

INFORMASI INTERAKTIF

JURNAL INFORMATIKA DAN TEKNOLOGI INFORMASI

PROGRAM STUDI INFORMATIKA – FAKULTAS TEKNIK -UNIVERSITAS JANABADRA

METODE KLASIFIKASI DATA MINING ALGORITMA C4.5 DAN PART UNTUK PREDIKSI WAKTU KELULUSAN MAHASISWA DI UNIVERSITAS DARWAN ALI

Selviana Yunita, Nurahman

PENERAPAN SISTEM PAKAR UNTUK IDENTIFIKASI ANAK BERKEBUTUHAN KHUSUS MENGGUNAKAN METODE *RULE BASED SYSTEM*

Yumarlin MZ, Hanang Indrianta

IMPLEMENTASI SMOTE UNTUK MENGATASI *IMBALANCED DATA* PADA SENTIMEN ANALISIS SENTIMEN HOTEL DI NUSA TENGGARA BARAT DENGAN MENGGUNAKAN ALGORITMA SVM

Erry Maricha Oki Nur Haryanto, Adhien Kenya Anima Estetikha, Rahmad Arif Setiawan

IMPLEMENTASI *DASHBOARD* MICROSOFT POWER BI UNTUK VISUALISASI DATA COVID 19 INDONESIA

Jemmy Edwin Bororing

RANCANG BANGUN APLIKASI KASIR USAHA MIKRO KECIL MENENGAH MENGGUNAKAN METODE *BLOCK PROGRAMMING* (STUDI KASUS : PELANGI STORE)

Agit Amrullah, Fata Aulia

PERANCANGAN APLIKASI PENGENALAN LITERASI COVID-19 MENGGUNAKAN *ACTIONSCRIPT 3.0* PADA *MACROMEDIA FLASH*

Jeffry Andhika Putra, Erry Maricha Oki Nur Haryanto

PENGARUH SMOTE DAN *FORWARD SELECTION* DALAM MENANGANI KETIDAKSEIMBANGAN KELAS PADA ALGORITMA KLASIFIKASI

Ika Nur Fajri, Femi Dwi Astuti

MEDIA KOMUNIKASI KESEHATAN UNTUK TUNA RUNGU DAN TUNA WICARA BERBASIS ANDROID

Ryan Ari Setyawan, Rizqi Mirza Fadilla

IMPLEMENTASI *USER EXPERIENCE DESIGN* PADA PERANCANGAN APLIKASI PEMBELAJARAN PRAKTIKUM *ONLINE* BERBASIS *MOBILE*

Eri Haryanto, Agustin Setiyorini

PROTOTYPE PENGENALAN CANDI DI YOGYAKARTA BERBASIS *AUGMENTED REALITY*

Fatsyahrina Fitriastut, Ryan Ari Setyawan, Helio Rofino Correia



DEWAN EDITORIAL

- Penerbit** : Program Studi Informatika Fakultas Teknik Universitas Janabadra
- Ketua Penyunting
(Editor in Chief)** : Fatsyahrina Fitriastuti, S.Si., M.T. (Universitas Janabadra)
- Penyunting (Editor)** : 1. Jemmy Edwin B, S.Kom., M.Eng. (Universitas Janabadra)
2. Ryan Ari Setyawan, S.Kom., M.Eng. (Universitas Janabadra)
3. Yumarlin MZ, S.Kom., M.Pd., M.Kom. (Universitas Janabadra)
- Alamat Redaksi** : Program Studi Informatika Fakultas Teknik
Universitas Janabadra
Jl. Tentara Rakyat Mataram No. 55-57
Yogyakarta 55231
Telp./Fax : (0274) 543676
E-mail: informasi.interaktif@janabadra.ac.id
Website : <http://e-journal.janabadra.ac.id/>
- Frekuensi Terbit** : 3 kali setahun

JURNAL INFORMASI INTERAKTIF merupakan media komunikasi hasil penelitian, studi kasus, dan ulasan ilmiah bagi ilmuwan dan praktisi dibidang Informatika. Diterbitkan oleh Program Studi Informatika Fakultas Teknik Universitas Janabadra di Yogyakarta, tiga kali setahun pada bulan Januari, Mei dan September.

