

IDENTIFIKASI SPESIES MANGROVE MENGGUNAKAN ALGORITME RANDOM FOREST

(Identification Mangrove Species use Forest Random Algoritm)

Ishak Ariawan^{1*}, Ayang Armelita Rosalia¹, Wildan Aprizal Arifin¹, Luthfi Anzani¹, La Ode Alam Minsaris¹, Lukman¹

¹Pendidikan Kelautan dan Perikanan, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No.229, Sukasari, Bandung, Jawa Barat 40154

*Corresponding author: ayang.armelita@upi.edu

ABSTRACT

Identification of mangrove species composition is an important topic in management and conservation of coastal ecosystem. The biodiversity of mangrove species affects the sustainability and balance of the entity-related in its ecosystem. This study aims to explore the potential of Machine Learning to identify mangrove species. Specifically, the Random Forest algorithm is used to classify six mangrove species, that are: *Avicennia eucalyptifolia*, *Bruguiera gymnorrhiza*, *Rhizophora apiculata*, *Rhizophora mucronata*, *Unrecorded Sp.*, and *Xylocarpus granatum*. Several approaches were taken to strengthen the performance of the Random Forest algorithm, that are preprocessing (SMOTE) and min-max normalization to balance the data distribution. The result shows that the projection of the normalization range (interval 0-1) has no effect in reducing the data pattern dimensionally. After preprocessing and normalization, five attributes (species, wood density, diameter at breast height, total of above ground biomass, and below-ground root) were classified and analyzed with species as the target attribute. The construction of model parameters is based on the total number of SMOTE results by specifying 100 and 500 as the number of single tree and 1000 as the number of nodes and the default predictor variable. The final result shows that the Random Forest algorithm obtained an optimal evaluation value with an average of 99.97% using the number of single tree and the predetermined cut-offs. The maximum accuracy of 100% is obtained from the number of single tree and cut-off with the following sizes: (1) 500 and 80:20; (2) 500 and 90:10; and (3) 100 and 80:20. These results indicate that it is very effective to apply the Random Forest algorithm as a machine learning classifier method to classify mangrove species.

Keywords: coastal ecosystem, data normalization, machine learning, mangrove, random forest

ABSTRAK

Identifikasi komposisi spesies mangrove adalah topik yang penting dalam manajemen dan konservasi ekosistem pesisir. Biodiversitas spesies mangrove berpengaruh terhadap keberlangsungan dan keseimbangan entitas-entitas yang terkait di dalam ekosistemnya. Penelitian ini bertujuan untuk menggali potensi *Machine Learning* untuk mengidentifikasi spesies mangrove. Secara spesifik, algoritme Random Forest digunakan untuk mengklasifikasikan enam spesies mangrove, yaitu: *Avicennia eucalyptifolia*, *Bruguiera gymnorrhiza*, *Rhizophora apiculata*, *Rhizophora mucronata*, *Unrecorded Sp.*, dan *Xylocarpus granatum*. Beberapa pendekatan dilakukan untuk memperkuat kinerja algoritme Random Forest, yaitu *preprocessing* (SMOTE) dan normalisasi *min-max* untuk menyeimbangkan distribusi data. Hasil tahap ini menunjukkan bahwa proyeksi dari rentang normalisasi (interval

0-1) tidak berpengaruh dalam mereduksi pola data secara dimensional. Setelah *preprocessing* dan normalisasi data, lima atribut (*spesies*, *wood density*, *diameter at breast height*, *total of above ground biomass*, dan *below-ground root*) diklasifikasi dan dianalisis dengan spesies sebagai atribut target. Pembangunan parameter model didasarkan pada jumlah total dari hasil SMOTE dengan menetapkan 100 dan 500 sebagai jumlah pohon tunggal dan 1000 sebagai jumlah *node* dan peubah *prediktor default*. Hasil akhir menunjukkan bahwa algoritme Random Forest memperoleh nilai evaluasi yang optimal dengan rata-rata 99.97% menggunakan jumlah pohon tunggal dan *cut-off* yang telah ditetapkan. Akurasi maksimal yaitu 100% diperoleh dari jumlah pohon tunggal dan *cut-off* dengan ukuran sebagai berikut: (1) 500 dan 80:20; (2) 500 dan 90:10; dan (3) 100 dan 80:20. Hasil ini menunjukkan bahwa algoritme Random Forest sangat efektif untuk diterapkan sebagai metode pengklasifikasi Machine Learning dalam menentukan spesies mangrove.

