



Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring

Omer Heranova

Program Pascasarjana, Ilmu Komputer, STMIK Nusa Mandiri
heranova.omer@gmail.com

Abstract

Bank or financial institution is a business entity whose activities are collecting funds from the public in the form of deposits and channeling them to the public in the form of credit and or other forms. In credit financing problems often occur and one of the problems faced in credit assessment is imbalance class data sets or dataset class imbalances. This problem can be overcome by resampling method, namely by using Oversampling, undersampling and hybrids that combine the two sampling approaches. This research proposes the method of applying SMOTE or Synthetic Minority Oversampling Technique on Averaged One Dependence estimators (AODE) to improve the performance of the accuracy of the credit rating classification on German Credit Creditsets. The results of this experimental study on the GermanCredit dataset with the classification method without the Resampling process on 13 classifiers produce an average performance value of 70%. The results of the classification with classification techniques that apply the SMOTE method on the AODE algorithm can increase the accuracy performance by 5.5% with an accuracy value of 0.817 or 81.69%. While the classification technique that applies the SpreadSubSample + AODE method decreased by 0.041 or 4.1% but still higher than the accuracy value of other methods with an accuracy value of 0.723 or 72.33%. The researcher concludes that by applying the Resampling technique with the SMOTE method on the AODE algorithm can increase the value of accuracy performance effectively on the imbalance class used for credit scoring or credit rating on GermanCredit datasets.

Keywords: imbalance class, credit scoring, resampling, synthetic minority oversampling technique, spreadsubsample

Abstrak

Bank adalah badan usaha yang kegiatannya menghimpun dana masyarakat dalam bentuk tabungan serta menyalurkannya ke dalam angsuran kredit atau bentuk lainnya. Dalam pembiayaan kredit sering terjadi masalah dan permasalahan pada *credit scoring* atau penilaian kredit yaitu adanya *imbalance class datasets* atau ketidakseimbangan kelas dataset. Problem ini dapat diatasi dengan metode *resampling*, yaitu dengan menggunakan *oversampling*, *undersampling* dan *hybrid* (menggabungkan kedua pendekatan *sampling*). Penelitian ini mengusulkan metode penerapan *SMOTE* atau *Synthetic Minority Oversampling Technique* pada *Averaged One Dependence estimators (AODE)* untuk meningkatkan kinerja akurasi klasifikasi *credit scoring* atau penilaian kredit pada *GermanCredit datasets*. Hasil penelitian eksperimen ini pada dataset *GermanCredit* dengan metode klasifikasi tanpa proses *Resampling* pada 13 pengklasifikasi menghasilkan nilai kinerja *accuracy* atau akurasi rata-rata 70%. Hasil klasifikasi dengan teknik klasifikasi yang menerapkan metode *SMOTE* pada *AODE* dapat meningkatkan kinerja akurasi sebesar 5,5% dengan nilai akurasi 0,817 atau 81,69%. Sedangkan klasifikasi yang menerapkan metode *SpreadSubSample + AODE* mengalami penurunan 0,041 atau 4,1% namun masih lebih tinggi dari nilai akurasi metode lainnya dengan nilai akurasi sebesar 0,723 atau 72,33%. Peneliti menyimpulkan bahwa dengan menerapkan teknik *Oversampling* dengan metode *SMOTE* pada *AODE* dapat meningkatkan nilai kinerja akurasi dengan efektif pada klasifikasi *imbalance class* atau kelas tidak seimbang yang digunakan untuk *credit scoring* atau penilaian kredit pada *datasets GermanCredit*.

Kata Kunci: imbalance class, credit scoring, resampling, synthetic minority oversampling technique, spreadsubsample.

© 2019 Jurnal RESTI

1. Pendahuluan

UU Perbankan No. X tahun 1998 menerangkan bahwa bank adalah perusahaan yang kegiatannya menghimpun dana masyarakat dalam bentuk tabungan dan

menyalurkannya kepada masyarakat dalam angsuran kredit dan atau bentuk lainnya dalam rangka meningkatkan kesejahteraan masyarakat. Berdasarkan UU tersebut, segala bentuk pembiayaan kredit yang

Diterima Redaksi : 20-09-2019 | Selesai Revisi : 26-11-2019 | Diterbitkan Online : 10-12-2019

dilakukan harus berdasarkan pada persetujuan pinjam meminjam, dan akan ada suatu analisis perhitungan yang dilakukan untuk menentukan sebuah pengambilan keputusan [1].

Dalam pembiayaan kredit sering terjadi masalah kredit dan permasalahan yang dihadapi pada *credit scoring* atau penilaian kredit yaitu adanya *imbalance datasets* atau ketidakseimbangan *dataset*. Problem ini dapat di atasi dengan metode *resampling*, yaitu dengan menggunakan *oversampling*, *undersampling* dan *hybrid* yang menggabungkan kedua pendekatan *sampling*. Penelitian ini mengusulkan metode penerapan metode *Synthetic Minority Oversampling Technique (SMOTE)* pada *Averaged One Dependence estimators (AODE)* untuk meningkatkan kinerja akurasi klasifikasi *credit scoring* atau penilaian kredit pada *GermanCredit datasets*. Tujuan *credit scoring* pada dasarnya adalah untuk mengklasifikasikan aplikasi pinjaman menjadi dua kelas, yaitu, pembayar yang baik atau “good” (yaitu, mereka yang cenderung untuk menjaga secara rutin pembayaran mereka) dan pembayar yang buruk atau “bad” (yaitu, mereka yang cenderung gagal bayar pada pinjaman mereka). *Credit scoring* atau penilaian kredit tersebut berdasarkan dari kemampuan calon nasabah untuk membayar kewajiban pembayarannya dengan melihat karakteristik seperti *gender*, usia, pekerjaan, dan gaji dan lain-lain [2].

