**VARIANSI: Journal of Statistics and Its Application on Teaching and Research (Online)** **ISSN 2684-7590**
**Vol. 4 No. 2 (2022), 121-127**

DOI: 10.35580/variansiunm31

# Penerapan Metode *Random Forest* Untuk Klasifikasi Varian Minuman Kopi Di Kedai Kopi Konijiwa Bantaeng

**Suci Amaliah[1], Muhammad Nusrang[2*], Aswi[3]**
*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia*

Kata Kunci: *Random Forest*, Varian Minuman, Akurasi.

**Abstrak:**

*Random Forest* (RF) adalah metode yang dapat meningkatkan hasil akurasi dalam membangkitkan atribut untuk setiap node yang dilakukan secara acak. Penelitian ini bertujuan untuk mengetahui tingkat akurasi metode RF dalam memprediksi varian minuman kopi di kedai Konijiwa Bantaeng yang paling diminati pelanggan. Berdasarkan hasil analisis diperoleh bahwa model dengan error klasifikasi terkecil adalah dengan menggunakan mtry 2 dan ntree 500. Model yang dihasilkan dievaluasi dengan menggunakan *confusion matrix* dimana diperoleh bahwa varian minuman kategori *coffee based* lebih diminati daripada *signature coffee* dengan nilai akurasi sebesar 94,12%.

## 1. Latar Belakang

Dalam klasifikasi suatu data, salah satu metode yang dapat digunakan adalah metode *Decision Tree* yang merupakan salah satu metode pengambilan keputusan. Pertumbuhan pada metode *Decision Tree* dapat ditumbuhkan sampai mencapai kehomogenan, tetapi hal tersebut dapat mengakibatkan *overfitting*, sehingga diperlukan suatu metode yang bisa mengatasi masalah tersebut. Metode yang dapat digunakan untuk menghindari *overfitting* adalah metode dengan memangkas pohon (*pruning*) atau menggunakan metode *Random Forest* (RF).

RF adalah metode klasifikasi dalam statistika yang berbasis komputasi. Metode klasifikasi digunakan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih ke dalam salah satu dari kategori kelas yang telah ditetapkan. Memasuki era *big data*, penggunaan metode statistika berbasis komputasi sangat banyak digunakan. Metode RF dipilih karena menghasilkan kesalahan yang lebih rendah, memberikan akurasi yang baik dalam klasifikasi, dapat menangani data yang jumlahnya sangat besar, dan efektif untuk mengatasi data yang tidak lengkap (Breiman, 2001). Pada penelitian ini, metode RF diterapkan dalam mengkasifikasi varian minuman kopi yang paling diminati di kedai Konijiwa serta tingkat akurasi dari hasil klasifikasi tersebut.

## 2. Kajian Pustaka

### 2.1 *Random Forest*

*Random Forest* (RF) pertama kali diperkenalkan oleh Leo Breiman (2001). RF merupakan salah satu metode yang dapat meningkatkan hasil akurasi dalam membangkitkan atribut untuk setiap node yang dilakukan secara acak. RF terdiri dari sekumpulan *decision tree*, dimana kumpulan pohon keputusan ini digunakan untuk mengklasifikasi data ke suatu kelas. Pohon keputusan dibuat dengan menentukan node akar dan berakhir dengan beberapa node daun untuk mendapatkan hasil akhir (Debby & Rahman, 2020).

Membentuk pohon keputusan pada metode RF sama dengan proses pada Classsification and Regression Tree (CART), hanya saja pada RF tidak dilakukan *prunning* (pemangkasan). Indeks Gini digunakan unruk memilih fitur di setiap simpul internal dari pohon keputusan. Nilai Indeks Gini dapat dihitung sebagai berikut:

$$\text{Gini}(S_i) = 1 - \sum_{i=0}^{c-1} p_i^2 \tag{2.1}$$

dengan $p_i$ merupakan frekuensi relative kelas $C_i$ di dalam set.
$C_i$ merupakan kelas untuk $i$ = 1, …, c-1, dan $c$ adalah jumlah kelas yang telah ditentukan.
Kualitas split pada fitur $k$ ke dalam subset $S_i$ merupakan jumlah sampel milik kelas $C_i$, kemudian dihitung sebagai jumlah pertimbangan indikasi Gini dari subset yang dihasilkan. Data dapat dihitung dengan rumus sebagai berikut:

$$\text{Gini}_{split} = \sum_{i=0}^{k-1} \left(\frac{n_i}{n}\right) \text{Gini}(S_i) \tag{2.2}$$

dimana $n_i$ merupakan jumlah sampel dalam subset $S_i$ setelah di split dan $n$ merupakan jumlah sampel di node yang dberikan.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

[1] Corresponding author.
*E-mail address*: muh.nusrang@unm.ac.id

Misalkan $\{h(x, \Theta_k), k= 1, \ldots\}$ dimana $\{ \Theta_k\}$ merupakan vector *random* yang *independent identically distributed* (iid) dan tiap pohon memilih kelas yang paling banyak dari rata rata (*majority vote*). Untuk RF, batas atas dapat diturunkan untuk kesalahan generalisasi dalam hal dua parameter yang mengukur seberapa kuat pengklasifikasian individu dan ketergantungan diantara keduanya (Breiman, 2001):
Fungsi margin untuk RF adalah

$$\mathrm{mr}(\mathrm{X},\mathrm{Y}) = \mathrm{P}_{\Theta}(\mathrm{h}\,(\mathrm{X},\Theta) = \mathrm{Y}) - \max_{j \neq Y} P_{\Theta}\,(\mathrm{h}\,(\mathrm{X},\Theta) = \mathrm{j}) \tag{2.3}$$

dan kekuatan himpunan pengklasifikasi $\{\mathrm{h}\,(\mathrm{X},\Theta)\}$ adalah

$$s = E_{X,Y} mr(X, Y) \tag{2.4}$$

Dengan asumsi $s \geq 0$, ketidaksamaan *Chebychev* serta penurunan variansi *mr* dari fungsi margin untuk metode RF, akan didapatkan persamaan batas atas kesalahan generalisasi sebagai berikut:

$$PE \leq \frac{\bar{\rho}(1-s^2)}{s^2} \tag{2.5}$$

Dimana $\bar{\rho}$ adalah nilai rata-rata korelasi, yaitu:

$$\bar{\rho} = \frac{E_{\Theta,\Theta'}\left(\rho(\Theta,\Theta')sd(\Theta)sd(\Theta')\right)}{E_{\Theta,\Theta'}\left(sd(\Theta)sd(\Theta')\right)} \tag{2.6}$$

## 2.2 *Confusion Matrix*

*Confusion matrix* adalah tabel yang digunakan untuk melihat akurasi serta seberapa baik algoritma yang dihasilkan dari klasifikasi yang sudah dibuat untuk mengklasifikasi dan memprediksi atribut dari data testing. Metode ini dikembangkan sebagai penilaian algoritma *machine learning* yang diterapkan dalam menyelesaikan masalah klasifikasi. Dalam *confusion matrix* terdapat *False Negative* (FN), *False Positive* (FP), *True Negative* (TN), dan *True Positive* (TP). Berikut merupakan tabel dari *confusion matrix*.

**Tabel 2.1**: Asumsi *Confusion Matrix*

| Kelas Prediksi | Kelas Aktual | |
|---|---|---|
| | *Positive* | *Negative* |
| *Positive* | TP | FP |
| *Negative* | FN | TN |

TP adalahkondisi dimana baik prediksi maupun nilai aktualnya benar; FN adalah kasus dimana nilai prediksi tidak benar tetapi nilai aktualnya benar; FP adalah kasus dimana nilai prediksi benar tapi nilai aktualnya tidak benar. Dalam mengevaluasi kinerja model, ada berbagai macam performa diantaranya akurasi, *recall*, dan presisi. Nilai akurasi, recall, dan presisi dapat diperoleh dengan menggunakan persamaan pada Tabel 2.2.