DAFTAR ISI

	<i>halaman</i>
Metode Klasifikasi Data Mining Algoritma C4.5 Dan Part Untuk Prediksi Waktu Kelulusan Mahasiswa Di Universitas Darwan Ali Selviana Yunita, Nurahman	1 - 7
Penerapan Sistem Pakar Untuk Identifikasi Anak Berkebutuhan Khusus Menggunakan Metode <i>Rule Based System</i> Yumarlin MZ, Hanang Indrianta	8 - 15
Implementasi SMOTE Untuk Mengatasi <i>Imbalanced Data</i> Pada Sentimen Analisis Sentimen Hotel Di Nusa Tenggara Barat Dengan Menggunakan Algoritma SVM Erry Maricha Oki Nur Haryanto, Adhien Kenya Anima Estetikha, Rahmad Arif Setiawan	16 - 20
Implementasi <i>Dashboard</i> Microsoft Power BI Untuk Visualisasi Data Covid 19 Indonesia Jemmy Edwin Bororing	21 - 29
Rancang Bangun Aplikasi Kasir Usaha Mikro Kecil Menengah Menggunakan Metode <i>Block Programming</i> (Studi Kasus : Pelangi Store) Agit Amrullah, Fata Aulia	30 - 37
Perancangan Aplikasi Pengenalan Literasi Covid-19 Menggunakan <i>Actionscript</i> 3.0 Pada <i>Macromedia Flash</i> Jeffry Andhika Putra, Erry Maricha Oki Nur Haryanto	38 - 44
Pengaruh SMOTE Dan <i>Forward Selection</i> Dalam Menangani Ketidakseimbangan Kelas Pada Algoritma Klasifikasi Ika Nur Fajri, Femi Dwi Astuti	45 - 49
Media Komunikasi Kesehatan Untuk Tuna Rungu Dan Tuna Wicara Berbasis Android Ryan Ari Setyawan, Rizqi Mirza Fadilla	50 - 59
Implementasi <i>User Experience Design</i> Pada Perancangan Aplikasi Pembelajaran Praktikum <i>Online</i> Berbasis <i>Mobile</i> Eri Haryanto, Agustin Setiyorini	60 - 69
Prototype Pengenalan Candi Di Yogyakarta Berbasis <i>Augmented Reality</i> Fatsyahrina Fitriastut, Ryan Ari Setyawan, Helio Rofino Correia	70 - 78

PENGANTAR REDAKSI

Puji syukur kami panjatkan kehadiran Allah Tuhan Yang Maha Kuasa atas terbitnya JURNAL INFORMASI INTERAKTIF Volume 7, Nomor 1, Edisi Januari 2022. Pada edisi kali ini memuat 10 (sepuluh) tulisan hasil penelitian dalam bidang informatika.

Harapan kami semoga naskah yang tersaji dalam JURNAL INFORMASI INTERAKTIF edisi Januari tahun 2022 dapat menambah pengetahuan dan wawasan di bidangnya masing-masing dan bagi penulis, jurnal ini diharapkan menjadi salah satu wadah untuk berbagi hasil-hasil penelitian yang telah dilakukan kepada seluruh akademisi maupun masyarakat pada umumnya.

Redaksi

PENGARUH SMOTE DAN FORWARD SELECTION DALAM MENANGANI KETIDAKSEIMBANGAN KELAS PADA ALGORITMA KLASIFIKASI

Ika Nur Fajri¹, Femi Dwi Astuti²

¹Universitas Amikom Yogyakarta, Jl.Padjajaran, Ring Road Utara, Condongcatur, Depok, Sleman, DIY

²Universitas Teknologi Digital Indonesia, Jl. Raya Janti (Majapahit) No.143 Bantul, DIY

Email : ¹fajri@amikom.ac.id , ²femi@utdi.ac.id

ABSTRACT

A high accuracy value in the classification process can ideally be obtained if the number of classes in the dataset is balanced. In fact, the data obtained do not all have a balanced number of classes, thus reducing the performance of the classification algorithm. In addition to the problem of an unbalanced number of classes, the attributes involved in the calculation also affect the accuracy value, so it is necessary to choose which attribute is the most influential. In this study, one method of feature selection is used, namely Forward Selection. This method is used to select which features are the most influential. SMOTE, which is one of the over-sampling algorithms, makes data with fewer classes equal to those with many classes. The results show that in the car evolution dataset the use of SMOTE can increase accuracy by 6.12% and the use of SMOTE with forward selection can increase accuracy by 6.09%. In the glass identification dataset the use of SMOTE can increase accuracy by 9.65% and the use of SMOTE with forward selection can increase accuracy by 12.6%. The use of forward selection with SMOTE is more effective for datasets that have a small number of classes.

