

Penerapan Metode *Random Over-Under Sampling* dan *Random Forest* untuk Klasifikasi Penilaian Kredit

Akhmad Syukron¹, Agus Subekti²

¹AMIK BSI Yogyakarta
e-mail: akhmad.khy@bsi.ac.id

²Lembaga Ilmu Pengetahuan Indonesia (LIPI)
e-mail: agus.subekti@lipi.go.id

Abstrak

Penilaian kredit telah menjadi salah satu cara utama bagi sebuah lembaga keuangan untuk menilai resiko kredit, meningkatkan arus kas, mengurangi kemungkinan resiko dan membuat keputusan manajerial. Salah satu permasalahan yang dihadapi pada penilaian kredit yaitu adanya ketidakseimbangan distribusi *dataset*. Metode untuk mengatasi ketidakseimbangan kelas yaitu dengan metode resampling, seperti menggunakan Oversampling, undersampling dan hibrida yaitu dengan menggabungkan kedua pendekatan sampling. Metode yang diusulkan pada penelitian ini adalah penerapan metode *Random Over-Under Sampling Random Forest* untuk meningkatkan kinerja akurasi klasifikasi penilaian kredit pada *dataset German Credit*. Hasil pengujian menunjukkan bahwa klasifikasi tanpa melalui proses resampling menghasilkan kinerja akurasi rata-rata 70 % pada semua *classifier*. Metode *Random Forest* memiliki nilai akurasi yang lebih baik dibandingkan dengan beberapa metode lainnya dengan nilai akurasi sebesar 0,76 atau 76%. Sedangkan klasifikasi dengan penerapan metode *Random Over-under sampling Random Forest* dapat meningkatkan kinerja akurasi sebesar 14,1% dengan nilai akurasi sebesar 0,901 atau 90,1 %. Hasil penelitian menunjukkan bahwa penerapan *resampling* dengan metode *Random Over-Under Sampling* pada algoritma *Random Forest* dapat meningkatkan kinerja akurasi secara efektif pada klasifikasi tidak seimbang untuk penilaian kredit pada *dataset German Credit*.

Kata kunci: Penilaian Kredit, *Random Forest*, Klasifikasi, ketidakseimbangan kelas, *Random Over-Under Sampling*

Abstract

Credit scoring has become one of the main ways for a financial institution to assess credit risk, improve cash flow, reduce the possibility of risk and make managerial decisions. One of the problems faced by credit scoring is the imbalance in the distribution of datasets. The method to overcome class imbalances is the resampling method, such as using Oversampling, undersampling and hybrids by combining both sampling approaches. The method proposed in this study is the application of the Random Over-Under Sampling Random Forest method to improve the accuracy of the credit scoring classification performance on German Credit dataset. The test results show that the classification without going through the resampling process results in an average accuracy performance of 70% for all classifiers. The Random Forest method has a better accuracy value compared to some other methods with an accuracy value of 0.76 or 76%. While classification by applying the Random Over-under sampling + Random Forest method can improve accuracy performance 14.1% with an accuracy value of 0.901 or 90.1%. The results showed that the application of resampling using Random Over-Under Sampling method in the Random Forest algorithm can improve accuracy performance effectively on an unbalanced classification for credit scoring on German Credit dataset.

Keywords: Imbalance Class, Credit Scoring, Random Forest, Classification, Resampling

1. Pendahuluan

Penilaian kredit telah menjadi salah satu cara utama bagi sebuah lembaga keuangan untuk menilai resiko kredit, meningkatkan arus kas, mengurangi kemungkinan resiko dan membuat keputusan manajerial (Lin, Wang, Xie, & Zhong, 2017). Penilaian kredit juga sebagai proses pengakuan pelanggan bank untuk memberikan kredit berdasarkan seperangkat kriteria tertentu (Koutanaei, Sajedi, & Khanabaei, 2015). Tujuan dari penilaian kredit adalah untuk mengklasifikasi para calon pemohon kredit yang diklasifikasikan menjadi dua jenis yaitu calon pemohon yang baik dan calon pemohon yang buruk. Penilaian kredit tersebut berdasarkan dari kemampuan calon pemohon untuk membayar kewajiban keuangannya dengan melihat karakteristik seperti jenis kelamin, umur, pekerjaan, dan gaji dan lain sebagainya (He, Zhang, & Zhang, 2018).

Ada dua teknik klasifikasi utama yang digunakan untuk penilaian kredit, yaitu dengan teknik statistik dan machine learning (He, 2018). Beberapa teknik statistik yang telah banyak digunakan dalam penilaian kredit yaitu seperti *Linear Discriminant Analysis (LDA)*, *Logistic Regression Analysis (LRA)* (Zhang, Yang, & Zhou, 2018). Selain itu penggunaan teknik *machine learning* dengan menggunakan *Artificial Intelligence (AI)* seperti *Artificial Neural Networks (ANN)*, *Decision tree (DT)*, *Case-Based Reasoning (CBR)*, *Support Vector Machine (SVM)* dan lainnya telah banyak digunakan diberbagai penelitian tentang penilaian kredit (Wang, Hao, Ma, & Jiang, 2011).

Salah satu permasalahan yang dihadapi dalam penggunaan teknik *machine learning* yaitu masalah ketidakseimbangan kelas. Ketidak- seimbangan kelas terjadi pada saat kelas minoritas jauh lebih kecil atau lebih jarang dari kelas mayoritas (Ren, Cao, Li, Zhao, & Zaiane, 2017). Model yang dibuat dengan menggunakan data tidak seimbang akan menghasilkan akurasi prediksi minoritas yang rendah. Informasi yang kaya dari kelas mayoritas mendominasi kelas minoritas sehingga menyebabkan batas-batas keputusan yang bias dalam sistem klasifikasi (Jian, Gao, & Ao, 2016).