**Tabel 2.2**: Rumus Evaluasi Performa Metode

| *Performance Metrics* | Rumus |
|---|---|
| Akurasi | $\frac{TP+TN}{TP+TN+FP+FN} \times 100\%$ |
| Recall | $\frac{TP}{TP+FN} \times 100\%$ |
| Presisi | $\frac{TP}{TP+FP} \times 100\%$ |

Akurasi merupakan rasio prediksi benar dengan keseluruhan data. *Recall* adalah rasio prediksi benar positif dibandingkan dengan keseluruhan data aktual positif. Presisi merupakan rasio prediksi benar positif dibandingkan dengan seluruh data yang diprediksi positif.

### 2.3 Minuman Kopi

Kopi salah satu komoditas yang banyak dibudidayakan di Indonesia. Kopi merupakan minuman yang cukup digemari, dilihat dari banyaknya produksi minuman berbahan dasar kopi. Provinsi Sulawesi Selatan merupakan salah satu provinsi di Indonesia yang menjadi penghasil produk kopi. Terdapat beberapa kabupaten di Provinsi Sulawesi Selatan yang menjadi penghasil produk kopi diantaranya Kabupaten Enrekang, Kabupaten Toraja Utara, Kabupaten Sinjai dan Kabupaten Bantaeng. Di era millennial ini, *cafe* atau kedai kopi hampir bisa kita temui di segala tempat serta berbagai daerah. Dari banyaknya peminat kopi tersebut menjadi landasan utama para pengusaha warung kopi atau kedai kopi untuk menciptakan inovasi dalam mengelola serta menyajikan minuman kopi, warung kopi atau kedai kopi yang membuat varian minuman kopi yang menjadi ciri khas dari kedai kopi itu sendiri. Salah satu contoh, kedai kopi Konijiwa yang bertempat di kabupaten Bantaeng menyediakan berbagai varian minuman kopi. Pada kedai tersebut, terdapat 2 kategori menu yaitu *coffee based* dan *signature coffee*. *Coffee based* terdiri dari affogato, americano, cafelatte rum, cafelatte cream cheese, cafelatte, cafelatte caramel, cafelatte jasmine, cafelatte vanilla, cappuccino, chocolate coffee, espresso, espresso doubel, espresso martini, espresso single, dan moccacino, sedangkan *Signature coffee* terdiri dari avocado coffee, kopi susu klasik, kopi susu konijiwa, kopi susu konijiwa seliter, dan kopi susu pandan.

## 3. Metode Penelitian

Penelitian yang digunakan yaitu pendekatan kuantitatif. Data yang digunakan merupakan data penjualan di kedai kopi Konijiwa pada bulan Januari–Maret tahun 2022. Variabel yang digunakan dalam penerapan metode RF ada dua yaitu variabel independen dan variabel dependen. Variabel independen terdiri dari variabel order, harga, cabang, dan usia, sedangkan variabel dependen adalah varian minuman kopi di kedai konijiwa. Pada proses klasifikasi data menggunakan metode RF. Data yang sudah didapatkan dibagi menjadi dua bagian yaitu data *training* dan data *testing*. Metode RF digunakan pada data *training* untuk menghasilkan algoritma. Kemudian algoritma yang terbentuk dievaluasi menggunakan *confusion matrix* dengan memakai data *testing*.

## 4. Hasil dan Pembahasan

### 4.1 *Random Forest* (RF) OOB estimate of error rate

Analisis yang dilakukan menggunakan tiga pembagian data training dan data testing yaitu data *training* 60%, 70%, dan 80%. Pada data *training* 60% terdapat 70 data dari 116 data keseluruhan; data training 70% terdapat 82 data dari 116 data, sedangkan data training 80% terdapat 93 data dari 116 jumlah keseluruhan data. Mencari nilai mtry dilakukan dengan cara $\sqrt{p}$ dimana p merupakan variabel independen, maka nilai mtry sebesar 2. Dengan nilai mtry = 2, dapat dikatakan bahwa pada proses pengklasifikasi terdapat 2 variabel yang dibandingkan untuk menentukan hasil splitting atau pemisahan terbaik.