Kata kunci: ekosistem pesisir, machine learning, mangrove, normalisasi data, random forest

PENDAHULUAN

Mangrove merupakan sekumpulan tanaman yang tumbuh pada area di antara garis batasan surut terendah serta batasan pasang paling tinggi, ataupun yang biasa disebut dengan area yang berada disekitar bibir pantai (Saenger, 2002). Mangrove adalah salah satu ekosistem sangat produktif serta unik, sebab dapat tumbuh subur serta sanggup bertahan hidup di area transisi antara lautan serta daratan (Robertson & Alongi, 1992).

Mangrove berperan sebagai peredam gejala alam yang ditimbulkan oleh gelombang, badai, abrasi, serta penyangga untuk kehidupan biota lainnya (Thampanya et al., 2006), selain itu ekosistem mangrove juga sebagai tempat *spawning ground*, *nursey ground*, *feeding ground* untuk udang, ikan, kepiting, kerang, serta yang lain (Abu El-Regal & Ibrahim, 2014). Mangrove mempunyai kemampuan dalam menyerap CO₂ ataupun karbon dioksida dari atmosfer (Donato 2011). Hal tersebut digunakan dalam proses fotosintesis untuk pertumbuhannya dan memberikan kontribusi dalam proses pengurangan dampak pergantian iklim yang dialami saat ini (Pramudji 2018).

Menurut Spalding et al (2010), Indonesia merupakan negara yang memiliki populasi mangrove yang beragam. Hal ini didasarkan pada jumlah spesies mangrove yang tumbuh di Indonesia sebanyak 45 spesies. Berdasarkan data dari IUCN hampir semua spesies mengalami penurunan jumlah. Hal tersebut disebabkan oleh pendegradasian lahan akibat penebangan mangrove secara berlebihan dan perubahan fungsi lahan untuk kegiatan pembangunan (Ulfa 2016).

Untuk mengetahui jumlah spesies mangrove pada suatu area hutan mangrove maka perlu pemahaman terkait spesies mangrove yang ada di area tersebut. Cara yang dapat dilakukan untuk mengetahui spesies mangrove tersebut dengan melakukan identifikasi.

Identifikasi hanya dapat dilakukan oleh orang-orang ahli atau menggunakan buku identifikasi mangrove.

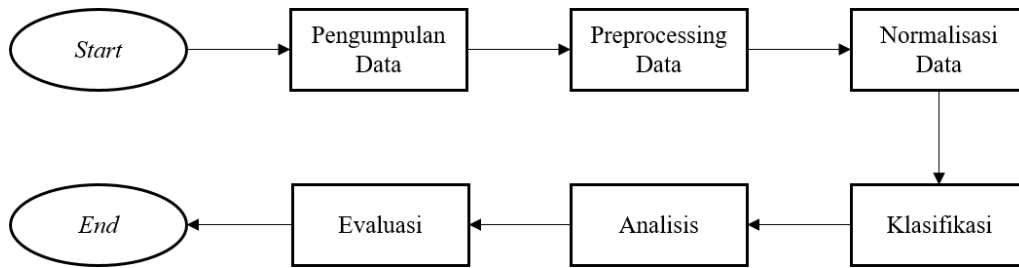
Kemajuan teknologi dapat membantu melakukan identifikasi spesies mangrove, yaitu dengan membuat sistem identifikasi secara otomatis (Ariawan et al., 2020). Salah satu bidang ilmu yang dapat membantu dalam pembuatan sistem identifikasi tersebut adalah *Machine Learning* (ML). *Machine Learning* adalah proses komputer untuk belajar dari data (Dey 2016).