Machine learning adalah sebuah program algoritma untuk memproses data, menemukan pola dalam kumpulan data dan menggunakannya untuk membuat keputusan, memprediksi dan mengambil keputusan. Atau pemrograman komputer untuk mengoptimalkan kinerja kriteria menggunakan data atau data pada masa lalu [3].

Credit Scoring menggunakan 2 (dua) metode klasifikasi, yaitu dengan metode statistik dan *machine learning* [2]. Metode statistik yang banyak digunakan dalam *creditscoring* atau penilaian kredit diantaranya adalah *Linear Discriminant Analysis (LDA)*, *Logistic Regression Analysis (LRA)* [4]. Selain itu *Machine Learning* dengan menggunakan *Artificial Intelligence (AI)* seperti *Artificial Neural Networks (ANN)*, *Decision tree (DT)*, *Case-Based Reasoning (CBR)*, *Support Vector Machine (SVM)* juga banyak digunakan dalam penelitian tentang *credit scoring* atau penilaian kredit [5].

Permasalahan dalam penggunaan *machine learning* yaitu *imbalance class* atau ketidakseimbangan kelas. *Imbalanceclass* atau ketidakseimbangan kelas yang terjadi pada kelas minoritas yang lebih kecil atau lebih jarang dari kelas mayoritas [6]. Model dibuat dengan menggunakan data yang tidak seimbang akan menghasilkan akurasi prediksi minoritas yang rendah. Informasi yang lebih banyak dari kelas mayoritas mendominasi kelas minoritas sehingga menyebabkan

batas-batas keputusan yang diragukan dalam sistem klasifikasi [7].

Metode yang digunakan dalam mengatasi *imbalanceclass* atau ketidakseimbangan kelas dapat dibagi menjadi tiga kategori. Pertama dengan menggunakan metode tingkat data yang menyeimbangkan distribusi data dengan metode *oversampling* dan *undersampling*. Kedua dengan pendekatan tingkatan algoritma yaitu mengembangkan algoritma baru atau modifikasi metode yang sudah ada dengan memperhitungkan arti dari kelas minor. Ketiga mengkombinasikan pendekatan algoritma dan pendekatan *level data* [8].

Dalam mengatasi masalah *imbalance class* atau ketidakseimbangan kelas, penelitian eksperimen ini menggunakan pendekatan *level data* dengan metode *SMOTE* pada *(AODE)* dimana kinerja akurasi pada klasifikasi *creditscoring* atau penilaian kredit tersebut dapat ditingkatkan. *Creditscoring* atau penilaian kredit adalah sebagai proses pengakuan nasabah bank untuk memberikan kredit berdasarkan seperangkat nilai kriteria tertentu. *Creditscoring* atau 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 yang minim [9].

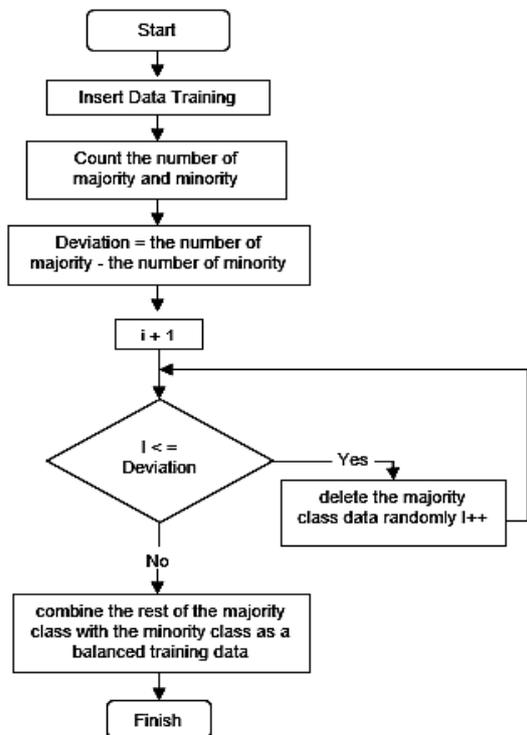
Credit scoring atau penilaian kredit bertujuan untuk mengklasifikasi para calon nasabah yang diklasifikasikan menjadi dua kriteria yaitu calon nasabah yang baik atau “good” dan calon nasabah yang buruk atau “bad”. *Credit scoring* didasarkan dari kemampuan calon nasabah untuk membayar kewajiban hutangnya. Calon nasabah dengan *credit scoring* yang bagus akan memiliki kemungkinan besar untuk membayar kewajiban hutangnya. Sedangkan calon nasabah dengan *credit scoring* yang buruk akan memiliki kemungkinan gagal dalam membayar kewajiban hutangnya. Keakuratan pemberian *credit scoring* sangatlah diperlukan untuk menjaga *performa* bank atau lembaga keuangan. Dengan adanya peningkatan nilai *accuracy credit scoring* terhadap calon nasabah dengan resiko gagal bayar atau macet maka dapat menurunkan kerugian bagi bank atau lembaga keuangan [5].

Klasifikasi merupakan proses menemukan model atau fungsi yang menggambarkan, membedakan *class* data atau konsep dengan tujuan agar bisa digunakan untuk *class prediction* dari objek yang *label class* tidak diketahui. Klasifikasi banyak digunakan untuk mendeteksi *fraud* atau penipuan, target pemasaran, prediksi kinerja, manufaktur, dan mendiagnosa kesehatan [10]. Tahapan klasifikasi data terdiri dari 2 langkah. Pertama yaitu tahap *learning* atau fase pembelajaran, dimana algoritma klasifikasi untuk

menganalisis data *training* atau data latih lalu direpresentasikan ke bentuk model klasifikasi. Kedua adalah proses tahapan klasifikasi, dimana *datatesting* digunakan untuk memprediksi nilai *accuracy* dari model klasifikasi. Jika nilai akurasi *acceptable* atau dapat diterima, maka *rule* bisa diterapkan pada klasifikasi *tupel* data baru [10].