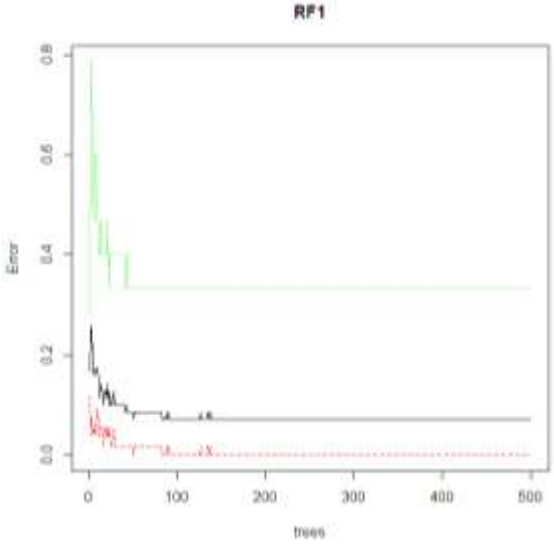
Penentuan ntree yang terbaik dapat dilakukan dengan melihat nilai error yang telah stabil atau tidak lagi berubah pada ntree yang ditentukan. Pada proses klasifikasi dengan menggunakan metode RF pada data penjualan varian minuman kopi, sudah stabil pada ntree ke-500, maka proses klasifikasi dilanjutkan ke tahap prediksi. Selain pohon yang terbentuk (ntree) serta mtry, juga terdapat nilai rata -rata *error* dari *Our of Bag* (OOB) yang menggambarkan estimasi *error* atau tingkat kesalahan dari proses klasifikasi dalam melakukan prediksi.

**Tabel 4.1.** Nilai *Error* OOB Berdasarkan Pembagian Data *Training*

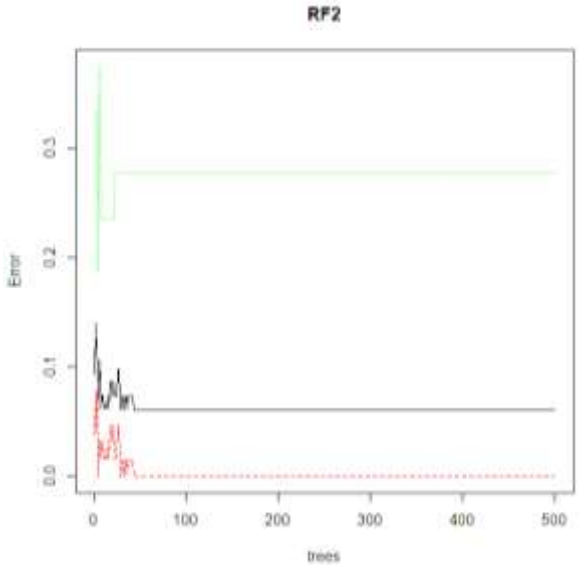
| ***Training*** | ***Error* OOB** |
|---|---|
| 60% | 7,14% |
| 70% | 3,66% |
| 80% | 4,30% |

Berdasarkan hasil analisis metode RF pada data *training* didapatkan nilai *error* OOB = 7,14% pada pembagian data *training* 60%, yang berarti nilai OOB tersebut merupakan peluang kesalahan klasifikasi metode RF yang diterapkan pada data penjualan di kedai kopi Konijiwa. Sedangkan pada pembagian data raining 70% dan 80% masing masing memberikan nilai error OOB sebesar 3,66% dan 4,3%. Berdasarkan hasil analisis diperoleh bahwa model dengan error klasifikasi terkecil adalah dengan menggunakan data training 70% dengan menggunakan mtry 2 dan ntree 500.

