

ANALISIS RANDOM FOREST PADA KLASIFIKASI CART KETIDAKTEPATAN WAKTU KELULUSAN MAHASISWA UNIVERSITAS TERBUKA

Analysis of Random Forest In Inaccuracies CART Classification of Terbuka University Student Graduates

Gede Suwardika¹, I Ketut Putu Suniantara^{2*}

¹Prodi Statistika – Universitas Terbuka – UPBJJ Denpasar
Jln. Gurita Sesetan Denpasar Selatan – Bali, 80223, Indonesia

²Prodi Sistem Informasi – STIKOM Bali
Jl. Raya Puputan No. 86 Renon, Denpasar Timur – Bali, 80234, Indonesia

e-mail: ¹isuwardika@ecampus.ut.ac.id; ^{2*} suniantara@stikom-bali.ac.id.

Corresponding author*

Abstrak

Classification and Regression Tree (CART) merupakan salah satu metode klasifikasi yang populer digunakan di berbagai bidang. Metode tersebut dinilai mampu menghadapi berbagai kondisi data. Akan tetapi metode *CART* memiliki kelemahan pada prediksi pohon klasifikasi yaitu kurang stabil pada perubahan *data learning* yang akan menyebabkan perubahan besar pada hasil prediksi pohon klasifikasi. Memperbaiki prediksi dari pohon klasifikasi *CART*, dikembangkan metode *ensemble random forest* yang mengkombinasikan banyak pohon klasifikasi untuk meningkatkan kestabilan dan menentukan prediksi klasifikasi. Penelitian ini bertujuan untuk meningkatkan kestabilan dan akurasi prediktif *CART* dengan *Random Forest*. Kasus yang digunakan dalam penelitian ini adalah klasifikasi ketidaktepatan waktu kelulusan mahasiswa Universitas Terbuka. Hasil analisis menunjukkan bahwa *random forest* mampu meningkatkan akurasi klasifikasi ketidaktepatan waktu kelulusan mahasiswa yang mencapai konvergen dengan prediksi klasifikasi mencapai 93.23%.

Kata Kunci : *Klasifikasi, CART, random forest.*

Abstract

Classification and Regression Tree (CART) is one of the classification methods that are popularly used in various fields. The method is considered capable of dealing with various data conditions. However, the *CART* method has weaknesses in the classification tree prediction, which is less stable in changes in learning data which will cause major changes in the results of the classification tree prediction. Improving the predictions of the *CART* classification tree, an ensemble random forest method was developed that combines many classification trees to improve stability and determine classification predictions. This study aims to improve *CART* predictive stability and accuracy with *Random Forest*. The case used in this study is the classification of inaccuracies in Open University student graduation. The results of the analysis show that random forest is able to increase the accuracy of the classification of the inaccuracy of student graduation that reaches convergence with the prediction of classification reaching 93.23%.

Keywords: *Classification, CART, random forest*

Submitted: 10th April 2019

Revised: 30th May 2019

Accepted: 15th July 2019

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



1. PENDAHULUAN

Classification and Regression Trees (CART) merupakan salah satu metode klasifikasi nonparametrik yang populer digunakan. Metode ini telah digunakan dalam penyelesaian masalah di berbagai bidang seperti bidang kesehatan, *marketing*, sosial, *financial*, dan lain sebagainya. Metode ini mampu menghadapi berbagai kondisi data. CART memiliki banyak kelebihan diantaranya mampu mengeksplorasi data berdimensi tinggi dengan komputasi yang efisien, dapat digunakan pada kombinasi data kontinu maupun kategorik, serta interpretasi yang mudah dilakukan. Diantara banyak kelebihan *CART*, kelemahan metode pohon klasifikasi ini yaitu kurang stabil pada perubahan *data learning* yang akan menyebabkan perubahan besar pada hasil prediksi pohon klasifikasi [12]. Untuk mengatasi kelemahan metode *CART* diperlukan suatu metode yang dapat digunakan untuk meningkatkan akurasi prediksi dari *classifier* yang kurang stabil .