Imbalance Class atau ketidakseimbangan kelas merupakan permasalahan pada data mining. Ini terjadi saat *class* minoritas yang jauh lebih kecil atau jarang dari *class* mayoritas [6]. Model yang dibuat dari data yang tidak seimbang akan menghasilkan akurasi prediksi minoritas yang rendah. Informasi yang lebih banyak dari kelas mayoritas akan mendominasi kelas minoritas sehingga menyebabkan batas keputusan yang diragukan dalam sistem klasifikasi [7]. Selain itu *Imbalance Class* atau ketidakseimbangan kelas serta *noise* juga dapat mempengaruhi kualitas data dalam kinerja klasifikasi.

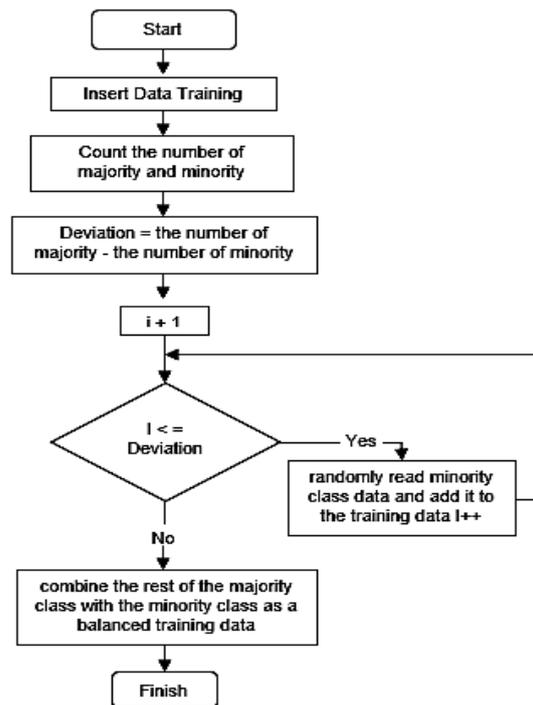
Teknik *resampling* adalah *preprocessing* yang mana menyeimbangkan distribusi data kembali untuk mengurangi efek distribusi kelas yang tidak seimbang dalam proses *data training* [7]. Teknik *resampling* digunakan untuk mengatasi masalah data tidak seimbang. Metode ini menyeimbangkan data yang asli berdasarkan serangkaian metode algoritma *sampling* menyesuaikan jumlah sampel kelas yang berbeda, kemudian melatih data seimbang baru yang mengadopsi algoritma klasifikasi.



Gambar 1. Flowchart Algoritma Random Under Sampling Sumber [12]

Metode *resampling* dibagi menjadi tiga kategori yaitu Pertama dengan Metode *OverSampling*, Kedua *UnderSampling*, dan Ketiga yaitu dengan *Hybrid* yang menggabungkan kedua Metode *Sampling* [7]. Teknik *resampling* yang digunakan yaitu *UnderSampling* yang secara acak memilih sampel kelas mayoritas dan menambahkannya ke dalam kelas minoritas sehingga membentuk sebuah *datasets training* baru. Metode *OverSampling* bertujuan meningkatkan sampel kelas minoritas sampai *equal* atau setara dengan kelas mayoritas lain yang menduplikasi secara acak sampel kelas minoritas [2]. Sedangkan Metode *UnderSampling* menghasilkan *subsampel* acak dari *instance* kelas mayoritas [11]. Metode *UnderSampling* secara acak memilih sampel kelas mayoritas dan menambahkannya ke kelas minoritas sehingga membentuk sebuah *datasets training* baru (1). [2].

Metode *OverSampling* yang mana sampel dari kelas minoritas dipilih secara acak dan diduplikasi, *instance* yang dihasilkannya meningkatkan besar jumlah kelas minoritas dengan mereplikasi informasi yang sama (2). Salah satu Teknik *OverSampling* yang populer adalah Teknik *SMOTE*.



Gambar 2. Flowchart Algoritma Random Over Sampling Sumber [12]

NaïveBayes merupakan *classifier* yang berdasarkan probabilitas dan Teorema Bayesian dengan asumsi setiap atribut bersifat berdiri sendiri atau *independence*. Namun, kondisi atribut *independent* pada *Naïve Bayes* sangat jarang terjadi. Sehingga, berkembanglah *AODE* yang dapat mengatasi kelemahan dalam penanganan ketergantungan atribut. Pada *AODE*, setiap atribut bergantung kepada kelas dan atribut tunggal lainnya.

AODE dapat mengurangi kerugian dalam hasil akurasi akibat penggunaan asumsi atribut saling *independent*. *AODE* memiliki tingkat akurasi tinggi dan *error rate* yang rendah dalam klasifikasi dan memiliki pembelajaran dan *efficiency* klasifikasi yang tinggi.

AODE adalah teknik pembelajaran klasifikasi probabilistik. Dikembangkan untuk mengatasi masalah atribut-atribut dari *classifier Naive Bayes*. Ini merupakan pengembangan pengklasifikasi yang jauh lebih akurat daripada *Naive Bayes* dengan peningkatan jumlah komputasi yang sederhana [13].

AODE mencapai klasifikasi yang sangat akurat dengan merata-ratakan semua ruang kecil model alternatif seperti *Naive* yang memiliki asumsi independensi yang lebih lemah daripada *Naive Bayes*. Algoritma yang dihasilkan efisien secara komputasi sambil memberikan klasifikasi yang sangat akurat pada banyak tugas pembelajaran.