Keywords: *Forward Selection, K-NN, Klasifikasi, SMOTE*

1. PENDAHULUAN

Bebagai macam kondisi data dengan karakteristik yang berbeda-beda, sehingga memerlukan penanganan yang berbeda pula, data yang ideal dengan jumlah kelas seimbang, cenderung menghasilkan nilai akurasi yang tinggi, berbeda ketika data yang diolah memiliki kelas yang tidak seimbang. Atribut/fitur yang terlibat dalam perhitungan juga mempengaruhi nilai akurasi, sehingga perlu ditentukan atribut mana yang paling berpengaruh, sehingga menghasilkan nilai akurasi yang tinggi. Untuk membuat kelas menjadi seimbang ada dua metode yang dapat digunakan, *under-sampling* dan *over-sampling*. *Under-sampling* berarti menurunkan jumlah data pada kelas yang lebih banyak sehingga sama jumlahnya dengan kelas yang lebih sedikit, sedangkan *over-sampling* adalah menjadikan data yang jumlah kelasnya lebih sedikit menjadi sama dengan kelas yang memiliki banyak. Pada penelitian ini menggunakan salah satu metode seleksi fitur/atribut *Forward Selection* untuk memilih fitur mana yang paling berpengaruh, dan salah satu algoritma *over-sampling* yaitu SMOTE untuk menanggapi ketidakseimbangan kelas. *Forward selection* mampu memilih fitur yang berpengaruh sehingga fitur-fitur yang tidak terlalu berpengaruh tidak

ikut masuk dalam proses klasifikasi, hal ini diharapkan dapat menaikkan nilai akurasi [1]. SMOTE digunakan karena metode ini terbukti mampu menyeimbangkan jumlah kelas sebelum proses klasifikasi [2].

2. ISI PENELITIAN

Data yang digunakan pada penelitian ini yaitu dataset *car evolution* dan *glass identification* yang diperoleh dari UCI machine learning. *Car evolution* memiliki 1728 record dengan 6 atribut yaitu (*buying price, price of maintance, doors, person capacity, lug boot, safety*), sedangkan dataset *glass identification* memiliki 214 record dengan 10 atribut yaitu (*Id number, refractive index, sodium, magnesium, aluminium, silicon, potassium, calcium, barium, iron*). Pembagian kelas pada dataset *car evolution* dan *glass identification* dapat dilihat pada Tabel 1.

Tabel 1. Pembagian Kelas Dataset

Dataset	Pembagian kelas
Car Evolution	Unacc : 1210 Acc : 384 Good : 69 V-good : 65
Glass	Negative : 138 Positive : 76

Berdasarkan Tabel 1 dapat dilihat jika jumlah pembagian kelas untuk masing-masing dataset tidak seimbang. Ada yang jumlahnya sangat banyak dan ada yang sangat sedikit sehingga dataset ini dapat digunakan untuk pengujian teknik SMOTE.

2.1 Tinjauan Pustaka

Forward selection pernah digunakan untuk menyeleksi fitur pada proses klasifikasi SVM dan Naïve Bayes untuk klasifikasi jalur minat SMA. Pengujian dari SMA ABC *feature* yang terpilih yaitu nilai IPA, tes akademik, abstrak konseptual, logika numerik, dan analisa sintesa. Sedangkan SMA XYZ yaitu nilai IPA, logika numerik dan analisa sintesa [1]. SMOTE digunakan untuk menyeimbangkan kelas pada proses klasifikasi *data mining*. Penggunaan metode SMOTE ini dapat berdampak performa hasil klasifikasi, metode ini mampu meningkatkan performa model klasifikasi [2]. Penggunaan SMOTE untuk klasifikasi dengan metode KNN juga pernah dilakukan untuk mengklasifikasi *car evolution*. Hasilnya menunjukkan bahwa penggunaan SMOTE mampu menaikkan nilai akurasi sebesar 9,97% [3]. Seleksi Fitur *forward Selection* dengan K-NN pernah digunakan untuk prediksi mahasiswa non aktif di Universitas Dian Nuswantoro Semarang. Hasilnya mampu meningkatkan nilai akurasi pada proses klasifikasi. Penggunaan seleksi fitur ini mampu menaikkan akurasi sebesar 8,4% [4]. *Forward selection* juga pernah digunakan untuk prediksi harga komoditi lada pada tahun 2015 dimana hasil penelitian juga menunjukkan bahwa K-NN yang berbasis *forward selection* mampu memberikan kinerja terbaik jika dibandingkan dengan K-NN berbasis *backward elimination* serta SVM yang berbasis seleksi atribut [5].

