maksimum atau ukuran minimum dari node.

- *Random Forest*

Model diinisialisasi dengan kelas 'RandomForestClassifier', di sini kumpulan pohon keputusan dibangun secara acak dengan menggunakan subset acak dari data training dan subset acak dari fitur-fitur yang tersedia. Setiap pohon dalam Random Forest dihasilkan dengan cara sama seperti pada Decision Tree, yaitu dengan membagi data berdasarkan fitur-fitur yang paling informatif. Bedanya *Random Forest* memperkenalkan variasi dengan menggunakan subset data dan subset fitur secara acak untuk mengurangi overfitting dan meningkatkan generalisasi.

- *K-Nearest Neighbors*

Model diinisialisasi dengan menggunakan kelas 'KNeighborsClassifier' dari library scikit-learn (*sklearn.neighbors*). Dengan kelas ini model mencari K tetangga terdekat dari suatu data uji berdasarkan jarak Euclidean atau metrik jarak lainnya. Kemudian dengan mempertimbangkan label tetangga-tetangga tersebut model dapat menentukan label prediksi untuk data uji.

4) Pelatihan Model

Setelah algoritma diinisialisasi pada masing-masing modelnya, tahap berikutnya model dilatih menggunakan data training dengan menggunakan metode 'fit' pada objek model yang telah diinisialisasi sebelumnya. Pada proses ini model dapat mempelajari pola dan struktur dari data training agar model dapat melakukan prediksi.

5) Prediksi

Sampai pada tahap ini model sudah bisa digunakan untuk melakukan prediksi menggunakan data testing dengan metode 'predict' untuk kemudian hasilnya bisa dievaluasi dan dibandingkan dengan satu sama lainnya.

F. Evaluasi Model

Setelah model melakukan prediksi gagal bayar untuk dataset koperasi dengan menggunakan seluruh atribut yang ada, kemudian dilakukan pengulangan dengan pengurangan beberapa

atribut pada dataset untuk mengetahui hasil lainnya untuk dibandingkan dengan menggunakan metrik-metrik seperti akurasi, presisi, recall, dan f1-score untuk mengevaluasi kinerja dan kemampuan prediktif model.

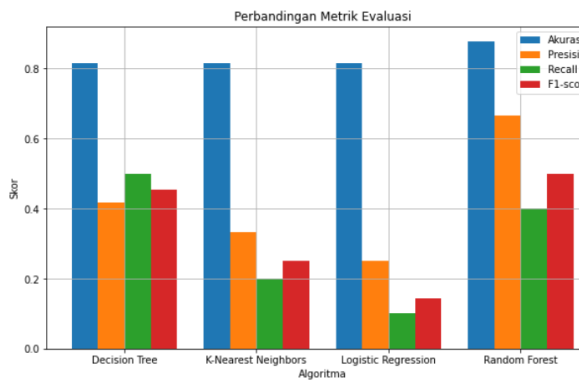
Dari hasil evaluasi ini, model dengan kinerja terbaik dapat diidentifikasi dan dipilih sebagai model prediksi gagal bayar pinjaman yang paling efektif dan optimal untuk memprediksi gagal bayar pinjaman di Koperasi Simpan-Pinjam Pampiran ataupun bisa digunakan juga untuk kasus serupa di beberapa lembaga keuangan kecil lainnya.