Beberapa penelitian pada bidang ML yang telah berhasil dilakukan untuk mengidentifikasi tumbuhan yaitu Cutler et al., (2007), menggunakan algoritme *Random Forest* (RF) dengan dataset tanaman *invasive* untuk mengidentifikasi spesies *Verbascum Thapsus*, *Cirsium vulgare*, *Urtica dioica*, *Marrubium vulgare*, *Lobaria oregano*, *Lobaria pulmonaria*, *Pseudocyphellaria anomala*, dan *Pseudocyphellaria anomala*. Kemudian, Raczko & Zagajewski, 2017, menggunakan algoritme RF dengan dataset *APEX Hyperspectral* untuk mengklasifikasikan lima spesies pohon yang berada di daerah Szklarska Poręba, yaitu *Picea alba L. Karst*, *Larix decidua Mill*, *Alnus Mill*, *Fagus sylvatica L*, dan *Betula pendula Roth*. Ambarwari et al., (2018), menggunakan algoritme *Support Vektor Machine* (SVM) dengan dataset daun tumbuhan obat untuk mengelompokkan bentuk venasi tumbuhan obat. Serta Ariawan et al 2020, menggunakan algoritme SVM dengan dataset daun shorea untuk mengidentifikasi 4 spesies shorea yaitu *acuminata*, *leprosula*, *ovalis*, dan *selanica*.

Oleh karena itu, penelitian ini berfokus pada identifikasi spesies mangrove (*Avicennia eucalyptifolia*, *Bruguiera gymnorrhiza*, *Rhizophora apiculate*, *Rhizophora mucronata*, *Unrecorded Sp.*, dan *Xylocarpus granatum*) menggunakan algoritme ML, yaitu RF dengan dataset Swamp Data Trees kuburaya dan Buruway (Adame et al., 2017, Murdiyarso et al. 2019). RF adalah algoritme ML yang telah banyak digunakan untuk mengklasifikasikan suatu objek. Metode ini telah sukses digunakan oleh Krishnaveni, 2018 mengklasifikasikan tanaman obat. Akurasi sebesar 99%. Sandika et al., 2016 mengklasifikasikan penyakit anggur. Akurasi sebesar 86%. Serta, Apao et al., 2020 untuk memprediksi tingkat aktivitas penderita stroke. Akurasi sebesar 88.24%.

METODE PENELITIAN

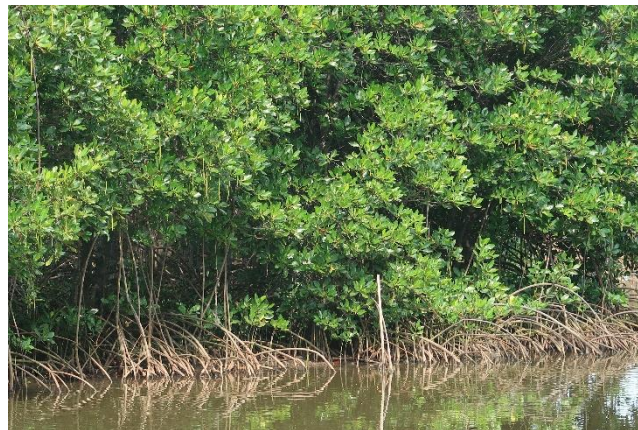
Tahapan penelitian yang dilakukan terdiri dari: pengumpulan data, *preprocessing* data, normalisasi data, klasifikasi, analisis, dan evaluasi. Tahapan ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Pada penelitian ini dataset yang digunakan adalah dataset Swamp Data Trees Kuburaya dan Buruway (dataset mangrove) yang bersumber dari <https://data.cifor.org/>. Data tersebut memiliki beberapa atribut, yaitu spesies, *wood density*, *diameter at breast height*, *total of above ground biomass*, dan *below-ground root*. Atribut spesies digunakan sebagai target klasifikasi. Spesies tersebut adalah *Avicennia eucalyptifolia*, *Bruguiera gymnorrhiza*, *Rhizophora apiculata*, *Rhizophora mucronata*, *Unrecorded Sp.*, dan *Xylocarpus granatum*. Gambar 2 menunjukkan spesies *Rhizophora apiculata*.