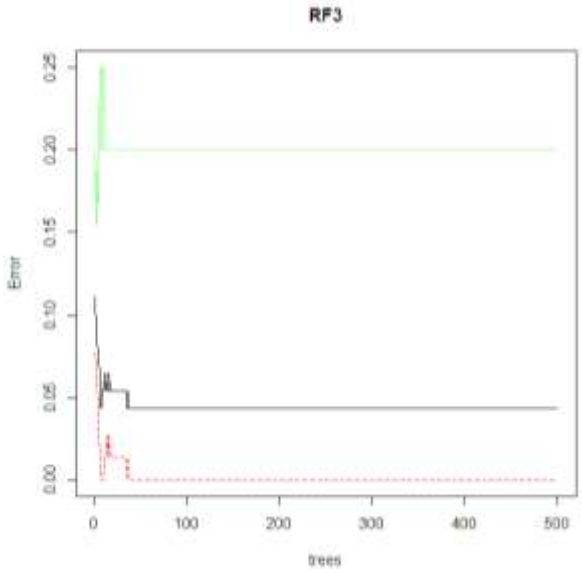
Dalam melakukan prediksi untuk analisis RF dapat dilihat pada plot dari jumlah pohon (ntree). Plot ntree dapat ditunjukkan pada Gambar 4..




**Gambar 4.1.** Plot Ntree dari Pembagian Data *Training* 60% dan Data *Testing* 40%




**Gambar 4.2.** Plot Ntree dari Pembagian Data *Training* 70% dan Data *Testing* 30%




**Gambar 4.3.** Plot Ntree dari Pembagian Data *Training* 80% dan Data *Testing* 20%

Berdasarkan Gambar 4.1, 4.2 dan 4.3 dapat disimpulkan bahwa proses klasifikasi sudah stabil pada ntree ke-500. Sehingga, dapat dikatakan bahwa model yang didapatkan pada analisis RF dapat dilanjutkan ke tahap prediksi dengan menggunakan *confusion matrix*.

### 4.2 *Confusion Matrix*

Metode yang digunakan untuk evaluasi adalah *confusion matrix* dengan menggunakan data *testing*. Berikut merupakan hasil evaluasi serta perhitungan manual nilai akurasi, recall, dan presisi berdasarkan pembagian data *training testing*:

**Tabel 4.1** Hasil *Confusion Matrix* Data *Testing* 40%

| Kelas Prediksi | Kelas Aktual | |
|---|---|---|
| | *Coffee Based* | *Signature Coffee* |
| *Coffee Based* | 35 | 2 |
| *Signature Coffee* | 1 | 8 |

Pada hasil analisis *confusion matrix* pada data *testing* 40%, menunjukkan bahwa kelas *positive* adalah *coffee based*, dengan nilai TP sebanyak 35 data pada kelas aktual *coffee based* berhasil diprediksi sebagai *coffee based*. Nilai FP sebanyak 2 data yang menunjukkan terdapat 2 data pada kelas aktual *signature coffee* tetapi di prediksi sebagai *coffee based*. Nilai FN sebanyak 1 data, sehingga dapat dikatakan terdapat 1 kategori aktual *coffee based* yang diprediksi sebagai *signature coffee*, dan nilai TN sebanyak 8 data yang kelas aktual *signature coffee* berhasil diprediksi sebagai *signature coffee*.

**Tabel 4.2** Hasil *Confusion Matrix* Data *Testing* 30%

| Kelas Prediksi | Kelas Aktual | |
|---|---|---|
| | *Coffee Based* | *Signature Coffee* |
| *Coffee Based* | 27 | 2 |
| *Signature Coffee* | 0 | 5 |

Pada hasil analisis *confusion matrix* pada data *testing* 30%, menunjukkan bahwa kelas *positive* adalah *coffee based*, dengan nilai TP sebanyak 27 data pada kelas aktual *coffee based* berhasil diprediksi sebagai *coffee based*. Nilai FP sebanyak 2 data yang menunjukkan terdapat 2 data pada kelas aktual *signature coffee* tetapi di prediksi sebagai *coffee based*. Nilai FN sebanyak 0 data, sehingga dapat dikatakan tidak terdapat aktual *coffee based* yang diprediksi sebagai *signature coffee*, dan nilai TN sebanyak 5 data yang kelas aktual *signature coffee* berhasil diprediksi sebagai *signature coffee*.