HASIL DAN PEMBAHASAN

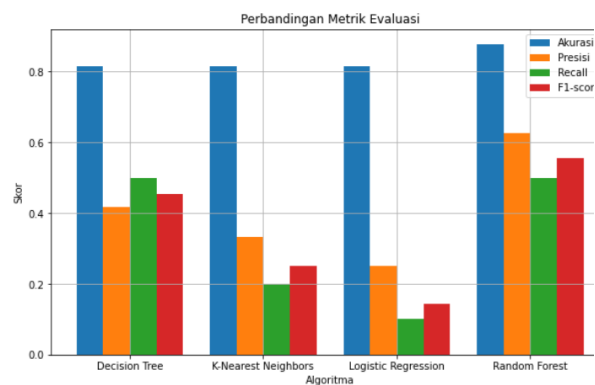
Dengan Model prediksi gagal bayar pinjaman yang telah dibuat dengan mengimplementasikan keempat algoritma dengan seluruh atributnya kemudian dilakukan pengujian dengan mengurangi atribut yang dirasa kurang berpengaruh untuk mengetahui hasil lain dari model yang dibuat jika atributnya berubah. Atribut yang digunakan pada pemodelan ini meliputi: (U)umur, (S)status_pekerjaan, (NK)nilai_kredit, (BK)banyak_angsuran, (TP), total_pinjaman, (SP)sisa_pinjaman, (GB)gagal_bayar. Pengujian dilakukan dengan membuat dataset yang telah siap tadi dijadikan tiga dataset dengan perbedaan atribut di dalamnya dengan rinciannya sebagai berikut :

- 1) dataset : U, SB, NK, BA, TP, SP, GB
- 2) dataset_2 : SB, NK, BA, TP, SP, GB
- 3) dataset_3 : NK, BA, TP, SP, GB

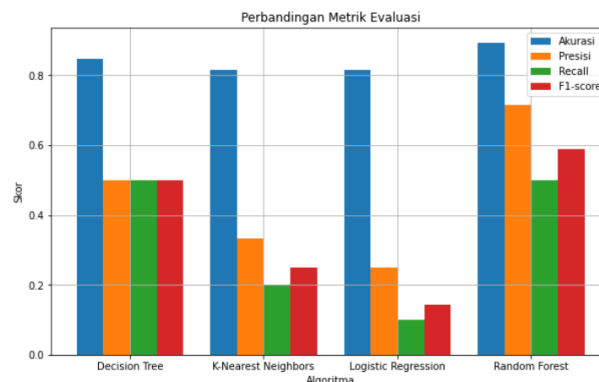
Berikut adalah hasil pengujian model dengan ketiga dataset tersebut:



Gambar 3 Pengujian dengan dataset pertama



Gambar 4 Pengujian dengan dataset kedua



Gambar 5 Pengujian dengan dataset ketiga

Dari hasil pengujian di atas pada **Gambar 3, 4, 5** dapat diuraikan hasilnya untuk setiap model seperti berikut:

A. Performa Model dengan Algoritma *Decision Tree*

Model pertama mengimplementasikan algoritma *Decision Tree* untuk membangun model prediksi gagal bayar pinjaman. Setelah melatih model dengan menggunakan data training dilakukan evaluasi performa model menggunakan data testing. Berikut disajikan tabel evaluasi dengan metrik akurasi, presisi, recall, dan f1-score dari model dengan seluruh atribut dan juga setelah pengurangan beberapa atribut:

Tabel 2 Evaluasi Model dengan algoritma *Decision Tree*

Atribut pada dataset	Akurasi	Presisi	Recall	F1-Score
U, SB, NK, BA, TP, SP, GB	0.8153	0.4166	0.5	0.4545
SB, NK, BA, TP, SP, GB	0.8153	0.4166	0.5	0.4545
NK, BA, TP, SP, GB	0.8461	0.5	0.5	0.5

Hasil di atas menunjukkan bahwa model dengan algoritma *Decision Tree* berhasil mencapai akurasi terbaiknya sebesar 84% dalam memprediksi gagal bayar pinjaman untuk dataset tanpa atribut 'umur' dan 'status_bekerja'.