2.2 SMOTE

Terdapat dua macam metode untuk menangani ketidakseimbangan data yaitu *over sampling* dan *under sampling*. *Syntetic Minority Oversampling Technique* atau biasa disebut dengan SMOTE biasa digunakan untuk menambah data sintesis pada dataset kelas minoritas, sehingga jumlah instances pada kelas minoritas menjadi sama dengan jumlah instances pada kelas mayoritas. Hal ini dilakukan agar data pada kedua kelas seimbang. Data sintesis yang akan ditambahkan hanya akan ditambahkan pada data latih.

2.3 Forward Selection

Metode *Forward Selection* merupakan salah satu metode yang biasa digunakan untuk melakukan seleksi fitur dari dataset yang memiliki banyak atribut. Seleksi fitur sendiri merupakan proses yang biasa digunakan di *machine learning* yang mana sekumpulan dari fitur/atribut yang dimiliki oleh data digunakan untuk proses pembelajaran algoritma. *Forward selection* ini merupakan pemodelan yang dimulai dari nol peubah (*empty model*), lalu satu persatu peubah dimasukan sampai kriteria tertentu dipenuhi [6].

2.4 Naïve Bayes

Naïve Bayes merupakan salah satu algoritma untuk melakukan klasifikasi. Untuk menggunakan teorema Bayes dapat digunakan persamaan (1).

$$P(X|H) = \frac{P(H|X)P(H)}{P(X)} \tag{1}$$

Dimana :

X : merupakan Data dengan class yang belum diketahui

H : Hipotesis data x yang merupakan suatu class spesifik

P(H|X) : merupakan Probabilitas hipotesis H berdasarkan kondisi X (*posteriori probability*)

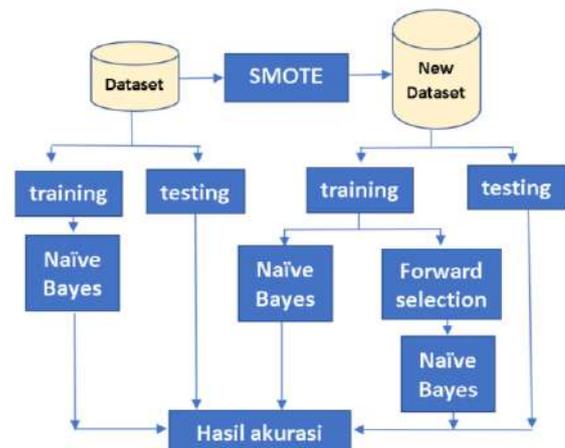
P(H) : merupakan Probabilitas hipotesis H (*prior probability*)

P(X|H) : merupakan Probabilitas X berdasar kondisi pada hipotesis H

P(X) : Probabilitas dari X

2.5 Metode penelitian

Alur penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1 dapat dilihat bahwa proses penelitian dilakukan dengan memilih dataset yang akan digunakan yang dalam hal ini adalah dataset *car evolution* dan *glass identification*. Penelitian diawali dengan melakukan klasifikasi dengan menggunakan algoritma naïve bayes sampai ditemukan nilai akurasi. Setelah itu dicoba menggunakan SMOTE untuk mengatasi ketidakseimbangan kelas pada dataset yang digunakan. Penggunaan SMOTE ini dilakukan selama *preprocessing* sehingga menghasilkan dataset baru (*new dataset*). *New dataset* yang telah terbentuk akan dilakukan klasifikasi dengan algoritma yang sama yaitu naïve bayes hingga ditemukan nilai akurasi. Selanjutnya *new dataset* digunakan untuk klasifikasi dengan memanfaatkan salah satu teknik seleksi fitur (*forward selection*) sebelum proses klasifikasi untuk mendapatkan atribut/fitur yang penting kemudian dicari nilai akurasi.