**Tabel 4.3** Hasil *Confusion Matrix* Data *Testing* 20%

| Kelas Prediksi | Kelas Aktual | |
|---|---|---|
| | *Coffee Based* | *Signature Coffee* |
| *Coffee Based* | 18 | 3 |
| *Signature Coffee* | 0 | 2 |

Pada hasil analisis *confusion matrix* pada data *testing* 20%, menunjukkan bahwa kelas *positive* adalah *coffee based*, dengan nilai TP sebanyak 18 data pada kelas aktual *coffee based* berhasil diprediksi sebagai *coffee based*. Nilai FP sebanyak 3 data yang menunjukkan terdapat 3 data pada kelas aktual *signature coffee* tetapi di prediksi sebagai *coffee based*. Nilai FN sebanyak 0 data, sehingga dapat dikatakan tidak terdapat aktual *coffee based* yang diprediksi sebagai *signature coffee*, dan nilai TN sebanyak 2 data yang kelas aktual *signature coffee* berhasil diprediksi sebagai *signature coffee*.

Pada proses evaluasi model RF menggunakan *confusion matrix*, kategori yang menjadi TP adalah *Coffee Based* serta memiliki nilai kelas prediksi yang paling banyak berdasarkan kelas aktualnya pada setiap pembagian data *training* dan data *testing*. Perbandingan nilai akurasi, *recall*, dan presisi dapat dilihat pada tabel 4.4.

**Tabel 4.4** Nilai akurasi, *recall*, dan Presisi

| Evaluasi | Pembagian Data *Training Testing* | | |
|---|---|---|---|
| | 60% : 40% | 70% : 30% | 80% : 20% |
| Akurasi | 93,48% | 94,12% | 86,96% |
| *Recall* | 97,22% | 100% | 100% |
| Presisi | 94,59% | 93,1% | 85,71% |

Pada data *testing* 60:40%, nilai akurasi yang didapatkan menunjukkan bahwa tingkat kepercayaan performa metode RF dalam memprediksi kelas sebesar 93,48%. Kemudian terdapat nilai recall yang menunjukkan bahwa tingkat kepercayaan kemampuan metode RF dalam menentukan kembali informasi dari data sebesar 97,22%. Serta, nilai presisi menunjukkan bahwa tingkat kemampuan metode RF dalam menggambarkan data yang diminta dengan hasil prediksi sebesar 94,59%. Begitupun penjelasan bagi data *training testing* sebanyak 70% : 30% dan 80% : 20%.

Dapat dilihat bahwa nilai-nilai akurasi yang paling tinggi terdapat pada pembagian data *training testing* sebanyak 70%:30% dengan tingkat akurasi sebesar 94,12% yang berarti persentase nilai tersebut menggambarkan seberapa akurat model dalam mengklasifikasi dengan benar, nilai *recall* sebesar 100% pada pembagian data 70% : 30% dan 80% : 20% yang menggambarkan keberhasilan model dalam menetukan kembali sebuah informasi, serta nilai presisi sebesar 94,59% pada pembagian data *training testing* 60% : 40% yang berarti nilai tersebut menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan model.

## 5. Kesimpulan

Berdasarkan hasil analisis diperoleh bahwa model dengan error klasifikasi terkecil adalah dengan menggunakan mtry 2 dan ntree 500. Model yang dihasilkan dievaluasi dengan menggunakan confusion matrix dimana diperoleh bahwa varian minuman kategori *coffee based* lebih diminati daripada *signature coffee* dengan nilai akurasi sebesar 94,12%. Jadi, klasifikasi menggunakan metode RF pada data penjualan minuman kopi di kedai kopi konijiwa Bantaeng dapat dikatakan akurat berdasarkan tingkat kepercayaan atau nilai akurasi yang tinggi.

## Referensi

Adrian, M., R., Putra, M., P., Rafialdy, M., H., dan Rakhmawati, N., A.. (2021). Perbandingan Metode Klasifikasi Random Forest Dan SVM Pada Analisis Sentimen PSBB. *Jurnal Informatika Upgris*. 7(1): 36–40.