Gambar 2. *Rhizophora apiculata*

B. Preprocessing Data

Pada tahap ini dilakukan penyeimbangan data dengan menggunakan algoritme *Synthetic Minority Oversampling Technique* (SMOTE) (Akbar et al., 2020). SMOTE adalah salah satu algoritme atau metode yang paling populer digunakan untuk menangani ketidakseimbangan distribusi data (Kasanah et al. 2019). Metode SMOTE bekerja dengan melakukan pencarian ketetanggaan terdekat dari data sebanyak K dalam setiap data pada kelas minoritas, setelah itu data sintetis dibuat sebanyak presentase duplikasi antara data minor dan ketetanggaan data sebanyak K yang dipilih secara acak (Chawla et al., 2002).

C. Normalisasi Data

Normalisasi data ialah metode yang digunakan untuk mengganti nilai numerik dalam dataset ke skala umum, tanpa mendistorsi perbandingan dalam rentang nilai. Normalisasi data membantu proses pembelajaran pada ML secara cepat (Li & Liu, 2011). Salah satu teknik normalisasi data ialah normalisasi *min-max*. Metode ini berfungsi untuk mengubah ukuran data pada rentang asli, yaitu semua nilai berada dalam kisaran 0 dan 1. Persamaan normalisasi *min-max* ditunjukkan pada persamaan (1) (Li & Liu, 2011).

$$w_{norm} = \left(\frac{w_i - w_{min}}{w_{max} - w_{min}} \right) \quad (1)$$

D. Klasifikasi

RF merupakan salah satu metode klasifikasi yang menggabungkan beberapa metode klasifikasi *tree*. RF dikembangkan oleh Leo Breiman, dalam jurnalnya disebutkan bahwa RF terdiri dari beberapa model klasifikasi *tree*. Setiap *tree* akan menampilkan hasil klasifikasi yang kemudian oleh RF dipilih hasil yang paling banyak muncul (Breiman, 2001).

Algoritme RF bekerja dengan melakukan pencarian *best split* menggunakan perhitungan *gini index* (Breiman, 2001). *Gini index* berfungsi untuk menghitung *impurity* dari sebuah data. Semakin tinggi nilai *gini index* suatu atribut maka semakin berpeluang atribut tersebut menjadi *root node* atau *node* pemisah selanjutnya. *Gini index* juga digunakan pada klasifikasi *tree* untuk menentukan target atau label akhir dalam klasifikasi. Perhitungan *gini index* ditunjukkan oleh Persamaan 1 (Fawagreh et al., 2014).

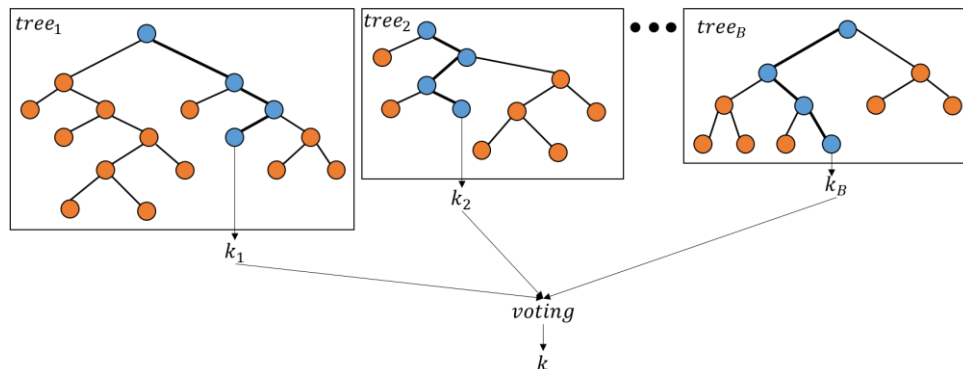
$$Gini(t) = 1 - \sum_{i=1}^N p \quad (1)$$

Hasil akhir klasifikasi RF merupakan penjumlahan atau *voting* dari masing-masing klasifikasi *tree*. Gambar 3 menjelaskan arsitektur dari metode RF. *Tree 1, tree 2, ..., tree b* adalah jumlah dari *tree* dalam RF yang melakukan klasifikasi. Sedangkan k_1, k_2, \dots, k_b adalah label kelas. RF menghasilkan label kelas k , setelah melakukan *voting* dari hasil masing-masing *tree* (Verikas et al., 2011).