2.6 Hasil dan Pembahasan

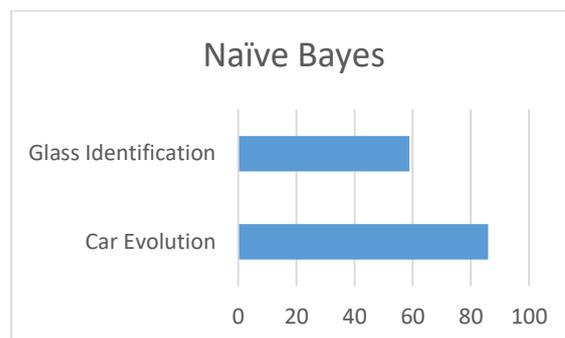
Pengujian dilakukan untuk tiga hal, pengujian pertama dilakukan untuk menghitung nilai akurasi dari klasifikasi dataset menggunakan Naïve Bayes. Pengujian kedua dilakukan untuk menghitung akurasi dari penggunaan teknik SMOTE untuk klasifikasi Naïve Bayes dan pengujian ketiga dilakukan untuk melihat nilai akurasi dari *new dataset* yang diperoleh dari teknik SMOTE lalu sebelum dilakukan klasifikasi dilakukan seleksi fitur dengan menggunakan *forward selection*. Hasil pengujian ini akan dibandingkan untuk melihat akurasi terbaik.

Hasil penelitian untuk dataset yang dilakukan klasifikasi menggunakan Naïve Bayes dapat dilihat pada Tabel 2. Berdasarkan Tabel 2. Dapat dilihat bahwa nilai akurasi yang dihasilkan dari proses klasifikasi untuk dataset *car evolution* sebesar 85,94% sedangkan untuk dataset *glass evolution* menghasilkan nilai yang jauh lebih rendah yaitu 58,83%.

Tabel 2. Klasifikasi dengan Naïve Bayes

Dataset	Akurasi
Car Evolution	85,94%
Glass Identification	58,83%

Grafik perbandingan hasil akurasi dari dua dataset yang digunakan dapat dilihat pada Gambar 2.



Gambar 2. Grafik hasil naïve bayes

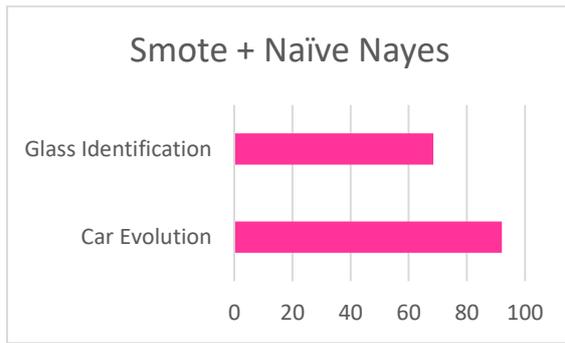
Berdasarkan Gambar 2 jika dilihat nilai akurasi antara kedua dataset sangat jauh berbeda. Metode naïve bayes jika diterapkan pada dataset yang memiliki jumlah kelas lebih banyak maka dapat menghasilkan akurasi yang tinggi, sedangkan untuk dataset yang jumlah kelasnya sedikit, maka nilai akurasi menjadi rendah.

Pengujian selanjutnya, melakukan klasifikasi dengan *new dataset* yang diperoleh dari teknik SMOTE untuk mengatasi ketidakseimbangan kelas. Klasifikasi yang digunakan masih menggunakan metode Naïve Bayes. Hasil pengujian kedua dapat dilihat pada Tabel 3. Berdasarkan Tabel 3. Dapat dilihat bahwa nilai akurasi yang dihasilkan dari proses klasifikasi dengan *new dataset* yang sudah dilakukan *preprocessing* mengalami peningkatan. Akurasi *car evolution* yang semula hanya bernilai 85,94% naik menjadi 92,06% sehingga dapat dikatakan penggunaan SMOTE untuk mengatasi ketidakseimbangan kelas dengan klasifikasi naïve bayes mampu menaikkan 7%. Nilai Akurasi *glass identification* yang semula hanya bernilai 58,83% naik menjadi 68,43% sehingga dapat dikatakan penggunaan SMOTE untuk mengatasi ketidakseimbangan kelas dengan jumlah kelas hanya 2 dengan klasifikasi naïve bayes juga mampu menaikkan akurasi sebesar 9,65%.