Beberapa metode untuk mengatasi ketidakseimbangan kelas dapat dibagi menjadi tiga kategori. Pertama yaitu dengan teknik tingkat data yang berusaha menyeimbangkan distribusi data dengan

metode *over-sampling* dan *undersampling*. Kedua adalah pendekatan tingkat algoritma yaitu dengan mengembangkan algoritma baru atau memodifikasi metode yang ada untuk memperhitungkan arti dari kelas minor. Ketiga dengan mengkombinasikan pendekatan algoritma dan pendekatan level data (Xiao, Xie, He, & Jiang, 2012).

Untuk mengatasi masalah ketidakseimbangan kelas, maka pada penelitian ini menggunakan pendekatan level data dengan menerapkan metode *resampling*, yaitu *Random over-under sampling* pada algoritma *random forest*. Sehingga dapat meningkatkan kinerja akurasi pada klasifikasi penilaian kredit.

Thomas dalam Koutanaei mendefinisikan penelanaan kredit sebagai proses pengakuan pelanggan bank untuk memberikan kredit berdasarkan seperangkat kriteria tertentu. Secara umum penilaian kredit memiliki beberapa keunggulan yaitu;

1. mengurangi biaya analisis kredit.
2. Penetapan kredit dengan proses pengambilan keputusan yang efektif dan cepat.
3. Probabilitas pembayaran kredit yang lebih tinggi.
4. kemungkinan resiko lebih rendah (Koutanaei et al., 2015).

Tujuan dari penilaian kredit adalah untuk mengklasifikasi para calon debitur yang diklasifikasikan menjadi dua jenis yaitu calon debitur yang baik dan calon debitur yang buruk. Penilaian kredit tersebut berdasarkan dari kemampuan calon debitur untuk membayar kewajiban keuangannya. Calon debitur dengan penilaian kredit yang bagus memiliki kemungkinan besar untuk membayar kewajiban keuangan. Sedangkan calon debitur dengan penilaian kredit yang buruk memiliki kemungkinan gagal untuk membayar kewajiban keuangan. Keakuratan pemberian nilai kredit sangatlah penting untuk profitabilitas lembaga keuangan. Dengan adanya peningkatan akurasi penilaian kredit terhadap calon debitur dengan resiko macet maka dapat menurunkan kerugian bagi lembaga keuangan (Wang et al., 2011).

Klasifikasi adalah proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. Klasifikasi telah banyak digunakan seperti pada deteksi

penipuan, target pemasaran, prediksi kinerja, manufaktur, dan diagnosis medis (Han,2012). Proses klasifikasi data terdiri dari 2 langkah. Pertama yaitu tahap *learning* (*fase training*), dimana algoritma klasifikasi dibuat untuk menganalisa data *training* lalu direpresentasikan dalam bentuk model klasifikasi. Proses kedua adalah tahap klasifikasi, dimana data tes digunakan untuk memperkirakan akurasi dari model klasifikasi. Jika akurasi dianggap dapat diterima, aturan bisa diterapkan pada klasifikasi tupel data baru (Han,2012).

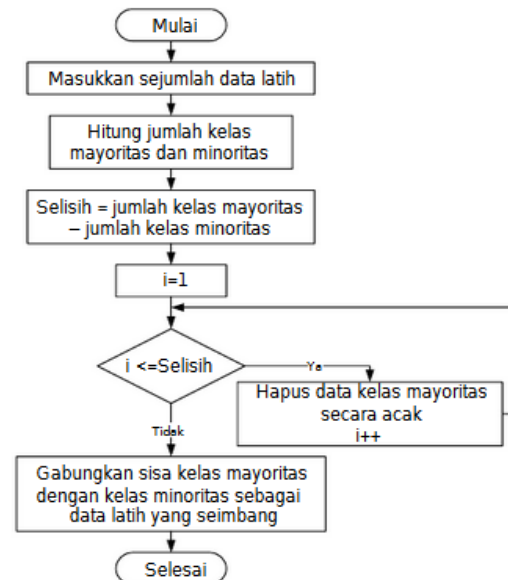
Class imbalance atau ketidakseimbangan kelas merupakan salah satu permasalahan yang ada pada data mining. Hal ini terjadi pada saat kelas minoritas jauh lebih kecil atau lebih jarang dari kelas mayoritas (Ren et al., 2017). Model yang dibuat dengan menggunakan data tidak seimbang akan menghasilkan akurasi prediksi minoritas yang rendah. Informasi yang kaya dari mayoritas mendominasi minoritas sehingga menyebabkan batas-batas keputusan yang bias dalam sistem klasifikasi (Jian et al., 2016). Selain itu ketidakseimbangan kelas (*class imbalance*) dan *noise* juga dapat berpengaruh pada kualitas data dalam hal kinerja klasifikasi.

Teknik *resampling* adalah salah satu teknik *preprocessing* dimana distribusi data diseimbangkan kembali untuk mengurangi efek distribusi kelas tidak seimbang dalam proses pembelajaran (Jian et al., 2016).

Teknik *resampling* secara luas digunakan untuk memecahkan masalah data yang tidak seimbang. Teknik ini dilakukan dengan mencoba menyeimbangkan data asli berdasarkan serangkaian algoritma sampling dengan menyesuaikan jumlah sampel dalam kelas yang berbeda, kemudian melatih data "seimbang" baru dengan mengadopsi algoritma klasifikasi.