AODE berupaya memperkirakan probabilitas setiap kelas y dengan serangkaian fitur x_1, \dots, x_n , $P(y | x_1, \dots, x_n)$ yang ditentukan. Untuk melakukannya menggunakan rumus (1):

$$\hat{P}(y|x_1, \dots, x_n) = \sum_{i: 1 \leq i \leq n} F(x_i) \geq m \hat{P}(y, x_i) \prod_{j=1}^n \hat{P}(x_j|y, x_i) / \sum_{y \in Y} \epsilon Y \sum_{i: 1}^n F(x_i) \geq m \hat{P}(y, x_i) \prod_{j=1}^n \hat{P}(x_j|y, x_i) \quad (1)$$

di mana $\hat{P}(\cdot)$ menunjukkan perkiraan $P(\cdot)$, $F(\cdot)$ adalah frekuensi dengan mana argumen muncul dalam data sampel dan m adalah frekuensi minimum yang ditentukan pengguna dengan istilah yang harus muncul untuk digunakan dalam penjumlahan luar. Dalam praktik baru-baru ini m biasanya ditetapkan pada angka 1.

Peneliti berupaya memperkirakan $P(y | x_1, \dots, x_n)$. Dengan definisi probabilitas bersyarat (2):

$$P(y|x_1, \dots, x_n) = \frac{P(y, x_1, \dots, x_n)}{P(x_1, \dots, x_n)} \quad (2)$$

Untuk semua $1 \leq i \leq n$

$$P(y, x_1, \dots, x_n) = P(y, x_i, \dots, x_n | y, x_i)$$

Dengan asumsi bahwa x_1, \dots, x_n bersifat independen karena y dan x_i , maka itu adalah sebagai berikut (3):

$$P(y, x_1, \dots, x_n) = P(y, x_i) \prod_{j=1}^n P(x_j | y, x_i) \quad (3)$$

Rumus ini mendefinisikan bentuk khusus *One Dependence Estimator (ODE)*, varian dari *classifier Naive Bayes* yang membuat asumsi independensi di atas menjadi lebih lemah (dan karenanya berpotensi sangat kurang) daripada asumsi independensi *Naive Bayes*. Karena itu, setiap *ODE* harus membuat penduga yang kurang bias dibandingkan *Naive Bayes*. Namun, karena estimasi probabilitas dasar masing-masing dikondisikan oleh dua variabel daripada satu, mereka dibentuk dari data yang lebih sedikit (contoh pelatihan

yang memenuhi kedua variabel) dan karenanya cenderung memiliki lebih banyak varian. *AODE* mengurangi varians ini dengan rata-rata estimasi semua *ODE* tersebut [13].

Fitur-fitur dari classifier *AODE*

Seperti *Naive Bayes*, *AODE* tidak melakukan pemilihan model dan tidak menggunakan parameter yang dapat disetel. Akibatnya, memiliki varian rendah. Ini mendukung pembelajaran tambahan di mana pengklasifikasi dapat diperbarui secara efisien dengan informasi dari contoh-contoh baru saat tersedia. Metode ini memprediksi probabilitas *class* daripada hanya memprediksi satu *class*, yang memungkinkan pengguna untuk menentukan kepercayaan dengan mana setiap klasifikasi dapat dibuat. Model probabilistiknya dapat langsung mengatasi situasi di mana beberapa data yang hilang.

AODE memiliki kompleksitas komputasi $O(\ln^2)$ pada waktu pelatihan dan $O(kn^2)$ pada klasifikasi waktu, di mana n adalah jumlah fitur, l adalah jumlah contoh pelatihan dan k adalah jumlah kelas. Ini membuatnya tidak layak untuk aplikasi ke data dimensi tinggi. Namun, dalam batasan itu, linier sehubungan dengan jumlah contoh pelatihan dan karenanya dapat secara efisien memproses sejumlah besar contoh pelatihan.

Pada penelitian eksperimen sebelumnya telah mengusulkan dengan menerapkan metode *Random Over-Under Sampling Random Forest* untuk memecahkan masalah *Imbalance Class* atau data tidak seimbang pada klasifikasi kredit.

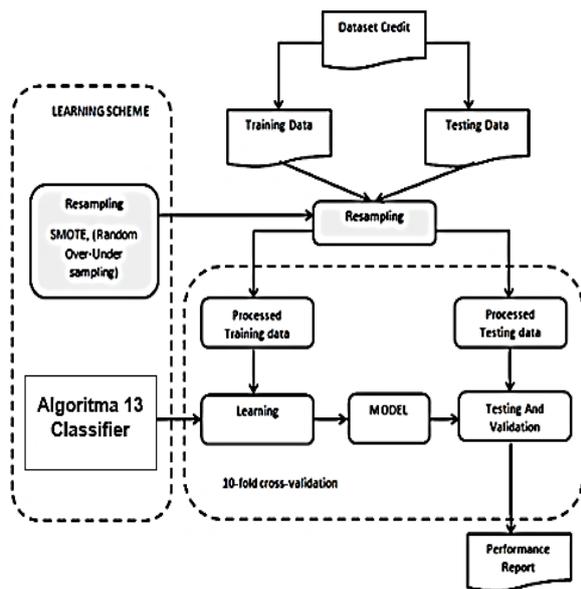
Penelitian tersebut telah melakukan klasifikasi tanpa proses *resampling* pada 12 pengklasifikasi yang menghasilkan kinerja akurasi rata-rata 70%. Hasil klasifikasi menjelaskan bahwa metode *Random Forest* memiliki nilai akurasi yang lebih baik dibandingkan dengan beberapa metode lainnya dengan nilai *accuracy* sebesar 0,76 (76%). Sedangkan klasifikasi dengan menerapkan metode *SMOTE* pada algoritma *Random Forest* dapat meningkatkan kinerja akurasi sebesar 0,48% dengan nilai *accuracy* 0,808 (80,8%). Sedangkan klasifikasi dengan menerapkan metode *Random Over-under sampling + Random Forest* dapat meningkatkan kinerja akurasi sebesar 14,1% dengan nilai akurasi sebesar 0,901 (90,1%). Penelitian tersebut menyimpulkan bahwa dengan menerapkan *resampling* dengan metode *Random Over-Under Sampling* pada algoritma *Random Forest* dapat meningkatkan kinerja akurasi dengan efektif pada klasifikasi tidak seimbang untuk *creditscoring* atau penilaian kredit pada dataset *German Credit*.