B. Performa Model dengan Algoritma *K-Nearest Neighbors*

Selanjutnya algoritma yang digunakan untuk model yaitu *K-Nearest Neighbors*, berikut adalah tabel evaluasi hasil yang didapat dari pengujian yang dilakukan:

Tabel 3 Evaluasi Model dengan algoritma *K-Nearest Neighbors*

Atribut pada dataset	Akurasi	Presisi	Recall	F1-Score
U, SB, NK, BA, TP, SP, GB	0.8153	0.3333	0.2	0.25
SB, NK, BA, TP, SP, GB	0.8153	0.3333	0.2	0.25
NK, BA, TP, SP, GB	0.8153	0.3333	0.2	0.25

Hasil pengujian di atas menunjukkan bahwa ketiga pengujian dengan model dengan algoritma *K-Nearest Neighbors* mempunyai hasil yang sama dan tidak terjadi perubahan hasil pengujian meskipun atribut pada datasetnya dikurangi. Nilai akurasinya pun sama 81% atau 0.8153.

C. Performa Model dengan Algoritma *Logistic Regression*

Selanjutnya model dengan algoritma *Logistic Regression*, berikut disajikan tabel evaluasi dengan metrik akurasi, presisi, recall, dan f1-score dari model dengan seluruh atribut dan juga setelah pengurangan beberapa atribut:

Tabel 4 Evaluasi Model dengan algoritma *Logitic Regression*

Atribut pada dataset	Akurasi	Presisi	Recall	F1-Score
U, SB, NK, BA, TP, SP, GB	0.8153	0.25	0.1	0.1428
SB, NK, BA, TP, SP, GB	0.8153	0.25	0.1	0.1428
NK, BA, TP, SP, GB	0.8153	0.25	0.1	0.1428

Hasil di atas menunjukkan bahwa model *Logistic Regression* berhasil memiliki hasil sama untuk semua metrik, mirip dengan yang terjadi pada model *K-Nearest Neighbors*. Nilai akurasi, presisi, recall maupun f1-scorenya pada ketiga kondisi yang berbeda hasilnya tetap sama meskipun atributnya ada yang dikurangi, dengan nilai akurasi yang cukup baik yaitu 81%.

D. Performa Model dengan Algoritma *Random Forest*

Kemudian algoritma terakhir yang digunakan untuk model yaitu *Random Forest* dan

berikut adalah tabel evaluasi hasil yang didapat dari pengujian yang dilakukan:

Tabel 4 Evaluasi Model dengan algoritma *Random Forest*

Atribut pada dataset	Akurasi	Presisi	Recall	F1-Score
U, SB, NK, BA, TP, SP, GB	0.8769	0.6666	0.4	0.5
SB, NK, BA, TP, SP, GB	0.8769	0.625	0.5	0.5555
NK, BA, TP, SP, GB	0.8923	0.714	0.5	0.5882

Hasil di atas menunjukkan bahwa model dengan algoritma *Random Forest* berhasil mencapai akurasi terbaiknya sebesar 89% untuk dataset yang atributnya mengabaikan atribut ‘umur’ dan ‘status_pekerjaan’. Dapat dilihat hasil pengujian berbeda untuk setiap dataset yang atributnya berbeda. Ada sedikit peningkatan nilai akurasi, presisi, recall, maupun f1-score nya terjadi ketika atributnya yang dirasa kurang berpengaruhnya dikurangi.

IV. KESIMPULAN

Berdasarkan hasil evaluasi kinerja dari keempat model yang digunakan, dapat disimpulkan bahwa model yang mengimplementasikan algoritma *Random Forest* menghasilkan nilai terbaik dalam memprediksi gagal bayar pinjaman pada semua metrik yang. Akurasi terbaik model ini adalah 89%, lebih tinggi dari model yang menggunakan algoritma *logistic regression* (akurasi 81%), *decision tree* (akurasi 84%), dan *k-nearest neighbors* (akurasi 81%).