Hasil klasifikasi akhir RF merupakan penjumlahan hasil pembelajar awal atau hasil klasifikasi setiap *tree* untuk masing-masing kelas. Kemudian hasil klasifikasi *tree* akan dijumlahkan berdasarkan kelasnya. Terakhir akan diambil kelas dengan jumlah terbanyak untuk kemudian dipilih sebagai hasil klasifikasi akhir. Pemilihan klasifikasi akhir dalam RF dapat dilihat pada Persamaan 2 (A. Cutler et al., 2012).

$$f(x) = \operatorname{arg}_{y \in Y}^{\max} \sum_{j=1}^i I(y = h_j(x)) \quad (2)$$

Dimana $f(x)$ adalah hasil dari klasifikasi RF dan $h_j(x)$ merupakan hasil dari klasifikasi masing-masing *tree*. Sedangkan $I(y = h_j(x))$ merupakan fungsi indikator yang akan memberikan nilai 1 jika hasil klasifikasi *tree* sama dengan kelas y dan nilai 0 jika sebaliknya.



Gambar 3. Arsitektur RF (Verikas et al. 2011).

HASIL DAN PEMBAHASAN

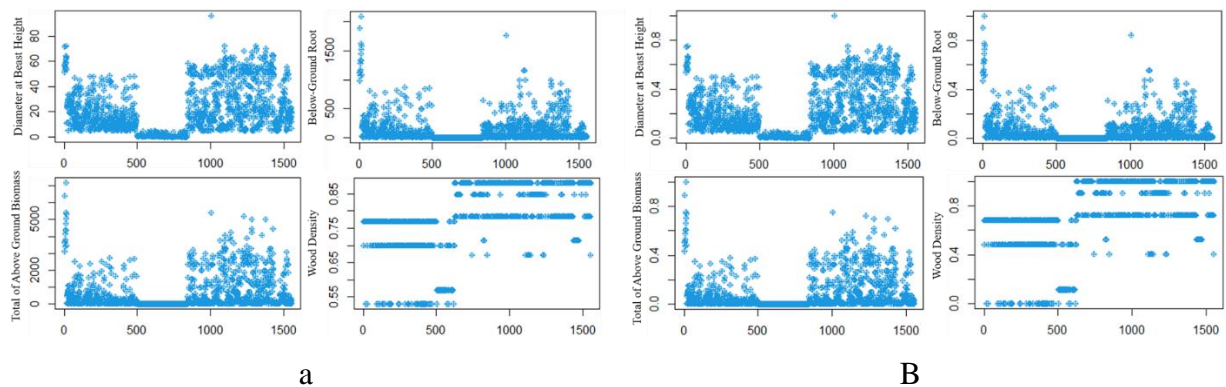
Dataset. yang bersumber dari <https://data.cifor.org/> terdiri dari 6 kelas, dengan rincian jumlah data dari setiap kelas antara lain spesies *Avicennia eucalyptifolia* 42, *Bruguiera gymnorrhiza* 470, *Rhizophora apiculata* 743, *Rhizophora mucronata* 139, *Unrecorded Sp.* 88 dan *Xylocarpus granatum* 72. Berdasarkan rincian data tersebut, dapat dilihat bahwa setiap spesies memiliki data yang tidak seimbang. Oleh karena itu, apabila proses klasifikasi dilakukan, maka data dari spesies *Avicennia eucalyptifolia*, *Unrecorded Sp.*, dan *Xylocarpus granatum* cenderung diklasifikasikan sebagai *noise* atau pencilan (*outlier*) yang dapat merusak hasil pemodelan. Upaya yang dilakukan untuk mengatasi data tidak seimbang tersebut adalah dengan menggunakan algoritme SMOTE. Adapun hasil SMOTE yang telah dilakukan dapat dilihat pada Tabel 1.