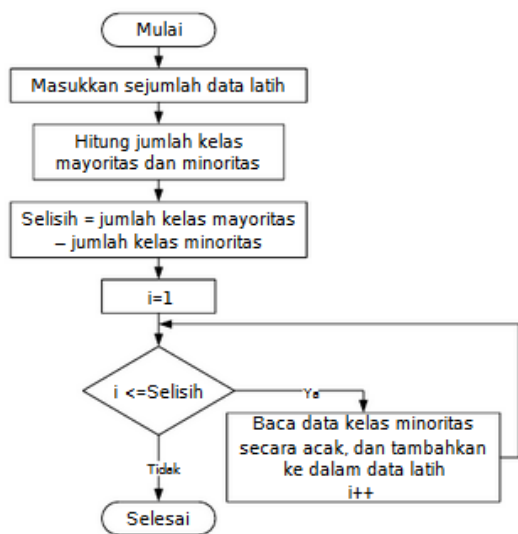
Pendekatan *resampling* dibagi menjadi tiga kategori: metode *over-sampling*, *under-sampling*, dan *hibrida* yang menggabungkan kedua pendekatan sampling (Jian et al., 2016). Salah satu teknik *resampling* yang umum digunakan yaitu *undersampling* yang secara acak memilih sampel di kelas mayoritas dan menambahkannya ke kelas minoritas, membentuk sebuah dataset pelatihan baru. *Oversampling* bertujuan untuk meningkatkan sampel kelas minoritas sampai sama dengan kelas mayoritas lain dengan menduplikasi secara acak sampel kelas minoritas (He et al., 2018). Sedangkan

random undersampling yaitu menghasilkan subsampel acak dari instance kelas mayoritas (Rajesh & Dhuli, 2018). *Undersampling* : metode sampling ini secara acak memilih sampel di kelas mayoritas dan menambahkannya ke kelas minoritas, membentuk sebuah dataset pelatihan baru (He et al., 2018).



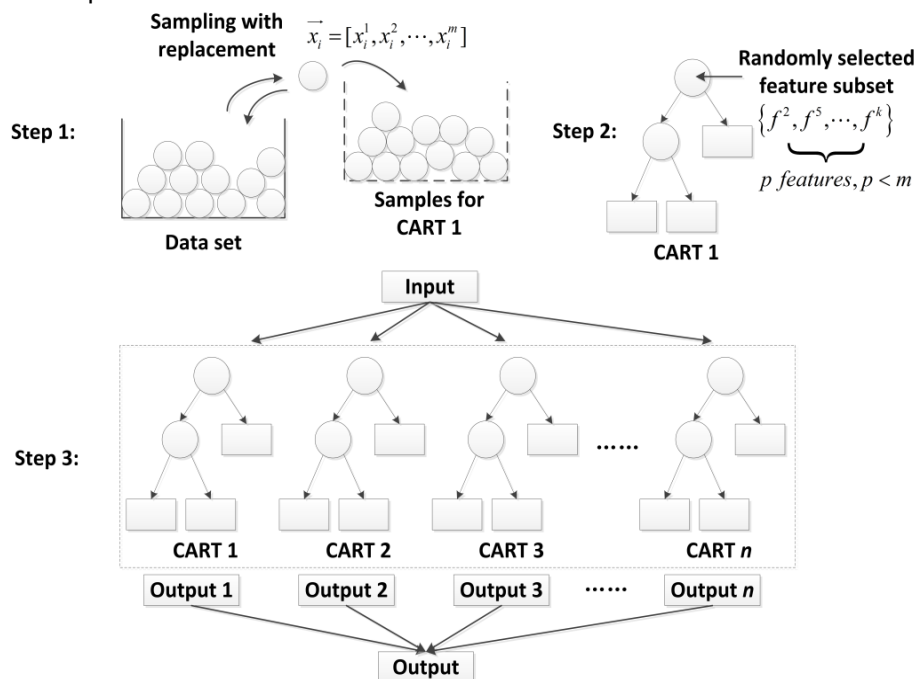
Gambar 1. Flowcart Algoritma *RandomUndersampling*
Sumber : (Saifudin et al., 2015)

Oversampling : Dalam metode ini, contoh dari kelas minoritas dipilih secara acak dan diduplikasi. instance yang dihasilkan hanya meningkatkan besarnya jumlah kelas minoritas dengan hanya mereplikasi informasi yang sama[1]. Salah satu teknik *overampling* yang populer dalam menangani ketidakseimbangan kelas yaitu *SMOTE* (*Synthetic Minority Over-sampling Technique*).



Gambar 2. Flowcart Algoritma *Random Oversampling*
Sumber : (Saifudin et al., 2015)

Random forest adalah metode *ensemble learning* yang pertama kali diusulkan oleh Breiman pada tahun 2001 yang merupakan kombinasi dari pohon klasifikasi sedemikian



Gambar 3. Flowcart Algoritma *Random Forest*
Sumber : (Lin et al., 2017)

Pada Penelitian Lean Yu, dkk mengusulkan penerapan *deep belief network (DBN)* berbasis *resampling support vector machine (SVM) ensemble learning* untuk memecahkan masalah data tidak seimbang

rupa sehingga setiap pohon bergantung pada nilai acak vektor sampel secara mandiri dan dengan distribusi yang sama untuk semua pohon di hutan (Xiao et al., 2012).

Random Forest telah banyak digunakan baik untuk klasifikasi dan regresi karena kinerjanya yang unggul dan strukturnya yang sederhana (Lin et al., 2017).

Langkah-langkah rinci dari *Random Forest* adalah sebagai berikut.