Menurut [6] mengembangkan suatu metode untuk meningkatkan akurasi prediksi dari *classifier* yang tidak stabil, yaitu metode *Ensemble*. Metode *Ensemble* merupakan metode kombinasi banyak *classifier* tunggal dimana hasil prediksi masing-masing *classifier* digabungkan menjadi prediksi akhir melalui proses *voting* mayoritas untuk klasifikasi atau *voting* rata-rata untuk kasus regresi [8]. Penelitian sebelumnya menunjukkan bahwa metode *Ensemble* seringkali menghasilkan prediksi yang lebih akurat dibandingkan dengan *classifier* tunggal [13]. Salah satu Metode *Ensemble* yang terbaru ialah *Random Forest* yang dikembangkan dari proses *Bagging* [2][12]. *Random Forest* pertama kali dikenalkan oleh Breiman pada Tahun 2003. *Random Forest* memiliki kelebihan dalam proses iterasi komputasi yang lebih cepat [10] [5].

Berbagai penelitian mengenai klasifikasi kelulusan yang telah dilakukan mengindikasikan bahwa banyak sekali faktor yang mempengaruhi ketepatan klasifikasi kelulusan mahasiswa. Menurut [9] faktor-faktor tersebut dapat bersumber dari diri mahasiswa (faktor internal) atau dari luar diri mahasiswa (lingkungan/faktor eksternal). Menurut [7] dalam Nurganita pada penelitiannya yang berjudul “identifikasi faktor-faktor yang ketepatan waktu kelulusan mahasiswa program pasca sarjana IPB” dengan menggunakan analisis regresi logistik biner bahwa faktor yang mempengaruhi ketepatan kelulusan adalah fakultas, jenis kelamin, jalur masuk, IPK, dan beasiswa dengan akurasi klasifikasi 85%.

Penelitian tentang klasifikasi ketepatan waktu lulusan lainnya dilakukan oleh [3] menyatakan bahwa faktor yang mempengaruhi lama studi mahasiswa adalah lama penyusunan skripsi, IPK dan jurusan mahasiswa dengan ketepatan klasifikasi 87%, akan tetapi hasil klasifikasi tidak stabil disaat ada penambahan variabel dan jumlah data. Menurut Margasari [1] mendapatkan hasil akurasi klasifikasi *CART* 94,2% dan hasil akurasi klasifikasi regresi logistik biner sebesar 86,7% pada data profil mahasiswa FMIPA Universitas Brawijaya. Pada penelitian lainnya yang dilakukan oleh [4] menyatakan bahwa *bagging CART* mampu meningkatkan akurasi klasifikasi sebesar 4,38% dari klasifikasi *CART* tanpa *bagging*.

Berdasarkan pendahuluan tersebut di atas peningkatan klasifikasi *CART* dapat ditingkatkan dengan menggunakan *Ensemble* salah satunya dengan *Random Forest*. Untuk itu pada penelitian ini akan diterapkan metode *Random Forest* pada *CART* dengan tujuan meningkatkan kestabilan dan akurasi prediktif pohon yang dihasilkan. Hasil klasifikasi *Random Forest CART* yang diperoleh akan dibandingkan dengan Metode *CART* sebelumnya.

2. METODE PENELITIAN

2.1 Data dan Variabel Penelitian

Penelitian ini dilakukan di Universitas Terbuka UPPJJ Denpasar dengan waktu penelitian sepuluh bulan. Data penelitian yang digunakan dalam penelitian ini adalah data sekunder berupa data wisuda tahun 2018 dengan jumlah data yang terverifikasi sebanyak 133 wisudawan. Adapun variabel – variabel yang digunakan yaitu 1). Variabel respon (Y) yang menyatakan status kelulusan mahasiswa tepat waktu dan tidak tepat waktu, 2). Variabel bebasnya yaitu terdiri dari: jenis kelamin (X_1), jurusan/prodi (X_2), IPK (X_3), Pendidikan orang tua (X_4), pekerjaan orang tua (bapak dan ibu) (X_5), penghasilan orang tua (X_6) dan status mahasiswa (X_7).