Hasil ini menunjukkan bahwa penggunaan algoritma *Random Forest* lebih efektif dalam memprediksi gagal bayar pinjaman. Model *Random Forest* mampu menggabungkan beberapa pohon keputusan untuk membuat prediksi yang lebih akurat. Saat memprediksi gagal bayar pinjaman, akurasi yang tinggi sangat penting agar Koperasi dapat mengidentifikasi risiko kredit yang tinggi dengan optimal dan mengambil tindakan yang tepat. Namun perlu dicatat bahwa pilihan algoritma bergantung pada karakteristik data, ukuran kumpulan data, kompleksitas masalah, dan tujuan prediksi yang dapat dicapai. Oleh karena itu, disarankan untuk melakukan eksperimen dan studi lebih lanjut untuk memvalidasi hasil ini dan mempertimbangkan faktor lain seperti waktu komputasi, interpretasi model, dan faktor terkait. Dalam penelitian mendatang direkomendasikan untuk memeriksa dan membandingkan algoritme lain seperti *Support Vector Machines (SVM)*, *Naive Bayes*, dan algoritme *Deep Learning* lainnya seperti *Neural Networks* dengan arsitektur yang lebih kompleks. Penelitian lebih lanjut akan memberikan pemahaman yang lebih dalam tentang algoritma untuk membuat model prediksi gagal bayar pinjaman yang paling tepat dan panduan yang lebih rinci untuk lembaga keuangan lainnya dalam keputusan pemberian pinjaman.

Referensi

- [1] Andi Nursyahriana, Michael Hadjat, Irsan Tricahyadinata, “ANALISIS FAKTOR PENYEBAB TERJADINYA KREDIT MACET”, *Forum Ekonomi: Jurnal Ekonomi, Manajemen dan Akuntansi Volume 19, No. 1, 2017*
- [2] Sukri Syafudin, Ranu Agastya Nugraha, Kartika Handayani, Windu Gata, Safitri Linawati, “PREDIKSI STATUS PINJAMAN BANK DENGAN DEEP LEARNING NEURAL NETWORK (DNN)”, *Jurnal Teknik Komputer AMIK BSI Volume 7, No.2, Juli 2021*

- [3] Peter Martey Addo, Dominique Guegan, Bertrand Hassan, "CREDIT RISK ANALYSIS USING MACHINE AND DEEP LEARNING MODELS", *University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 08/WP/ 2018*
- [4] Kadek Dwi Pradnyana, Raden Aswin Rahadi, *et al.*, "LOAN DEFAULT PREDICTION IN MICROFINANCE GROUP LENDING WITH MACHINE LEARNING," *International Journal of Business and Technology Management*, Vol.4, No. 4, 83-95, 2022, 31 December 2022
- [5] Lkhagvadorj Munkhdalai, Tsendsuren Munkhdalai, Oyun-Erdene Namsrai, Jong Yun Lee, and Keun Ho Ryu, *et al.*, "AN EMPIRICAL COMPARISON OF MACHINE-LEARNINGMETHODS ON BANK CLIENT CREDIT ASSESSMENTS," *MDPI*, 29 January 2019
- [6] Lin Zhua, Dafeng Qiua, Daji Ergua, Cai Yinga, Kuiyi Liu, "A STUDY ON PREDICTING LOAN DEFAULT BASED ON THE RANDOM FOREST ALGORITHM," *7th International Conference on Information Technology and Quantitative Managemet*, 2019
- [7] Aida Krichene Abdelmoula, "BANK CREDIT RISK ANALYSIS WITH K-NEARESTNEIGHBOR CLASSIFIER: CASE OF TUNISIAN BANKS" *Accounting and Management Information Systems* Vol. 14, No. 1, pp. 79-106, 2015
- [8] Laura Maria Badea Stroie, "TECHNIQUES FOR CUSTOMER BEHAVIOUR PREDICTION: A CASESTUDY FOR CREDIT RISK ASSESSMENT" *New Techniques and Technologies for Statistics*, 2013
- [9] Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo, "CONSUMER CREDIT-RISK MODELS VIA MACHINE-LEARNING ALGORITHMS" *Journal of Banking & Finance* 34, 2010