1. Menghasilkan Sebuah set-set pelatihan baru dengan sampel acak dengan penggantian (*bootstrap*) dari set pelatihan asli.
2. Untuk setiap set pelatihan baru, dibangun sebuah pohon dengan pemilihan fitur acak di setiap simpul pohon dan tanpa pemangkasan.
3. Setelah sejumlah besar pohon dihasilkan, data baru diprediksi dengan menggabungkan hasil semua pohon, dengan strategi voting mayoritas.

pada klasifikasi kredit . Beberapa model data *minning* dibandingkan untuk menilai kinerja klasifikasi. Hasil penelitiannya menunjukkan bahwa kinerja klasifikasi dapat ditingkatkan secara efektif ketika strategi *Deep belief*

network (DBN) berbasis *ensemble* terintegrasi dengan teknik *re-sampling*, terutama pada masalah ketidakseimbangan data.

Penelitian Jie Sun, dkk ini mengusulkan penerapan model *ensemble Decision Tree (DTE)* yang efektif dengan menggabungkan *SMOTE (Synthetic Minority over-sampling technique)* dan *Bagging* dengan *DSR (Differentiated sampling rates)*, yang disingkat sebagai *DTE-SBD*, sehingga dapat meningkatkan kinerja ketidakseimbangan class pada evaluasi kredit perusahaan dengan membandingkan enam model yaitu *Decision Tree* murni, *Decision Tree over-sampling*, *Decision Tree over-under-sampling*, *SMOTE Decision Tree*, *Bagging Decision Tree*, dan *DTE-SBD*.

Penelitian (Yun Qian, dkk) mengusulkan penerapan model *Resampling Ensemble algorithm (REA)* dengan *Naïve Bayes* sebagai *base learning method* dan *ensemble learning* menggunakan *Bagging* yang berfokus pada masalah klasifikasi pada *dataset* yang tidak seimbang. Pada metode ini, kelas-kelas dengan data yang kecil di-*oversampling* dan kelas-kelas dengan data yang besar di-*undersampling* dan *resampling* skala ditentukan berdasarkan rasio jumlah kelas minimal dan jumlah kelas maksimal. Hasil penelitian dari permasalahan multiple class menunjukkan bahwa *ensemble* dari berbagai jenis metode dapat meningkatkan kinerja algoritma secara efisien

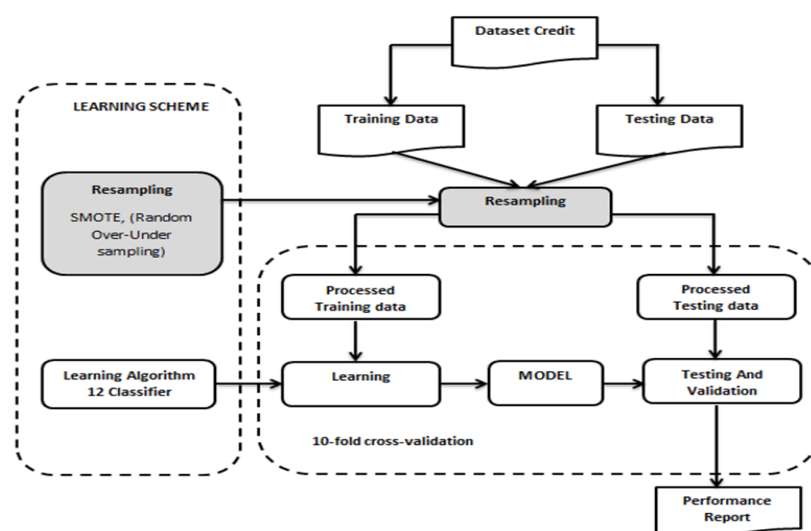
2. Metode Penelitian

Pada penelitian ini jenis penelitian yang dilakukan adalah jenis Eksperimen komparatif akan membandingkan dua objek yang berbeda, misalnya membandingkan dua algoritma yang berbeda dengan melihat hasil statistik masing-masing mana yang lebih baik. Dimana Melibatkan penyelidikan hubungan kausal menggunakan tes yang dikontrol sendiri (Dawson,2009).

Penelitian ini bertujuan untuk melakukan komparasi dan evaluasi model *ensemble* menggunakan *random forest* dan metode *resampling* untuk mengetahui algoritma yang memiliki nilai akurasi lebih tinggi pada klasifikasi tidak seimbang penilain kredit. Data yang digunakan dalam penelitian ini adalah data sekunder yang didapat dari UCI (*University of California, Irvine*) *Machine Learning Repository*.

Metode pengumpulan data yang dilakukan pada penelitian ini adalah menggunakan data sekunder yang diperoleh dari *University of California Irvine machine learning data repository* yang dapat diperoleh melalui alamat web [https://archive.ics.uci.edu/ml/datasets/statlog+\(german+credit+data\)](https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)).

Pada penelitian ini model yang diusulkan untuk mengatasi masalah ketidakseimbangan kelas yaitu dengan menerapkan kombinasi teknik pendekatan level data dengan metode *resampling* dan menggunakan pendekatan algoritma dengan metode *ensemble* berbasis *random forest*. Kerangka model yang diusulkan ditunjukkan pada gambar 3.1.