Tabel 1. Hasil SMOTE

Spesies mangrove	Jumlah data
<i>Avicennia eucalyptifolia</i>	743
<i>Bruguiera gymnorrhiza</i>	743
<i>Rhizophora apiculata</i>	743
<i>Rhizophora mucronata</i>	743
<i>Unrecorded Sp.</i>	743
<i>Xylocarpus granatum</i>	743
Total	4458

A. Normalisasi Data

Pada ML, setiap dataset tidak harus dilakukan normalisasi. Normalisasi data dapat dilakukan ketika atribut memiliki rentang nilai berbeda seperti pada atribut dari dataset mangrove. Dataset tersebut memiliki rentang nilai berbeda pada setiap atribut-atributnya. Mulai dari kisaran 0 sampai dengan puluhan ribu (Gambar 4a). Perbedaan ini dapat dikecilkan dengan melakukan normalisasi data. Normalisasi data diperlukan untuk mengoptimalkan kinerja dari algoritme RF. Semakin kecil rentang nilai dari setiap atribut maka semakin cepat proses komputasi yang dilakukan. Hasil normalisasi data menggunakan min-max pada nilai dari setiap atribut dataset mangrove ditunjukkan pada Gambar 4b.



Gambar 4. Visualisasi dataset (a) Sebelum Normalisasi dan (b) Setelah Normalisasi

Normalisasi *min-max* mengubah data kedalam interval 0 sampai dengan 1. Berdasarkan Gambar 4 dapat dilihat bahwa pola data tidak ada yang berubah. Perubahan terlihat hanya pada rentang nilai dari setiap atributnya.

B. Klasifikasi

Pada tahap ini klasifikasi dengan metode RF dilakukan dengan menggunakan *package randomForest* yang tersedia pada R. Pembuatan model menggunakan *package* RF yang terdapat pada R ada beberapa langkah.

Langkah pertama adalah mendefinisikan *package* RF untuk memanggil *library* RF yang akan digunakan dalam pembuatan model klasifikasi. Kedua adalah melakukan pengaturan parameter pemodelan, yakni jumlah pohon tunggal yang dibangun sebanyak 500 dan 1000 dengan jumlah *node* dan peubah *prediktor default*. Kemudian model dibuat berdasarkan label sebagai acuan klasifikasi dan *train data* sebagai data latih. Terakhir menganalisa dan mengevaluasi hasil dari prediksi.

Penentuan beberapa *cut-off* (90:10, 80:20, dan 70:30) dilakukan untuk melihat seberapa sensitif kinerja metode RF dalam mengklasifikasikan spesies mangrove sehingga diperoleh nilai evaluasi yang optimal. Tabel 2 menunjukkan hasil perhitungan akurasi klasifikasi yang

diperoleh dari model yang menggunakan jumlah pohon tunggal 100 dan 500 serta nilai evaluasi model validasi silang $k=5$.

Tabel 2. Hasil akurasi model RF menggunakan jumlah pohon tunggal 100 dan 500 serta *cut-off* (90:10, 80:20, dan 70:30)

Fold	Jumlah pohon tunggal = 100			Jumlah pohon tunggal = 500		
	70:30	80:20	90:10	70:30	80:20	90:10
1	100%	100%	100%	99.92%	100%	100%
2	100%	100%	100%	100%	100%	100%
3	99.92%	100%	99.78%	100%	100%	100%
4	99.92%	100%	100%	100%	100%	100%
5	100%	100%	99.78%	100%	100%	100%
Rataan	99.97%	100%	99.91%	99.98%	100%	100%

Hasil evaluasi model RF pada Tabel 2 menunjukkan bahwa nilai rataan akurasi dari seluruh percobaan sangat baik. Hal tersebut dibuktikan dengan akurasi dari seluruh percobaan berada pada presentasi 99 - 100%. Hasil ini juga menunjukkan bahwa kenaikan jumlah pohon tunggal dapat meningkatkan akurasi dari sebuah model.