2.2 Analisis Data

Langkah–langkah untuk mencapai tujuan penelitian tersebut dituangkan pada analisis data, dimana proses awal terlebih dahulu dilakukan *pre-processing*. Data *missing* berskala nominal dan ordinal akan diisi

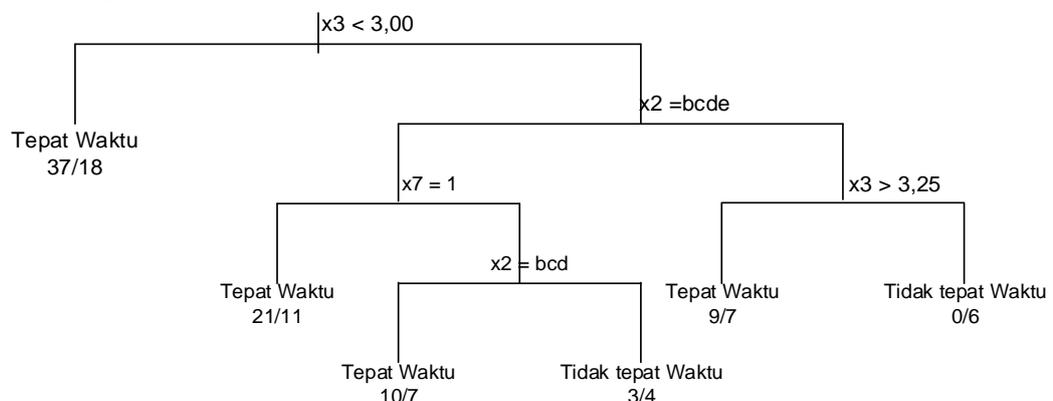
dengan nilai modus sedangkan data *missing* berskala interval akan diisi dengan nilai *mean*. Tahapan analisis data sebagai berikut:

- a. Analisis *CART*, dengan tahapan sebagai berikut:
 1. Menentukan semua kemungkinan penyekatan pada tiap variabel penjelas.
 2. Menghitung kehomogenan simpul pada semua variabel penjelas.
 3. Jika simpul induk telah didapat, maka simpul anak kiri dan kanan dibuat dengan cara yang sama untuk semua variabel penjelas berdasarkan data yang sudah dikelompokkan oleh simpul induk.
 4. Pembentukan pohon dilakukan sampai dipenuhi suatu aturan penghentian tertentu.
 5. Pemangkasan pohon dilakukan untuk mendapatkan pohon akhir yang lebih sederhana.
 6. Setelah diperoleh pengelompokan, kemudian melakukan interpretasi.
- b. Analisis klasifikasi *Random Forest* dilakukan sesuai dengan tahapan berikut:
 1. Menentukan jumlah variabel prediktor yang nanti akan diambil secara acak untuk digunakan dalam proses pemilahan pohon klasifikasi yaitu sebanyak 3.
 2. Mengambil n sampel dengan teknik *resampling* dengan pengembalian sehingga diperoleh *dataset* baru D^* .
 3. Membentuk model *CART* tanpa *pruning* dari *dataset* D^* dimana setiap pemilahan *node* dipilih pemilah terbaik dari 3 variabel prediktor yang diambil secara acak.
 4. Mengamati prediksi pengelompokan untuk tiap data sampel.
 5. Mengulangi langkah 1-4 dari 2-20 kali replikasi.
 6. Melakukan *voting* mayoritas terhadap dugaan pengelompokan tiap sampel data dari 2-20 iterasi yang telah dilakukan.
 7. Menghitung 1-APER klasifikasi *Random Forest*, *sensitivity*, *specificity*, *standar deviasi*, 1-APER, dan *G-means*.
- c. Berdasarkan analisis klasifikasi dengan metode *Random Forest* yang telah dilakukan, selanjutnya membandingkan kinerja klasifikasi *CART* dengan *Random Forest*. Perbandingan akurasi klasifikasi dilakukan berdasarkan 1-APER 20 replikasi, *sensitivity*, *specificity*, dan *G-means*. Sedangkan perbandingan Kecepatan akurasi klasifikasi untuk mencapai konvergen dilakukan berdasarkan standar deviasi 1-APER.