2. Metode Penelitian

Pada penelitian ini yang dilakukan adalah jenis eksperimen komparasi yang akan membandingkan dua objek yang berbeda, misalnya dengan menganalisis

nilai statistik yang dihasilkan dari masing-masing mana yang unggul dengan menyelidiki hubungan kausal dengan menggunakan tes yang dikontrol sendiri [14].

Penelitian mempunyai tujuan untuk melakukan komparasi serta mengevaluasi model *ensemble* dengan menggunakan *Averaged One Dependence estimators* atau *AODE* dan metode *resampling* untuk analisis algoritma mana yang memiliki nilai akurasi lebih tinggi pada klasifikasi *imbalance class* atau data tidak seimbang dalam *credit scoring* atau penilaian kredit. *Datasets* yang digunakan merupakan data sekunder yang didownload dari *UCI (University of California, Irvine) Machine Learning Repository*.



Gambar 3. Metode Usulan,
Sumber : Penelitian (2019)

Pada penelitian ini pengumpulan data dengan menggunakan data sekunder yang di download dari *University of California Irvine machine learning data repository* 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)).

Model yang diusulkan oleh peneliti untuk mengatasi masalah *imbalance class* atau ketidakseimbangan kelas yaitu dengan menerapkan kombinasi teknik pendekatan level data metode *resampling* dan menggunakan pendekatan algoritma dengan metode *Averaged One Dependence estimators* atau *AODE*. Gambar *Flowchart model* yang diusulkan pada Gambar 3.

3. Hasil dan Pembahasan

Pada bab ini akan membahas hasil penelitian eksperimen yang sudah dilakukan dengan menggunakan aplikasi *WEKA 3.8.3. Datasets* yang digunakan pada penelitian atau eksperimen ini adalah menggunakan *datasets GermanCredit* yang terdiri dari 21 *atribut* dan 1000 *instance* serta 2 kelas atau *class* yaitu kelas atau *class* “good” atau “baik” dan kelas atau *class* “bad” atau “buruk”. Dengan jumlah *instance* masing-masing kelas atau *class* yaitu 700 untuk kelas atau *class* “good” atau “baik”, dan 300 *instance* untuk kelas atau *class* “bad” atau “buruk”. Pengaturan eksperimen penelitian dilakukan untuk menghasilkan nilai akurasi atau *accuracy* yang paling tinggi dengan metode yang diusulkan yaitu membandingkan teknik metode *AODE* dengan beberapa teknik metode lainnya dengan menerapkan teknik *resampling* seperti *SMOTE* dan *SpreadSubSample* ataupun tanpa teknik *resampling*.

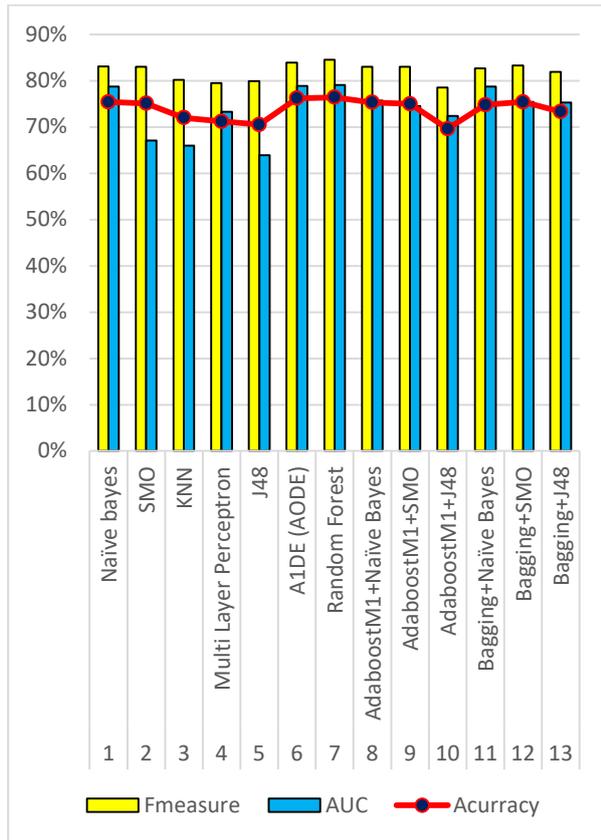
3.1. Klasifikasi *Averaged One Dependence Estimators*

Pada tahap ini penelitian eksperimen dilakukan tanpa menggunakan teknik *resampling*. Melainkan langsung dilakukan pemodelan dan *testing* atau pengujian menggunakan *AODE*. Hasil *testing* atau pengujian dibandingkan dengan 12 teknik lainnya untuk menghasilkan model dengan nilai kinerja akurasi yang terbaik atau tertinggi. Hasil *testing* atau pengujian menggunakan model klasifikasi tanpa teknik *resampling* dijelaskan pada tabel 1.