Badan Pusat Statistik. (2020). *Statistik Kopi Indonesia*. Indonesia.

Breiman, L. (2001). *Random Forests*. *Manufactured in The Netherlands*. 5–32.

Debby, Alita, & Rahman, A.. (2020). Pendeteksian Sarkasme Pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi*. 8(2): 50–58.

Devella, Siska, Yohannes, & Rahmawati, F. N.. (2020). Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT. *Jurnal Teknik Informatika dan Sistem Informasi*. 7(2): 310–20.

Fadilah, Laili. (2018). *Klasifikasi Random Forest Pada Data Imbalanced*. Jakarta: Unuversitas Islam Negeri Syarif Hidayatullah.

Fachruddin, M. I. (2015). *Classification Dan Support Vector Machine Untuk Deteksi Epilepsi Menggunakan Data Rekaman Electroencephalograph ( EEG )*. Surabaya: Institut Teknologi Sepuluh Nopember.

Fanani, Zainal, N., Sooai, A. G., Sumpeno, S., & Purnomo, M. H.. (2021). Penentuan Kemampuan Motorik Halus

Anak Dari Proses Menulis Hanacaraka Menggunakan Random Forest. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*. 9(2): 148-154

Genuer, Robin. (2020). *Use R ! Random Forests with R*. Switzerland: Springer Nature

Geneur, R., Poggi, J. M., & Malot, C., T. (2009). Variable Selection using Random Forests. France : Laboratoire de Mathematiques, Universite Paris-Sud 11, Bat. 425, 91405.

Grandini, Margherita, Bagli, E., & Visani, G.. 2020. *Metrics for Multi-Class Classification: An Overview*. Italy: Bologna (BO). 1–17.

Hanun, Nugraha, L., & Zailani, A., U,. (2020). Penerapan Algoritma Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera. *Journal of Technology Information*. 6(1): 7–14.

Haristu, R., & Rosa, P.,H., P.. (2019). Penerapan Metode Random Forest Untuk Prediksi Win Ratio Pemain Player Unknown Battleground. *Jurnal MEANS (Media Informasi Analisa dan Sistem)*. 4(2): 120–28.

Kahpi, Ashabul. (2017). Budidaya Dan Produksi Kopi Di Sulawesi Bagian Selatan Pada Abad Ke-19. *Lensa Budaya*. 12(1): 13–26.

Kantardzic, Mehmed. (2020). *DATA MINING Concepts, Models, Methods, and Algorithms*. Canada: Wiley.

Luque, Amalia, Carrasco, A., Martín, A., & Ana De. (2019). The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix. *Jurnal Homepage*. 91: 216–31.

Markoulidakis, Ioannis et al. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *Article Technologies MDPI*.

Octhaviani, Shisillia, & Sibarani, H., J.. (2021). Analisis Pengaruh Citra Merek Produk, Kualitas Produk, Dan Harga Diskon Terhadap Keputusan Pembelian Produk Minuman Kopi Pada Pengguna Aplikasi Grab Food Di Kota Medan. *Scientific Journal*. 4(3): 521–28.

Pradadimara, Dias. (2015). Rice in Colonial and Post-Colonial Southeast Asia, *Paramita*, 25, (1).

Prasetyo, W. B. (2020). 2020 Kedai Kopi Diprediksi Tumbuh 15%. beritasatu.com. https://www.beritasatu.com/ekonomi/601687/2020-kedai-kopidiprediksi-tumbuh-15.

Primajaya, Aji, & Betha N., S.. (2018). Random Forest Algorithm for Prediction of Precipitation. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*. 1(1): 27–31.

Siburian, Wanika, V., & Mulyana, I., E.. (2018). *Prediksi Harga Ponsel Menggunakan Metode Random Forest*. Repository Proceeding Seminar Fakultas Imu Komputer. 4(1): 978–79.

Tebibel, B., Thouraya, & Rubin, H., S.. (2016). *Theoretical information reuse dan integration*. Switzerland: Springer.