Gambar 4. Metode Usulan
Sumber : Penelitian (2018)

Pada metode yang diusulkan, tahapan awal akan dilakukan yaitu menerapkan metode *resampling* seperti *SMOTE*, dan *Random Over-Under Sampling (ROUS)* pada dataset *German Credit* agar dataset menjadi lebih seimbang. Tahapan berikutnya dilakukan pemodelan pada data *training* dengan menggunakan metode *Random Forest* dan 11 model algoritma klasifikasi lainnya sebagai pembandingan dan diuji dengan data uji melalui proses validasi. Hasil validasi digunakan untuk mengukur kinerja masing-masing model. Komparasi dilakukan dengan membandingkan kinerja *Random Forest* dengan 11 model algoritma dengan metode *resampling* seperti *SMOTE*, *Random Over-Under Sampling* dan juga tanpa melalui proses *resampling*. Untuk mengukur kinerja algoritma pengklasifikasi dilakukan dengan menggunakan *confusion matrix* yang diperoleh dari proses validasi menggunakan *k-fold cross validation* dengan metode evaluasi standarnya yaitu *10-fold cross validation* yang merupakan pilihan terbaik untuk mendapatkan hasil validasi yang akurat. Hasil yang diperoleh dari pengukuran kinerja model digunakan untuk dibandingkan antara algoritma *random forest* dengan beberapa model algoritma lainnya. Baik model yang dibentuk dengan kombinasi metode *resampling* maupun tanpa metode *resampling*. Kualitas model dapat dilihat

berdasarkan nilai *Accuracy*, *F-measure* dan juga nilai *Area Under curve (AUC)*.

3.

4. Hasil Dan Pembahasan

Pada bab ini membahas hasil eksperimen yang telah dilakukan dengan menggunakan aplikasi WEKA 3.8.1. Dataset yang digunakan pada penelitian ini adalah *dataset German Credit* yang terdiri dari 20 atribut, 1000 instance dan 2 class yaitu class "good" dan class "bad". Dengan jumlah *instance* masing-masing class yaitu 700 untuk class "good", dan 300 *instance* untuk class "bad". Pengaturan eksperimen dilakukan untuk menghasilkan akurasi yang paling tinggi pada metode yang diusulkan yaitu dengan membandingkan metode *random forest* dengan beberapa metode lainnya dengan penerapan *resampling* seperti *SMOTE* dan *Random Over-Under Sampling* ataupun tanpa *resampling*.

3.1. Klasifikasi *Random Forest*

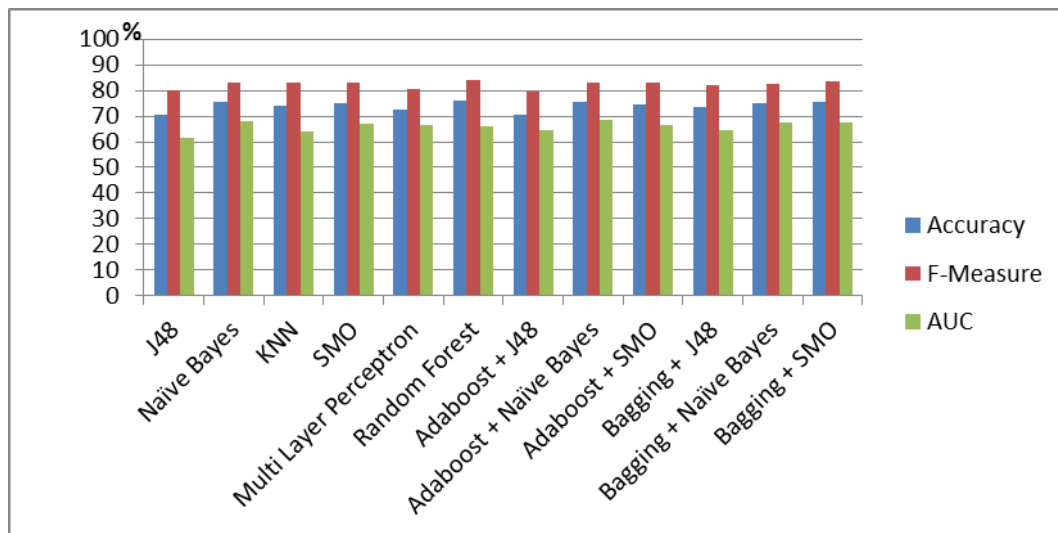
Pada tahapan ini eksperimen dilakukan tanpa menggunakan metode *resampling*. Tetapi langsung dilakukan pemodelan dan pengujian menggunakan metode *random forest*. Hasil pengujian dibandingkan dengan 11 metode lainnya untuk menghasilkan model dengan kinerja akurasi terbaik. Hasil pengujian model pengklasifikasi tanpa *resampling* disajikan pada tabel 1.

Tabel 1. Perbandingan kinerja model Pengklasifikasi Tanpa *Resampling*

| No | Algoritma | TPRate | FPRate | Precision | Recall | Accuracy | F-Measure | AUC |
|----|------------------------|--------|--------|-----------|--------|----------|-----------|-------|
| 1 | J48 | 0,843 | 0,610 | 0,763 | 0,843 | 0,707 | 0,801 | 0,617 |
| 2 | Naïve Bayes | 0,864 | 0,503 | 0,800 | 0,864 | 0,754 | 0,831 | 0,681 |
| 3 | KNN | 0,891 | 0,607 | 0,774 | 0,891 | 0,742 | 0,829 | 0,642 |
| 4 | SMO | 0,873 | 0,530 | 0,794 | 0,873 | 0,752 | 0,831 | 0,672 |
| 5 | Multi Layer Perceptron | 0,809 | 0,477 | 0,798 | 0,809 | 0,723 | 0,803 | 0,666 |
| 6 | Random Forest | 0,913 | 0,597 | 0,781 | 0,913 | 0,760 | 0,842 | 0,658 |
| 7 | Adaboost + J48 | 0,803 | 0,517 | 0,784 | 0,803 | 0,707 | 0,793 | 0,643 |
| 8 | Adaboost + Naïve Bayes | 0,860 | 0,493 | 0,803 | 0,860 | 0,754 | 0,830 | 0,684 |
| 9 | Adaboost + SMO | 0,870 | 0,540 | 0,790 | 0,870 | 0,747 | 0,828 | 0,665 |
| 10 | Bagging + J48 | 0,863 | 0,570 | 0,779 | 0,863 | 0,733 | 0,819 | 0,647 |
| 11 | Bagging + Naïve Bayes | 0,859 | 0,510 | 0,797 | 0,859 | 0,748 | 0,827 | 0,675 |
| 12 | Bagging + SMO | 0,876 | 0,530 | 0,794 | 0,876 | 0,754 | 0,833 | 0,673 |