Tabel 1. Perbandingan nilai kinerja model Pengklasifikasi tanpa *Resampling*

| NO | ALGORITMA | TPRate | FPRate | Precision | Recall | Accuracy | FMeasure | AUC |
|----|------------------------|--------|--------|-----------|--------|----------|----------|-------|
| 1 | Naïve bayes | 0,864 | 0,503 | 0,800 | 0,864 | 0,754 | 0,831 | 0,787 |
| 2 | SMO | 0,871 | 0,530 | 0,793 | 0,871 | 0,751 | 0,830 | 0,671 |
| 3 | KNN | 0,810 | 0,490 | 0,794 | 0,810 | 0,720 | 0,802 | 0,660 |
| 4 | Multi Layer Perceptron | 0,797 | 0,487 | 0,793 | 0,797 | 0,712 | 0,795 | 0,733 |
| 5 | J48 | 0,840 | 0,610 | 0,763 | 0,840 | 0,705 | 0,799 | 0,639 |
| 6 | AIDE (AODE) | 0,884 | 0,523 | 0,798 | 0,884 | 0,762 | 0,839 | 0,789 |
| 7 | Random Forest | 0,917 | 0,593 | 0,783 | 0,917 | 0,764 | 0,845 | 0,791 |
| 8 | AdaboostM1+Naïve Bayes | 0,859 | 0,493 | 0,802 | 0,859 | 0,753 | 0,830 | 0,757 |
| 9 | AdaboostM1+SMO | 0,871 | 0,533 | 0,792 | 0,871 | 0,750 | 0,830 | 0,745 |
| 10 | AdaboostM1+J48 | 0,791 | 0,527 | 0,778 | 0,791 | 0,696 | 0,785 | 0,724 |
| 11 | Bagging+Naïve Bayes | 0,859 | 0,510 | 0,797 | 0,859 | 0,748 | 0,827 | 0,787 |
| 12 | Bagging+SMO | 0,876 | 0,530 | 0,794 | 0,876 | 0,754 | 0,833 | 0,754 |
| 13 | Bagging+J48 | 0,863 | 0,570 | 0,779 | 0,863 | 0,733 | 0,819 | 0,753 |

Pada Tabel 1 menunjukkan bahwa model *Averaged One Dependence estimators* atau *AODE* memiliki nilai akurasi atau *accuracy* dan *F-Measure* yang hampir menyamai dengan model *Random Forest* yang sama-sama lebih tinggi dibandingkan dengan beberapa model lainnya dengan nilai *accuracy AODE* sebesar 0,76, nilai *F-Measure* sebesar 0,84 dan nilai *accuracy Random Forest* sebesar 0,76, nilai *F-Measure* 0,85. Grafik perbandingan nilai kinerja 13 model pengklasifikasi dijelaskan pada Gambar 4.



Gambar 4. Perbandingan nilai kinerja permodelan tanpa Resampling

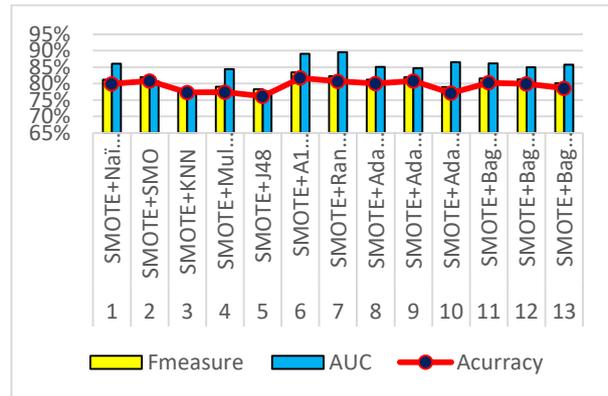
3.2. Klasifikasi SMOTE AIDE (AODE)

Penelitian *Experiment* dilakukan dengan Metode SMOTE yaitu data latih atau *training* menjadi lebih seimbang antara kelas "baik" atau "good" dengan kelas "buruk" atau "bad" dengan jumlah *instance* pada kelas "baik" atau "good" 700 *instance*, sedangkan kelas "buruk" atau "bad" menjadi 600 *instance*. Hasil *testing* pengujian dibandingkan dengan 12 metode lainnya untuk menghasilkan model dengan kinerja *accuracy* terbaik. Hasil pengujian teknik model pengklasifikasi dengan SMOTE atau *Synthetic Minority Oversampling Technique* dijelaskan pada Tabel 2.

Tabel 2. Perbandingan nilai kinerja model Pengklasifikasi dengan metode Resampling SMOTE

| NO | ALGORITMA | TPRate | FPRate | Precision | Recall | Accuracy | FMeasure | AUC |
|----|------------------------------|--------|--------|-----------|--------|----------|----------|-------|
| 1 | SMOTE+Naïve bayes | 0.807 | 0.210 | 0.818 | 0.807 | 0.799 | 0.812 | 0.860 |
| 2 | SMOTE+S | 0.813 | 0.198 | 0.827 | 0.813 | 0.808 | 0.820 | 0.807 |
| 3 | SMOTE+KNN | 0.744 | 0.193 | 0.818 | 0.744 | 0.773 | 0.779 | 0.775 |
| 4 | SMOTE+Multi Layer Perceptron | 0.797 | 0.253 | 0.786 | 0.797 | 0.774 | 0.791 | 0.844 |
| 5 | SMOTE+J48 | 0.799 | 0.282 | 0.768 | 0.799 | 0.762 | 0.783 | 0.777 |
| 6 | SMOTE+A1DE (AODE) | 0.854 | 0.227 | 0.815 | 0.854 | 0.817 | 0.834 | 0.891 |
| 7 | SMOTE+Random Forest | 0.836 | 0.227 | 0.811 | 0.836 | 0.807 | 0.823 | 0.895 |
| 8 | SMOTE+AdaboostM1+Naive Bayes | 0.803 | 0.203 | 0.822 | 0.803 | 0.800 | 0.812 | 0.851 |
| 9 | SMOTE+AdaboostM1+S | 0.813 | 0.198 | 0.827 | 0.813 | 0.808 | 0.820 | 0.847 |
| 10 | SMOTE+AdaboostM1+J48 | 0.801 | 0.265 | 0.779 | 0.801 | 0.771 | 0.790 | 0.865 |
| 11 | SMOTE+Bagging+Naive Bayes | 0.813 | 0.210 | 0.819 | 0.813 | 0.802 | 0.816 | 0.861 |
| 12 | SMOTE+Bagging+S | 0.813 | 0.217 | 0.814 | 0.813 | 0.799 | 0.813 | 0.850 |
| 13 | SMOTE+Bagging+J48 | 0.803 | 0.235 | 0.799 | 0.803 | 0.785 | 0.801 | 0.858 |