Tabel 3. Klasifikasi dengan Naïve Bayes dan SMOTE

Dataset	Akurasi
Car Evolution	92,06%
Glass Identification	68,48%

Grafik perbandingan hasil akurasi penggunaan teknik SMOTE untuk mengatasi ketidakseimbangan kelas pada dua dataset dapat dilihat pada Gambar 3.



Gambar 3. Grafik hasil naïve bayes dengan SMOTE

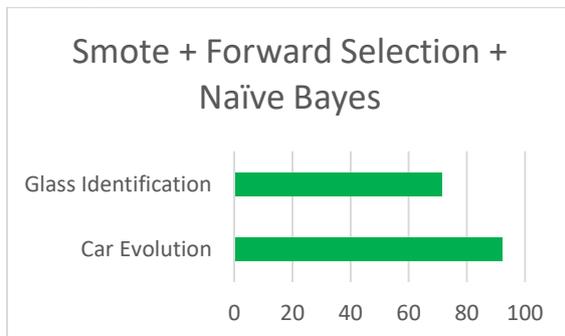
Pengujian ketiga dilakukan dengan menggunakan SMOTE beserta forward selection untuk menentukan atribut yang akan digunakan dalam proses klasifikasi. Hasil pengujian dapat dilihat pada Tabel 4.

Tabel 4. Klasifikasi dengan Naïve Bayes, SMOTE dan Forward Selection

Dataset	Akurasi	Jumlah atribut yang digunakan
Car Evolution	92,03%	5
Glass Identification	71,43%	7

Berdasarkan Tabel 4 dapat dilihat bahwa penggunaan seleksi fitur *forward selection* pada dataset *car evolution* dengan memilih antribut menjadi 5 justru dapat menurunkan nilai akurasi sebesar 0,03%. Sedangkan untuk dataset *glass identification* yang memiliki jumlah kelas sedikit (2 kelas) dapat menaikkan nilai akurasi hasil klasifikasi dari 68,48% menjadi 71,43% sehingga mengalami kenaikan sebesar 2,95%.

Grafik perbandingan nilai akurasi hasil klasifikasi dari new dataset yang sudah menggunakan SMOTE serta sudah menerapkan seleksi fitur *forward selection* dapat dilihat pada Gambar 4.



Gambar 4. Grafik hasil naïve bayes dengan SMOTE dan *forward selection*

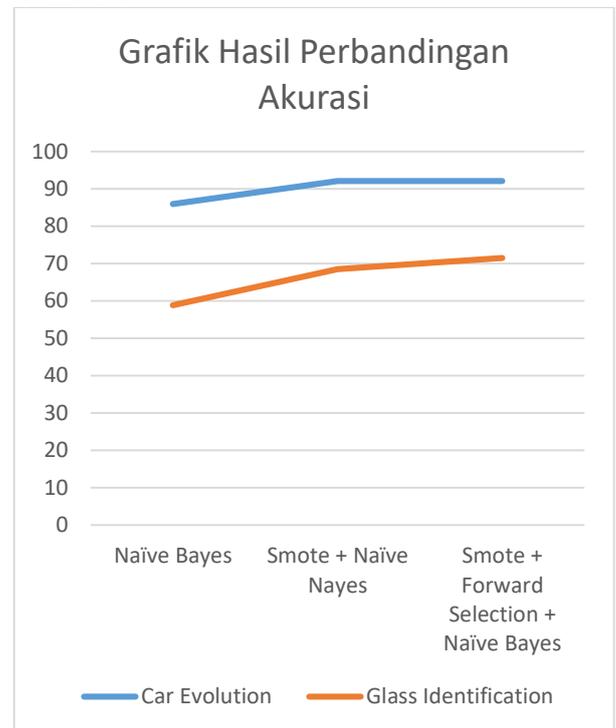
Setelah melalui beberapa pengujian, perbandingan akurasi dari kedua dataset dengan

metode yang berbeda-beda dapat dilihat pada Tabel 4.