Pada tabel 1 menunjukkan bahwa model *Random forest* memiliki nilai *accuracy* dan *F-Measure* yang lebih baik daibandingkan dengan beberapa model

lainnya dengan nilai *accuracy* sebesar 0,76 dan nilai *F-Measure* sebesar 0,842. Grafik perbandingan kinerja 12 model pengklasifikasi disajikan pada gambar 5.



Gambar 5. Perbandingan kinerja pemodelan tanpa *Resampling*

3.2. Klasifikasi *SMOTE Random Forest*

Experimen dilakukan dengan menerapkan *SMOTE (Synthetic Minority Over-sampling Technique)* pada metode *random forest*. Dimana data *training* menjadi lebih seimbang antara kelas “good” dengan kelas “bad” dengan jumlah *instance* pada class “good”

700 *instance*, sedangkan class “bad” menjadi 600 *instance*. Hasil pengujian dibandingkan dengan 11 metode lainnya untuk menghasilkan model dengan kinerja *accuracy* terbaik. Hasil pengujian model pengklasifikasi dengan *SMOTE* disajikan pada tabel 2.

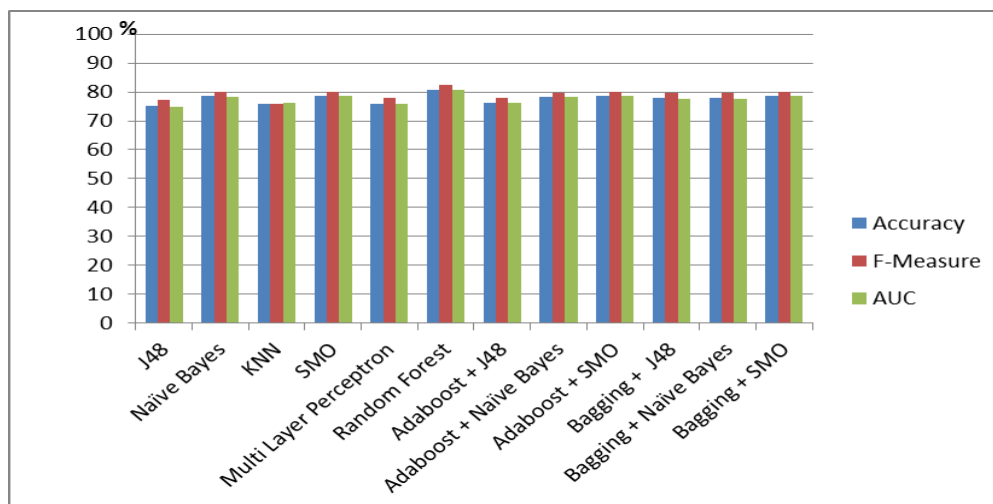
Tabel 2. Perbandingan kinerja model Pengklasifikasi dengan metode *Resampling SMOTE(Synthetic Minority Over-sampling Technique)*

| No | Algoritma | TP-rate | FP-rate | Precision | Recall | Accuracy | F-Measure | AUC |
|----|--------------------------------|---------|---------|-----------|--------|----------|-----------|-------|
| 1 | SMOTE+J48 | 0,786 | 0,290 | 0,760 | 0,786 | 0,751 | 0,772 | 0,748 |
| 2 | SMOTE+Naive Bayes | 0,801 | 0,235 | 0,799 | 0,801 | 0,785 | 0,800 | 0,783 |
| 3 | SMOTE+KNN | 0,706 | 0,180 | 0,821 | 0,706 | 0,758 | 0,759 | 0,763 |
| 4 | SMOTE+SMO | 0,786 | 0,212 | 0,812 | 0,786 | 0,787 | 0,799 | 0,787 |
| 5 | SMOTE+Multilayer perceptron | 0,784 | 0,268 | 0,773 | 0,784 | 0,760 | 0,779 | 0,758 |
| 6 | SMOTE+Random Forest | 0,833 | 0,220 | 0,815 | 0,833 | 0,808 | 0,824 | 0,807 |
| 7 | SMOTE+(Adaboost + J48) | 0,781 | 0,260 | 0,778 | 0,781 | 0,762 | 0,780 | 0,761 |
| 8 | SMOTE+(Adaboost + Naive Bayes) | 0,797 | 0,233 | 0,799 | 0,797 | 0,783 | 0,798 | 0,782 |
| 9 | SMOTE+(Adaboost + SMO) | 0,786 | 0,212 | 0,812 | 0,786 | 0,787 | 0,799 | 0,787 |

| | | | | | | | | |
|----|-------------------------------|-------|-------|-------|-------|-------|-------|-------|
| 10 | SMOTE+(Bagging + J48) | 0,801 | 0,248 | 0,790 | 0,801 | 0,778 | 0,796 | 0,777 |
| 11 | SMOTE+(Bagging + Naïve Bayes) | 0,797 | 0,243 | 0,793 | 0,797 | 0,778 | 0,795 | 0,777 |
| 12 | SMOTE+(Bagging + SMO) | 0,796 | 0,225 | 0,805 | 0,796 | 0,786 | 0,800 | 0,786 |

Pada tabel 2 menunjukkan bahwa dengan penerapan metode resampling SMOTE (*Synthetic Minority Over-sampling Technique*) menghasilkan kinerja yang lebih baik dibandingkan dengan tanpa menggunakan metode *resampling*. Hal ini ditunjukkan dengan adanya peningkatan *accuracy* pada semua model yang diuji. Hasil

kinerja yang diperoleh menunjukkan bahwa model pengklasifikasi *SMOTE+Random Forest* memiliki nilai *accuracy* yang lebih baik dengan beberapa model lainnya dengan nilai *accuracy* sebesar 0,808 atau 80,8% yang mengalami kenaikan 0,048 atau 4,8%. Grafik perbandingan kinerja 12 model pengklasifikasi disajikan pada gambar 6.