3. HASIL DAN PEMBAHASAN

3.1. Klasifikasi CART

Metode klasifikasi *CART* menggunakan algoritma pohon keputusan (*decision tree*). Hasil analisis *CART* berupa pohon klasifikasi pada kasus ketidaktepatan waktu lulusan mahasiswa Universitas Terbuka ditunjukkan pada Gambar 1. Berdasarkan pohon klasifikasi tersebut diketahui bahwa variabel yang digunakan sebagai pemilah pohon klasifikasi *CART* dan paling menentukan waktu kelulusan yaitu IPK (X_3), program studi (X_2), status mahasiswa (X_7).



Gambar 1. Pohon Klasifikasi *CART*

Berdasarkan pohon klasifikasi *CART* didapatkan hasil klasifikasi yang menyatakan klasifikasi *CART* terdapat sebanyak 77 data sampel waktu kelulusan diklasifikasikan tepat waktu sedangkan 3 lainnya salah diklasifikasikan. Pada data sampel kelulusan tidak tepat waktu, terdapat 10 data sampel tepat diklasifikasikan tidak tepat waktu namun data sampel yang salah diklasifikasikan lebih banyak yaitu 43 (Tabel 1).

Tabel 1. Hasil Klasifikasi *CART*

		Kelas Sebenarnya	
		Tepat waktu	tidak tepat waktu
Kelas	tepat waktu	77	43
Prediksi	tidak tepat waktu	3	10

Pada Tabel 1, juga digunakan untuk mengukur akurasi klasifikasi pada pohon klasifikasi *CART* dimana dapat dihitung nilai $1-APER$, *sensitivity*, *specificity*, dan *G-means*. Hasil perhitungan menunjukkan bahwa $1-APER$ pohon klasifikasi *CART* bernilai 0,6541 atau dengan kata lain, data sampel yang tepat diklasifikasikan secara keseluruhan sebanyak 65,41%. *Sensitivity* pohon klasifikasi *CART* bernilai 0,1887 yang artinya hanya 18,87% data waktu kelulusan tepat diklasifikasikan tidak tepat waktu. Adapun *specificity* bernilai 0,9625 yang artinya sebanyak 96,25% waktu lulusan tepat diklasifikasi tepat waktu. Hal tersebut menunjukkan bahwa *CART* hanya bagus dalam mengklasifikasikan tepat waktu kelulusan mahasiswa. Sedangkan akurasi klasifikasi untuk waktu kelulusan tidak tepat waktu sangat rendah. Berikut merupakan perhitungan nilai $1-APER$, *sensitivity*, *specificity*, dan *G-means* hasil klasifikasi dengan *CART*.

$$1 - APER = \frac{TP+TN}{TP+FP+TN+FN} = \frac{10+77}{133} = 0,6541 \quad (1)$$

$$sensitivity = \frac{TP}{TP+FN} = \frac{10}{53} = 0,1887 \quad (2)$$

$$specificity = \frac{TN}{TN+FP} = \frac{77}{80} = 0,9625 \quad (3)$$

$$G - means = \sqrt{sensitivity \times specificity}$$

$$G - means = \sqrt{0,1887 \times 0,9626} = 0,4261 \quad (4)$$

Pengujian signifikansi prediksi pohon klasifikasi *CART* digunakan uji *Press's Q*, dengan jumlah pengamatan total $N = 133$, jumlah pengamatan yang benar diklasifikasikan $n = 80$, dan jumlah kelas $K = 2$, maka perhitungan statistik *Press's Q* bernilai 12,6391. Adapun nilai kritis pada tingkat signifikansi 0,01 ialah 6,63.

$$Press'sQ = \frac{[N-(nK)]^2}{N(K-1)} = \frac{[133-(87.2)]^2}{133(2-1)} = 12,6391 \quad (5)$$