Pada Tabel 2 menjelaskan bahwa dengan menerapkan metode *Resampling SMOTE* menghasilkan nilai kinerja yang lebih baik dibandingkan dengan tanpa menggunakan metode *Resampling*. Terbukti dengan adanya peningkatan *accuracy* pada semua model yang diuji. Hasil nilai kinerja yang diperoleh membuktikan bahwa model SMOTE+AODE memiliki nilai *accuracy* yang lebih baik dengan beberapa model lainnya dengan nilai *accuracy* sebesar 0,817 atau 81,69% yang mengalami kenaikan 0,055 atau 5,492%. Grafik perbandingan nilai kinerja 13 model pengklasifikasi disajikan pada gambar 5.



Gambar 5. Perbandingan kinerja Permodelan dengan Metode SMOTE

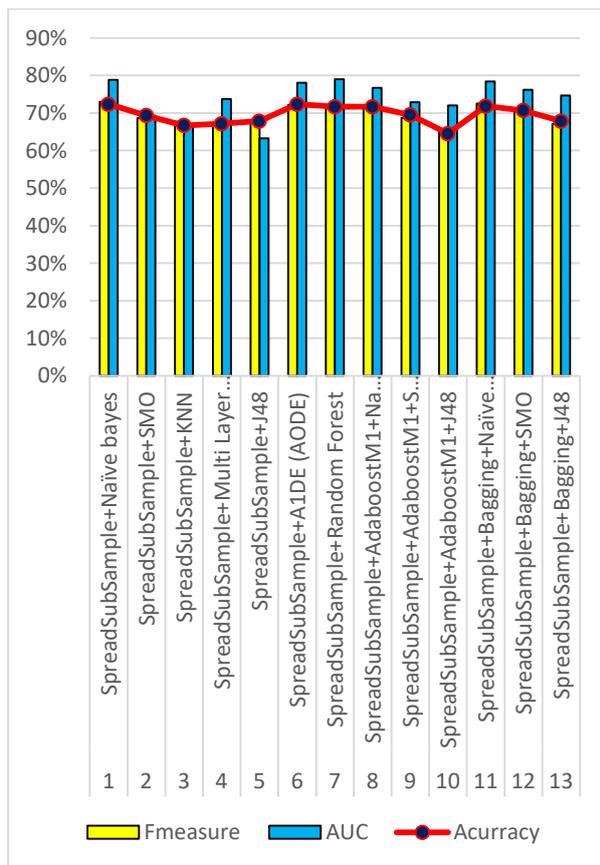
3.3. Klasifikasi SpreadSubSample AODE

Penelitian *Experiment* ini menerapkan teknik *SpreadSubSample* pada metode *Averaged One Dependence estimators* atau AODE. Di mana data latih atau *training* menjadi seimbang antara kelas "baik" atau "good" dengan kelas "buruk" atau "bad" dengan jumlah *instance* masing-masing kelas atau *class* menjadi 600 *instance*. Hasil pengujian akan dibandingkan dengan teknik 12 metode lainnya yang menghasilkan model dengan nilai kinerja akurasi terbaik. Hasil pengujian model pengklasifikasi dengan teknik *SpreadSubSample* disajikan pada tabel 3.

Tabel 3. Perbandingan hasil nilai kinerja model Pengklasifikasi teknik SpreadSubSample

| NO | ALGORITMA | TPRate | FPRate | Precision | Recall | Accuracy | FMeasure | AUC |
|----|--|--------|--------|-----------|--------|----------|----------|-------|
| 1 | SpreadSubSample+Naive bayes | 0.747 | 0.300 | 0.713 | 0.747 | 0.723 | 0.730 | 0.788 |
| 2 | SpreadSubSample+S | 0.670 | 0.283 | 0.703 | 0.670 | 0.693 | 0.686 | 0.693 |
| 3 | SpreadSubSample+KNN | 0.700 | 0.367 | 0.656 | 0.700 | 0.667 | 0.677 | 0.667 |
| 4 | SpreadSubSample+Multi Layer Perceptron | 0.660 | 0.317 | 0.676 | 0.660 | 0.672 | 0.668 | 0.737 |
| 5 | SpreadSubSample+J48 | 0.667 | 0.310 | 0.683 | 0.667 | 0.678 | 0.675 | 0.633 |
| 6 | SpreadSubSample+A1DE (AODE) | 0.720 | 0.273 | 0.725 | 0.720 | 0.723 | 0.722 | 0.781 |
| 7 | SpreadSubSample+Random Forest | 0.703 | 0.270 | 0.723 | 0.703 | 0.717 | 0.713 | 0.790 |
| 8 | SpreadSubSample+AdaboostM1+Naive Bayes | 0.720 | 0.287 | 0.715 | 0.720 | 0.717 | 0.718 | 0.767 |
| 9 | SpreadSubSample+AdaboostM1+S | 0.670 | 0.280 | 0.705 | 0.670 | 0.695 | 0.687 | 0.729 |
| 10 | SpreadSubSample+AdaboostM1+J48 | 0.660 | 0.370 | 0.641 | 0.660 | 0.645 | 0.650 | 0.720 |
| 11 | SpreadSubSample+Bagging+Naive Bayes | 0.743 | 0.307 | 0.708 | 0.743 | 0.718 | 0.725 | 0.784 |
| 12 | SpreadSubSample+Bagging+S | 0.710 | 0.297 | 0.705 | 0.710 | 0.707 | 0.708 | 0.762 |
| 13 | SpreadSubSample+Bagging+J48 | 0.657 | 0.300 | 0.686 | 0.657 | 0.678 | 0.671 | 0.747 |

Pada Tabel 3 menjelaskan dengan menerapkan teknik metode *SpreadSubSample* pada *datasets* sebelum dilakukan pemodelan menghasilkan nilai kinerja yang lebih rendah dibandingkan jika tanpa menggunakan metode *resampling*. Nilai kinerja yang didapat menunjukkan bahwa model teknik pengklasifikasi metode *SpreadSubSample+AODE* menghasilkan nilai akurasi yang lebih baik dibandingkan dengan beberapa model lainnya dengan nilai *accuracy* sebesar 0,723 atau 72,33% yang mengalami penurunan 0,041 atau 4,1%. Grafik perbandingan kinerja 13 model pengklasifikasi disajikan pada gambar 6.