Gambar 6. Perbandingan kinerja Pemodelan dengan Metode SMOTE(*Synthetic Minority Over-sampling Technique*)

3.3. Klasifikasi *Random Over-Under Sampling Random Forest*

Experimen dilakukan dengan menerapkan *random over-under sampling* pada metode *random forest*. Dimana data *training* menjadi seimbang antara class "good" dengan class "bad" dengan jumlah *instance* pada masing-

masing class menjadi 500 *instance*. Hasil pengujian dibandingkan dengan 11 metode lainnya untuk menghasilkan model dengan kinerja akurasi terbaik. Hasil pengujian model pengklasifikasi dengan *Random Over-Under Sampling* disajikan pada tabel 3.

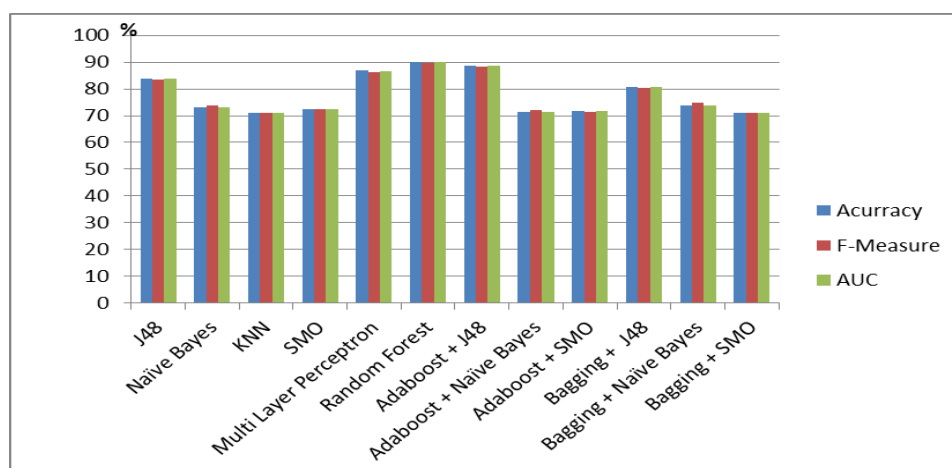
Tabel 3. Perbandingan kinerja model Pengklasifikasi dengan metode *Random Over-Under Sampling (ROUS)*

| No | Algoritma | TP-Rate | FP-Rate | Precision | Recall | Accuracy | F-Measure | AUC |
|----|----------------------------|---------|---------|-----------|--------|----------|-----------|-------|
| 1 | ROUS+J48 | 0,824 | 0,148 | 0,848 | 0,824 | 0,838 | 0,836 | 0,838 |
| 2 | ROUS+Naïve Bayes | 0,760 | 0,300 | 0,717 | 0,760 | 0,730 | 0,738 | 0,73 |
| 3 | ROUS+KNN | 0,710 | 0,290 | 0,710 | 0,710 | 0,710 | 0,710 | 0,71 |
| 4 | ROUS+SMO | 0,720 | 0,272 | 0,726 | 0,720 | 0,724 | 0,723 | 0,724 |
| 5 | ROUS+Multilayer Perceptron | 0,830 | 0,100 | 0,893 | 0,836 | 0,868 | 0,864 | 0,865 |

| | | | | | | | | |
|----|-------------------------------|-------|-------|-------|-------|-------|-------|-------|
| 6 | ROUS+Random Forest | 0,864 | 0,062 | 0,933 | 0,864 | 0,901 | 0,897 | 0,901 |
| 7 | ROUS+(Adaboost + J48) | 0,854 | 0,082 | 0,912 | 0,854 | 0,886 | 0,882 | 0,886 |
| 8 | ROUS+(Adaboost + Naive Bayes) | 0,740 | 0,310 | 0,705 | 0,740 | 0,715 | 0,722 | 0,715 |
| 9 | ROUS+(Adaboost + SMO) | 0,710 | 0,276 | 0,720 | 0,710 | 0,717 | 0,715 | 0,717 |
| 10 | ROUS+(Bagging + J48) | 0,780 | 0,166 | 0,825 | 0,780 | 0,807 | 0,802 | 0,807 |
| 11 | ROUS+(Bagging + Naive Bayes) | 0,772 | 0,296 | 0,723 | 0,772 | 0,738 | 0,747 | 0,738 |
| 12 | ROUS+(Bagging + SMO) | 0,712 | 0,292 | 0,709 | 0,712 | 0,710 | 0,711 | 0,71 |

Pada tabel 3 menunjukkan bahwa dengan penerapan metode *Random Over-Under Sampling (ROUS)* pada *dataset* sebelum dilakukan pemodelan dapat menghasilkan kinerja yang lebih baik dibandingkan dengan tanpa menggunakan metode *resampling*. Hasil kinerja yang diperoleh menunjukkan bahwa model pengklasifikasi *Random Over-*