Berdasarkan statistik *Press's Q*, didapatkan nilainya lebih besar dari nilai kritis. Sehingga dapat disimpulkan bahwa prediksi pohon klasifikasi *CART* signifikan. Berdasarkan pohon klasifikasi yang terbentuk, ketidaktepatan waktu kelulusan dibagi menjadi dua kelas yaitu lulus tepat waktu dan tidak tepat waktu.

a. Kelulusan Tepat Waktu

Lulus tepat waktu menjadi 4 simpul dengan karakteristik sebagai berikut:

1. Kelulusan tepat waktu dengan IPK kurang dari 3.
2. Kelulusan tepat waktu dengan IPK lebih dari 3; program studi selain Ilmu Administrasi Negara; dengan status mahasiswa bekerja.
3. Kelulusan tepat waktu dengan IPK lebih dari 3; program studi selain Ilmu Administrasi Negara dan ilmu komunikasi; dengan status mahasiswa bekerja.
4. Kelulusan tepat waktu dengan IPK lebih dari 3,25; program studi selain akuntansi; dengan status mahasiswa tidak bekerja.

b. Kelulusan Tidak Tepat Waktu

Lulus tidak tepat waktu menjadi 2 simpul dengan karakteristik sebagai berikut:

1. Kelulusan tidak tepat waktu dengan IPK lebih dari 3; status mahasiswa bekerja pada program studi manajemen.
2. Kelulusan tidak tepat waktu dengan IPK lebih dari 3,25; program studi selain akuntansi; dengan status mahasiswa tidak bekerja.

3.2. Klasifikasi dengan *Random Forest*

Klasifikasi dengan *Random Forest* merupakan modifikasi dari *Bagging CART*. Pemilihan pemilah pada *random forest* tidak melibatkan semua variabel namun sebagian saja yang terambil secara acak dalam setiap pemilihan. Setiap pemilihan dilakukan pengambilan acak sebanyak 2 variabel dari 7 variabel yang kemudian dicari pemilah terbaik dari 2 variabel tersebut.

Hasil klasifikasi *random forest* pohon klasifikasi yang dilakukan pada data ketidaktepatan waktu kelulusan mahasiswa Universitas Terbuka, hampir sama dengan menggunakan *bagging CART* dimana terdapat 3 variabel yang paling sering digunakan sebagai pemilah pohon dan paling menentukan ketepatan kelulusan secara berurutan adalah Indeks Prestasi Kumulatif (X_3), Program Studi (X_2), dan status mahasiswa (X_7). Klasifikasi *random forest CART* dapat dilihat pada Tabel 2. Sebanyak 81 data sampel ketepatan waktu kelulusan diklasifikasikan tepat waktu sedangkan 1 lainnya salah diklasifikasikan sebagai ketidaktepatan waktu kelulusan. Pada data sampel tidak tepat waktu, hanya 43 data sampel tepat diklasifikasikan tidak tepat waktu. Sedangkan 8 data sampel salah diklasifikasi sebagai kelulusan tepat waktu.

Tabel 2. Hasil Klasifikasi *Random Forest CART*

		Kelas Sebenarnya	
		Tepat waktu	Tidak tepat waktu
Kelas Prediksi	Tepat waktu	81	8
	Tidak tepat waktu	1	43

Hasil perhitungan menunjukkan bahwa $1-APER$ pohon klasifikasi *random forest CART* bernilai 0,9323 atau dengan kata lain, data sampel yang tepat diklasifikasikan secara keseluruhan sebanyak 93,23%. *Sensitivity* pohon klasifikasi *random forest CART* bernilai 0,8431 yang artinya hanya 84,31% data waktu kelulusan tepat diklasifikasikan tidak tepat waktu. Adapun *specificity* bernilai 0,9878 yang artinya sebanyak 98,78% waktu lulusan diklasifikasi tepat waktu. Nilai ketiganya tidak jauh berbeda atau relatif berimbang yang artinya *random forest CART* bagus dalam mengklasifikasikan baik untuk kelas lulus tepat waktu maupun lulus tidak tepat waktu. Perhitungan nilai $1-APER$, *sensitivity*, *specificity*, dan *G-means* hasil klasifikasi dengan *random forest* sebagai berikut:

$$1 - APER = \frac{TP+TN}{TP+FP+TN+FN} = \frac{43+81}{133} = 0,9323 \quad (6)$$

$$sensitivity = \frac{TP}{TP+FN} = \frac{43}{33+8} = 0,8431 \quad (7)$$

$$specificity = \frac{TN}{TN+FP} = \frac{81}{81+1} = 0,9878 \quad (8)$$

$$G - means = \sqrt{sensitivity \times specificity}$$

$$G - means = \sqrt{0,8431 \times 0,9878} = 0,9126 \quad (9)$$

Mengukur kecepatan akurasi prediksi untuk mencapai konvergen pada metode *Random Forest* dilihat dengan menggunakan standar deviasi $1-APER$ *random forest CART* yang diperoleh sebesar 0,0192. Pengujian signifikansi prediksi klasifikasi *random forest CART* digunakan uji *Press's Q*. Diketahui bahwa jumlah pengamatan total $N = 133$, jumlah pengamatan yang benar diklasifikasikan $n = 124$, dan jumlah kelas $K = 2$, maka perhitungan statistik *Press's Q* bernilai 99,4361.

$$Press's Q = \frac{[N-(nK)]^2}{N(K-1)} = \frac{[133-(124 \cdot 2)]^2}{133(2-1)} = 99,4361 \quad (10)$$

Adapun nilai kritis pada tingkat signifikansi 0,01 adalah 6,63. Berdasarkan statistik *Press's Q*, didapatkan nilai yang lebih besar dari nilai kritis. Karena statistik *Press's Q* melebihi nilai kritis, maka dapat disimpulkan bahwa prediksi *Random Forest* signifikan secara statistik.

3.3. Perbandingan Klasifikasi

Kinerja metode *Random Forest* diukur dari akurasi prediksi dan kecepatan akurasi prediksi untuk mencapai konvergen. Akurasi prediksi secara umum diukur dari nilai 1-*APER* dimana semakin besar 1-*APER* maka semakin baik akurasi prediksinya. Semakin besar *sensitivity* maka semakin baik akurasi prediksi klasifikasi tidak tepat waktu sedangkan semakin besar *specificity* maka semakin baik akurasi prediksi klasifikasi tepat waktu. Kemampuan dalam prediksi data diukur dari keseimbangan *sensitivity* dan *specificity* yang dirangkum dalam nilai *G-means* dimana semakin besar nilainya maka semakin baik. Sedangkan kecepatan akurasi prediksi untuk mencapai konvergen diukur dari standar deviasi 1-*APER* dimana semakin kecil semakin cepat atau semakin sedikit pohon yang dibutuhkan untuk mencapai akurasi yang konvergen.

Tabel 3. Ukuran Kinerja Metode Klasifikasi

Metode Klasifikasi	1 - APER	Sensitivity	Specificity	G-Means
CART	0,6541*	0,1887	0,9625	0,4261
Random Forest	0,9323* (0,031)	0,8431	0,9878	0,9126

Keterangan: *Statistik *Press's Q* signifikan
() standar deviasi

Berdasarkan Tabel 3, dapat diketahui bahwa 1-*APER* pohon klasifikasi *CART* bernilai 0,6541 yang artinya sebanyak 65,41% data sampel tepat diklasifikasikan. Sedangkan *random forest* menghasilkan 1-*APER* bernilai 93,23%. Melihat nilai 1-*APER* dari kedua metode dapat dikatakan metode *random forest* mampu meningkatkan akurasi klasifikasi *CART* secara umum.