Gambar 6. Perbandingan kinerja Permodelan dengan Metode *SpreadSubSample*

4. Kesimpulan

Berdasarkan hasil penelitian *experiment* ini pada *datasets GermanCredit*. Metode klasifikasi tanpa proses *Resampling* pada 13 pengklasifikasi menghasilkan nilai kinerja *accuracy* atau akurasi rata-rata 70%. Hasil klasifikasi menunjukkan bahwa metode *RandomForest* memiliki nilai *accuracy* atau akurasi yang lebih baik dibandingkan dengan beberapa metode lainnya dengan nilai *accuracy* sebesar 0,764 (76%). Sedangkan dengan metode klasifikasi yang menerapkan metode *SMOTE* pada *AODE* dapat meningkatkan kinerja akurasi sebesar 5,5% dengan nilai akurasi 0,817

atau 81,69%. Sedangkan teknik klasifikasi yang menerapkan metode *SpreadSubSample + AODE* mengalami penurunan 0,041 atau 4,1% namun masih lebih tinggi dari nilai akurasi metode lainnya dengan nilai *accuracy* atau akurasi sebesar 0,723 atau 72,33%.

Peneliti menyimpulkan bahwa dengan menerapkan teknik *Resampling* dengan metode *SMOTE* pada *AODE* dapat meningkatkan nilai kinerja akurasi dengan efektif pada klasifikasi *imbalance class* atau kelas tidak seimbang yang digunakan untuk *credit scoring* atau penilaian kredit pada *datasets GermanCredit*.

Untuk penelitian eksperimen selanjutnya ada beberapa hal berikut yang bisa ditambahkan supaya bisa menghasilkan nilai *accuracy* atau akurasi yang lebih baik sebagai berikut:

1. Pada penelitian eksperimen selanjutnya dapat menerapkan teknik metode lainnya seperti metode seleksi fitur atau *feature selection* yang dapat meningkatkan nilai *accuracy* atau akurasi klasifikasi *creditscoring* atau penilaian kredit kelas tidak seimbang atau *imbalance class*.
2. Menggunakan *datasets* yang jumlah instance-nya lebih banyak, dan nilai rasio *imbalance class* atau kelas tidak seimbang yang lebih tinggi.
3. Mencoba dan bereksperimen untuk menerapkan teknik metode *ensemble learning* lain dan dengan dasar *base classifier* yang lain dan berbeda.

Daftar Rujukan

- [1] Peraturan OJK No.29/POJK.05/2014. Tersedia di <https://www.ojk.go.id/id/kanal/perbankan/Pages/Bank-Umum.aspx>. [Accessed 22Agust 2019].
- [2] He, H, Zhang, W, & Zhang, S.(2018). *A novel ensemble method for creditscoring: Adaption of different imbalance ratios*. *Expert Systems with Applications*, 98, 105–117. <https://doi.org/10.1016/j.eswa.2018.01.012>.
- [3] Ethem Alpaydin, (2014). *Introduction to Machine Learning*, 3rd ed., MIT Press.
- [4] Zhang, X, Yang, Y, & Zhou, Z. (2018). *A Novel CreditScoring Model based on Optimized Random Forest*. *Computing and Communicating Workshop and Conference (CCWD), 2018 IEEE 8th Annual*, 978(1), 60–65.
- [5] 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>.
- [6] Ren, F, Cao, P, Li, W, Zhao, D, & Zaiane, O. (2017). *Ensemble based adaptive oversampling method for imbalance 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>.
- [7] Jian, C, Gao, J, & Ao, Y. (2016). *A new sampling method for classifying imbalance data based on supportvector machine ensemble*. *Neurocomputing*, 193, 115–122. <https://doi.org/10.1016/j.neucom.2016.02.006>.
- [8] Xiao, J., Xie, L., He, C., & Jiang, X. (2012). *Dynamic classifier ensemble model for customer classification with imbalance class distribution*. *ExpertSystems with Applications*, 39(3), 3668–3675. <https://doi.org/10.1016/j.eswa.2011.09.059>.

- [9] Koutanaei, F. N, Sajedi, H, & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for creditscoring. *Journal of Retailing and Consumer Services*, 27, 11–23. <https://doi.org/10.1016/j.jretconser.2015.07.003>.
- [10] Han, J, Kamber, M., Pei, J,(2012). *Data Mining Concept And Techniques*. California: Morgan Kaufmann.
- [11] Rajesh, K.N.V.P.S, & Dhuli, R. (2018). Classification of imbalance ECG beats using resampling techniques and AdaBoost ensemble classifier. *Biomedical Signal Processing and Control*, 41,242-254. <https://doi.org/10.1016/j.bspc.2017.12.004>.
- [12] 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.
- [13] Webb, G. I., J. Boughton, and Z. Wang (2005). "Not So Naive Bayes: Aggregating One-Dependence Estimators".
- [14] Dawson, C. W. (2009). *Projects in Computing and Information Systems A Student's Guide (2nd ed.)*. Pearson Education Limited.