Under Sampling (ROUS)+Random Forest memiliki nilai akurasi yang lebih baik dibandingkan dengan beberapa model lainnya dengan nilai *accuracy* sebesar 0,901 atau 90,1% yang mengalami kenaikan 0,093 atau 9,3%. Grafik perbandingan kinerja 12 model pengklasifikasi disajikan pada gambar 4



Gambar 4. Perbandingan kinerja Pemodelan dengan Metode *Random Over-Under Sampling(ROUS)*

5. Kesimpulan

Berdasarkan hasil eksperimen yang telah dilakukan pada *dataset German Credit*. Klasifikasi tanpa melalui proses *resampling* pada 12 pengklasifikasi menghasilkan kinerja akurasi rata-rata 70 %. Hasil klasifikasi menunjukkan bahwa metode *Random Forest* memiliki nilai akurasi yang lebih baik dibandingkan dengan beberapa metode lainnya dengan nilai *accuracy* sebesar 0,76 atau 76%. Sedangkan klasifikasi dengan penerapan metode *SMOTE (Synthetic Minority Over-sampling Technique)* pada algoritma *Random Forest* dapat meningkatkan kinerja akurasi sebesar 0,48%

dengan nilai akurasi 0,808 atau 80,8%. Sedangkan klasifikasi dengan penerapan metode *Random Over-under sampling+ Random Forest* dapat meningkatkan kinerja akurasi sebesar 14,1% dengan nilai akurasi sebesar 0,901 atau 90,1 %. Maka dapat disimpulkan bahwa penerapan *resampling* dengan metode *Random Over-Under Sampling* pada algoritma *Random Forest* dapat meningkatkan kinerja akurasi secara efektif pada klasifikasi tidak seimbang untuk penilaian kredit pada *dataset German Credit*.

Namun untuk penelitian selanjutnya hal-hal berikut bisa ditambahkan untuk

menghasilkan nilai akurasi yang lebih baik seperti berikut ini.

1. Pada penelitian selanjutnya dapat diterapkan teknik tingkat data lainnya seperti metode seleksi fitur (feature selection) untuk dapat meningkatkan akurasi pada klasifikasi penilaian kredit dengan ketidakseimbangan kelas.
2. Menggunakan dataset dengan jumlah instance yang lebih banyak, dan dengan rasio ketidakseimbangan kelas yang lebih tinggi.
3. Menerapkan metode ensemble learning lainnya dengan base classifier yang berbeda.

Referensi

- Agrawal, K., Baweja, Y., Dwivedi, D., Saha, R., Prasad, P., Agrawal, S., ... Dutt, V. (2018). A Comparison of Class Imbalance Techniques for Real-World Landslide Predictions. *Proceedings - 2017 International Conference on Machine Learning and Data Science, MLDS 2017, 2018-Janua*, 1–8. <https://doi.org/10.1109/MLDS.2017.21>
- Dawson, C. W. (2009). *Projects in Computing and Information Systems A Student's Guide (2nd ed.)*. Pearson Education Limited.
- Han, J., Kamber, M., Pei, J., (2012). *Data Mining Concept And Techniques*. California: Morgan Kaufmann.
- He, H., Zhang, W., & Zhang, S. (2018). A novel ensemble method for credit scoring: Adaption of different imbalance ratios. *Expert Systems with Applications*, 98, 105–117. <https://doi.org/10.1016/j.eswa.2018.01.012>
- Jian, C., Gao, J., & Ao, Y. (2016). A new sampling method for classifying imbalanced data based on support vector machine ensemble. *Neurocomputing*, 193, 115–122. <https://doi.org/10.1016/j.neucom.2016.02.006>
- Koutanaei, F. N., Sajedi, H., & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, 27, 11–23. <https://doi.org/10.1016/j.jretconser.2015.07.003>
- Lin, L., Wang, F., Xie, X., & Zhong, S. (2017). Random forests-based extreme learning machine ensemble for multi-regime time series prediction. *Expert Systems with Applications*, 83, 164–176. <https://doi.org/10.1016/j.eswa.2017.04.013>
- Rajesh, K. N. V. P. S., & Dhuli, R. (2018). Classification of imbalanced ECG beats using re-sampling techniques and AdaBoost ensemble classifier. *Biomedical Signal Processing and Control*, 41, 242–254. <https://doi.org/10.1016/j.bspc.2017.12.004>
- Ren, F., Cao, P., Li, W., Zhao, D., & Zaiane, O. (2017). Ensemble based adaptive over-sampling method for imbalanced data learning in computer aided detection of microaneurysm. *Computerized Medical Imaging and Graphics*, 55, 54–67. <https://doi.org/10.1016/j.compmedimag.2016.07.011>
- Saifudin, A., Teknik, F., Pamulang, U., Komputer, F. I., Nuswantoro, U. D., & Software, P. C. (2015). Pendekatan Level Data untuk Menangani Ketidakseimbangan Kelas pada Prediksi Cacat Software. *Journal of Software Engineering*, 1(2), 76–85.
- Wang, G., Hao, J., Ma, J., & Jiang, H. (2011). A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, 38(1), 223–230. <https://doi.org/10.1016/j.eswa.2010.06.048>
- Xiao, J., Xie, L., He, C., & Jiang, X. (2012). Dynamic classifier ensemble model for customer classification with imbalanced class distribution. *Expert Systems with Applications*, 39(3), 3668–3675. <https://doi.org/10.1016/j.eswa.2011.09.059>

Zhang, X., Yang, Y., & Zhou, Z. (2018). A Novel Credit Scoring Model based on Optimized Random Forest. Computing and Communicating Workshop and Conference (CCWD), 2018 IEEE 8th Annual, 978(1), 60–65.