Keseimbangan akurasi prediksi untuk tiap kelas dapat diukur dari nilai *G-means*. Aspek ini perlu diperhitungkan karena metode klasifikasi cenderung baik dalam memprediksi kelas dengan data sampel yang lebih banyak namun buruk dalam memprediksi kelas dengan data sampel sedikit. Berdasarkan Tabel 3 nampak bahwa pohon klasifikasi *CART* memiliki *G-means* bernilai 0,4261. Sedangkan *random forest* menghasilkan pohon klasifikasi yang memiliki nilai *G-means* (0,9126) lebih besar dibanding *CART* yang artinya akurasi prediksi metode *random forest* cenderung lebih seimbang dalam melakukan prediksi di tiap kelas.

4. KESIMPULAN

Berdasarkan hasil pembahasan dapat disimpulkan sebagai berikut:

- Penerapan klasifikasi *CART* pada klasifikasi ketepatan waktu lulusan mahasiswa Universitas Terbuka menunjukkan bahwa variabel yang digunakan sebagai pemilah pohon klasifikasi *CART* dan paling menentukan waktu ketepatan lulusan adalah IPK, program studi dan status mahasiswa.
- Klasifikasi *random forest* mampu menghasilkan akurasi klasifikasi yang lebih baik dibandingkan dengan pohon klasifikasi *CART*.

UCAPAN TERIMA KASIH

Pada kesempatan ini, penulis mengucapkan terima kasih yang sebesar – besarnya kepada Universitas Terbuka telah memberi dukungan finansial terhadap penelitian ini, sehingga penulis dapat menyelesaikan penelitian ini dengan baik dan tepat waktu.

DAFTAR PUSTAKA

- [1] A. Margasari, Penerapan metode CART (Classification and Regression Trees) dan analisis regresi logistik Biner pada klasifikasi profil mahasiswa FMIPA Universitas Brawijaya, *Studen J*, vol. 2, no. 4, 2014.
- [2] D. R. C. A. J. R. S. A. Cutler, *Ensemble Machine Learning*, no. 2, 2012.

- [3] I. K. P. Suniantara, Penerapan Metode Regresi Berstruktur Pohon Dalam Memprediksi Berat Badan Bayi Lahir, Studi Kasus: Rumah Sakit Umum Daerah Wangaya, Bali: Jurusan Matematika, Universitas Udayana, 2008.
- [4] I. K. P. S. A. M. Rusli, Ketepatan Klasifikasi Bagging CART Pada Klasifikasi Ketidaktepatan Waktu Kelulusan Mahasiswa STIKOM Bali, in *Konferensi Nasional Sistem dan Informatika*, Bali, 2017.
- [5] J. A. B. a. J. R. P. Oskar Gislason, Random Forests for land cover classification, *Pattern Recognit. Lett.*, vol. 27, no. 4, p. 294–300.
- [6] M. Wezel and R. Potharst, Improved Customer Choice Predictions using Ensemble Methods, *Eur. J. Oper. Res.*, vol. 181, no. 1, pp. 436–452, 2007.
- [7] N. Nurganita, Identifikasi Faktor-faktor Yang Mempengaruhi Ketepatan Waktu Kelulusan Mahasiswa Program Sarjana IPB, Bogor: IPB Bogor, 2015.
- [8] N. T. V. K. D. S. D. P. a. A. F. A. Zhukov, Ensemble methods of classification for power systems security assessment, *Appl. Comput. Informatics*, vol. 15, no. 1, p. 45–53, 2019.
- [9] S. Suryabrata, Psikologi Pendidikan, Jakarta: Raja Grafindo Persada, 2008.
- [10] R. S. E. J. Cheung-Wai Chan and D. Paelinckx, Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery, vol. 112, no. 6, p. 299–301.
- [11] R. Timofeev, Classification and Regression Trees (CART) Theory and Applications, Berlin: Humboldt University, 2004.
- [12] V. Y. K. a. P. K. Sinha, Random Forest Classifier: A Survey and Future Research Directions, *Int. J. Adv. Comput.*, vol. 36, no. 1, p. 1144–1156, 2013.
- [13] Z.-H. Zhou, Ensemble Learning, *Berlin: Springer*, p. 270–273, 2009.