Berdasarkan hasil evaluasi, Tabel 2 menunjukkan bahwa performa model RF dengan menggunakan dataset mangrove, mendapatkan hasil akurasi yang maksimal dengan menggunakan jumlah pohon tunggal 500 dan *cut-off* 80:20, jumlah pohon tunggal 500 dan *cut-off* 90:10, serta jumlah pohon tunggal 100 dan *cut-off* 80:20, yaitu akurasi 100%.

Selain menggunakan akurasi, untuk mengukur sebuah kinerja model juga dapat menggunakan *precision* dan *sensitivity*. Perhitungan *precision* dilakukan agar dapat diketahui seberapa besar proporsi dari klasifikasi suatu kelas yang benar sedangkan *sensitivity* dilakukan untuk mengetahui seberapa besar proporsi dari kelas sesungguhnya yang terklasifikasi dengan benar. Pengujian *precision* dan *sensitivity* menggunakan jumlah pohon tunggal 500 dan *cut-off* 80:20. Tabel 3 dan 4 menyajikan nilai *precision* dan *sensitivity* masing-masing kelas.

Tabel 3. Nilai rataan *precision* menggunakan jumlah pohon tunggal 500 dan *cut-off* (80:20)

<i>Fold</i>	AEU	BGY	RAP	RMU	USP	XGR	Rataan
1	100%	100%	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%	100%	100%
3	100%	100%	100%	100%	100%	100%	100%
4	100%	100%	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%	100%	100%
Rataan	100%	100%	100%	100%	100%	100%	100%

Tabel 4. Nilai rataan *sensitivity* menggunakan jumlah pohon tunggal 500 dan *cut-off* (80:20)

<i>Fold</i>	AEU	BGY	RAP	RMU	USP	XGR	Rataan
1	100%	100%	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%	100%	100%
3	100%	100%	100%	100%	100%	100%	100%
4	100%	100%	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%	100%	100%
Rataan	100%	100%	100%	100%	100%	100%	100%

Pada Tabel 3 dan 4 menunjukkan bahwa nilai rataan *precision* dan *sensitivity* pada semua kelas adalah 100%. Hal tersebut menunjukkan bahwa secara keseluruhan data sesungguhnya masing-masing kelas dapat terklasifikasi dengan benar mencapai kategori tinggi dan proporsi dari klasifikasi suatu kelas yang benar juga mencapai kategori tinggi.

KESIMPULAN

Berdasarkan hasil dan pembahasan diperoleh akurasi yang optimal dari implementasi algoritme RF untuk mengidentifikasi spesies mangrove. Namun, klasifikasi spesies mangrove tidak dipengaruhi oleh tahapan normalisasi sehingga tidak terjadi perubahan terhadap pola data. Berbagai hasil akurasi dari setiap kombinasi model menunjukkan bahwa tidak terjadinya selisih nilai yang signifikan antar model. Kesimpulan utama dalam penelitian ini, yaitu: (1) RF terbukti sangat akurat dengan pencapaian presentasi 99 - 100% dalam mengidentifikasi spesies mangrove. (2) Terjadinya peningkatan akurasi dengan diimbangi kenaikan jumlah pohon. (3) Transformasi dimensional dari normalisasi min-max tidak berpengaruh terhadap pola data. Secara keseluruhan, hasil penelitian ini menunjukkan algoritme RF sebagai alat yang potensial untuk mengidentifikasi spesies mangrove dalam level individual.

DAFTAR PUSTAKA

- K. P., & Krishnaveni, K. (2018). *Random Forest Classification of Medicinal Plant Leaves Using Shape and Texture Features*.
- Abu El-Regal, M. A., & Ibrahim, N. K. (2014). Role of mangroves as a nursery ground for juvenile reef fishes in the southern Egyptian Red Sea. *Egyptian Journal of Aquatic Research*, 40(1), 71–78. <https://doi.org/10.1016/j.ejar.2014.01.001>
- Adame, M. F., Cherian, S., Reef, R., & Stewart-Koster, B. (2017). Mangrove root biomass and the uncertainty of belowground carbon estimations. In *Forest Ecology and Management* (Vol. 403, pp. 52–60). Elsevier B.V. <https://doi.org/10.1016/j.foreco.2017.08.016>
- Akbar, A. T., Husaini, R., Akbar, B. M., & Saifullah, S. (2020). A proposed method for handling an imbalance data in classification of blood type based on Myers-Briggs type