Tabel 4. Tabel Perbandingan Akurasi

Metode	Akurasi	
	Car evolution	Glass Identification
Naïve Bayes	85,94%	58,83%
SMOTE dengan Naïve Bayes	92,06%	68,48%
SMOTE, Forward Selection dengan Naïve Bayes	92,03%	71,43%

Grafik perbandingan hasil akurasi dari ketiga pengujian yang dilakukan dapat dilihat pada Gambar 5.



Gambar 5. Perbandingan hasil akurasi

Berdasarkan Gambar 5 dapat dilihat bahwa terjadi perbedaan antara hasil nilai akurasi hasil klasifikasi dari dua dataset yang digunakan. Pada dataset *car evolution* nilai akurasi dari penggunaan Naïve Bayes dibandingkan dengan Naïve Bayes menggunakan Smote mengalami kenaikan sedangkan dari naïve Bayes dengan SMOTE dibandingkan dengan Naïve Bayes dengan SMOTE dan *Forward Selection* mengalami penurunan meskipun sangat sedikit. Sedangkan pada dataset *Glass Identification*, nilai akurasi selalu naik baik nilai akurasi naïve bayes, naïve bayes dengan SMOTE maupun naïve bayes dengan SMOTE dan *Forward Selection*.

3. KESIMPULAN

Berdasarkan hasil penelitian yang sudah dilakukan, maka dapat disimpulkan bahwa pada dataset *car evolution* penggunaan SMOTE dapat meningkatkan akurasi sebesar 6,12% dan penggunaan SMOTE dengan forward selection dapat meningkatkan akurasi sebesar 6,09%. Pada dataset *glass identification* penggunaan SMOTE dapat meningkatkan akurasi sebesar 9,65% dan penggunaan SMOTE dengan forward selection dapat meningkatkan akurasi sebesar 12,6%. Sehingga dapat dikatakan penggunaan forward selection dengan SMOTE ini mampu menaikkan akurasi hasil klasifikasi dan lebih efektif digunakan pada dataset yang memiliki jumlah pembagian kelas yang sedikit.

DAFTAR PUSTAKA

- [1] T. B. Sasongko and O. Arifin, "Implementasi Metode Forward Selection pada Algoritma Support Vector Machine (SVM) dan Naive Bayes Classifier Kernel Density (Studi Kasus Klasifikasi Jalur Minat SMA)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 4, pp. 383–388, 2019, doi: 10.25126/jtiik.201961000.
- [2] A. A. Arifiyanti and E. D. Wahyuni, "Smote: Metode Penyeimbang Kelas Pada Klasifikasi Data Mining," *SCAN - J. Teknol. Inf. dan Komun.*, vol. 15, no. 1, pp. 34–39, 2020, doi: 10.33005/scan.v15i1.1850.
- [3] F. Dwi Astuti, Femi and Nova Lenti, "Implementasi SMOTE untuk mengatasi," *JUPITER (Jurnal Penelit. Ilmu dan Teknol. Komputer)*, vol. 13, pp. 89–98, 2021.
- [4] A. Salam, F. B. Nugroho, and J. Zeniarja, "Implementasi Algoritma K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Mahasiswa Non Aktif Universitas Dian Nuswantoro Semarang," *JOINS (Journal Inf. Syst.*, vol. 5, no. 1, pp. 69–76, 2020, doi: 10.33633/joins.v5i1.3351.
- [5] M. Nanja and P. Purwanto, "Metode K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Harga Komoditi Lada," *Pseudocode*, vol. 2, no. 1, pp. 53–64, 2015, doi: 10.33369/pseudocode.2.1.53-64.
- [6] M. F. Nugroho and S. Wibowo, "Fitur Seleksi Forward Selection Untuk Menentukan Atribut Yang Berpengaruh Pada Klasifikasi Kelulusan Mahasiswa Fakultas Ilmu Komputer UNAKI Semarang Menggunakan Algoritma Naive Bayes," *J. Inform. Upgris*, vol. 3, no. 1, pp. 63–70, 2017, doi: 10.26877/jiu.v3i1.1669.