- indicator. *Jurnal Teknologi Dan Sistem Komputer*, 8(4), 276–283. <https://doi.org/10.14710/jtsiskom.2020.13625>
- Ambarwari, A., Herdiyeni, Y., & Hermadi, I. (2018). Biometric analysis of leaf venation density based on Digital Image. *Telkonnika (Telecommunication Computing Electronics and Control)*, 16(4), 1735–1744. <https://doi.org/10.12928/TELKOMNIKA.v16i4.7322>
- Apao, N. J., Feliscuzo, L. S., Lyn Sta Romana, C. C., & Aurea Tagaro, J. S. (n.d.). *Multiclass Classification Using Random Forest Algorithm To Prognosticate The Level Of Activity Of Patients With Stroke*. www.ijstr.org
- Ariawan, I., Herdiyeni, Y., & Siregar, I. Z. (2020). Geometric morphometric analysis of leaf venation in four shorea species for identification using digital image processing. *Biodiversitas*, 21(7), 3303–3309. <https://doi.org/10.13057/biodiv/d210754>
- Breiman, L. (2001). *Random Forests* (Vol. 45).
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. In *Ensemble Machine Learning* (pp. 157–175). Springer US. https://doi.org/10.1007/978-1-4419-9326-7_5
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY. In *Ecology* (Vol. 88, Issue 11).
- Dey, A. (n.d.). *Machine Learning Algorithms: A Review*. www.ijcsit.com
- Donato 2011. Mangrove adalah salah satu hutan terkaya karbon di kawasan tropis. www.cifor.org
- Fawagreh, K., Gaber, M. M., & Elyan, E. (2014). Random forests: From early developments to recent advancements. *Systems Science and Control Engineering*, 2(1), 602–609. <https://doi.org/10.1080/21642583.2014.956265>
- Kasanah AN, Muladi, Pujianto U. 2019. Penerapan Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Objektivitas Berita Online Menggunakan Algoritma KNN *Jurnal Rekayasa Sistem dan Teknologi Informasi*. 3(2):196-201
- Li, W., & Liu, Z. (2011). A method of SVM with normalization in intrusion detection. *Procedia Environmental Sciences*, 11(PART A), 256–262. <https://doi.org/10.1016/j.proenv.2011.12.040>
- Pramudji. 2018. Mangrove di Indonesia. <http://lipi.go.id/publikasi/mangrove-di-indonesia-/27339>
- Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing*, 50(1), 144–154. <https://doi.org/10.1080/22797254.2017.1299557>
- Robertson, A. I. (Alistar I. ., & Alongi, D. M. (Daniel M. . (1992). *Tropical mangrove*

ecosystems. American Geophysical Union.

- Saenger, P. (2002). Mangrove Ecology, Silviculture and Conservation. In *Mangrove Ecology, Silviculture and Conservation*. Springer Netherlands. <https://doi.org/10.1007/978-94-015-9962-7>
- Sandika, B., Avil, S., Sanat, S., & Srinivasu, P. (2016). Random forest based classification of diseases in grapes from images captured in uncontrolled environments. *International Conference on Signal Processing Proceedings, ICSP, 0*, 1775–1780. <https://doi.org/10.1109/ICSP.2016.7878133>
- Spalding, M., M. Kainuma, and L. Collins (2010) World Atlas of Mangrove. Earthscan. London, UK.
- Thampanya, U., Vermaat, J. E., Sinsakul, S., & Panapitukkul, N. (2006). Coastal erosion and mangrove progradation of Southern Thailand. *Estuarine, Coastal and Shelf Science*, 68(1), 75–85. <https://doi.org/10.1016/j.ecss.2006.01.011>
- Verikas, A., Gelzinis, A., & Bacauskiene, M. (2011). Mining data with random forests: A survey and results of new tests. *Pattern Recognition*, 44(2), 330–349. <https://doi.org/10.1016/j.patcog.2010.08.011>