

SENASTIKA Universitas Malikussaleh

KLASIFIKASI JENIS TANAH YANG SESUAI TERHADAP TANAMAN PANGAN MENGGUNAKAN METODE RANDOM FOREST DI KABUPATEN MANDAILING NATAL

Yulia Fatimah^{*1}, Asrianda², Rini Meiyanti³

^{1,2,3} Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh

Email: ¹yulia.200170224@mhs.unimal.ac.id, ²asrianda@unimal.ac.id, ³rinimeiyanti@unimal.ac.id

Abstrak

Pertanian di Indonesia, terutama di Kabupaten Mandailing Natal, memainkan peran sentral dalam mendukung perekonomian dan ketahanan pangan. Keberhasilan pertanian sangat dipengaruhi oleh pemilihan jenis tanah yang sesuai. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi jenis tanah yang cocok untuk tanaman pangan dengan menerapkan algoritma *Random Forest* guna membantu petani memilih jenis tanah yang tepat. Adapun langkah – langkah yang dilakukan pada penelitian ini yaitu studi literatur, data *collecting*, data *preprocessing*, analisis & *modelling* dan *evaluation* sistem dan implementasi sistem. Pada penelitian ini variabel/kriteria yang digunakan antara lain drainase, nilai kelembapan, kelembapan, tekstur tanah, dan unsur hara dan jenis tanaman sebagai input untuk memberikan rekomendasi jenis tanah yang paling sesuai.

Keywords: Klasifikasi, Pertanian, *Random Forest*

1. PENDAHULUAN

Dinas Pertanian di Kabupaten Mandailing Natal merupakan unsur pelaksana Otonomi daerah, yang mempunyai tugas menyelenggarakan kewenangan desentralisasi dan dekonsentrasi di bidang pertanian dan juga sebagai suatu organisasi pemerintah yang mengemban tugas meningkatkan kesejahteraan masyarakat melalui peningkatan ketersediaan kebutuhan masyarakat umum, yaitu ketersediaan hasil-hasil pertanian yang menjadi kebutuhan pokok manusia lebih khusus kebutuhan masyarakat yang ada disekitar.

Klasifikasi adalah teknik pengolahan data yang membagi objek menjadi beberapa kelas sesuai dengan jumlah kelas yang diinginkan [1]. Klasifikasi merupakan suatu teknik menemukan suatu pola yang mampu memisahkan kelas data yang satu dengan yang lainnya untuk menentukan objek yang masuk dengan kategori tertentu dengan melihat kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini mampu mengklarifikasi data baru dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Klasifikasi melibatkan penggunaan berbagai algoritma seperti Naïve Bayes, Support Vector Machine, Decision Tree, Random Forest, dan lain-lain.

Random Forest merupakan metode klasifikasi yang terdiri dari kombinasi pohon klasifikasi (CART) yang saling independen, berasal dari distribusi yang sama, dan melakukan proses *voting* (menggunakan aturan jumlah terbanyak) untuk mendapatkan prediksi klasifikasi.

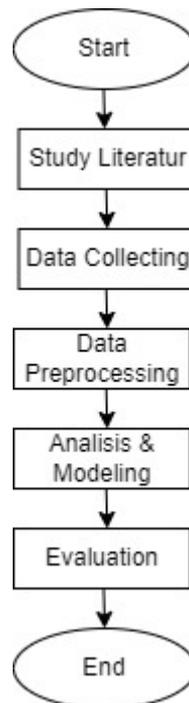
Metode *random forest* biasa digunakan dalam menangani masalah klasifikasi. *Random Forest* memiliki beberapa kelebihan dibandingkan dengan metode lain. Kelebihan dari metode *random forest* seperti memiliki tingkat akurasi yang tinggi, dan lain-lain.

Penelitian ini dilakukan untuk mengetahui jenis tanah mana yang cocok untuk tanaman pangan di Kabupaten Mandailing Natal. Tujuan dari penelitian adalah agar para petani dapat mengetahui jenis tanah yang cocok untuk tanaman pangan di Kabupaten Mandailing Natal dan menghasilkan sebuah sistem yang dapat mengklasifikasi jenis tanah untuk tanaman pangan dengan menggunakan metode Random Forest.

2. METODE PENELITIAN

Penelitian ini dilakukan di Dinas Pertanian Kabupaten Mandailing Natal, Sumatra Utara. Teknik pengumpulan data yang digunakan untuk mendapatkan data-data sehingga data dapat diolah dan dianalisa. Teknik yang digunakan ada dua yaitu data primer dan data sekunder. Data primer adalah data yang didapatkan secara langsung pada objek penelitian, data yang diambil dari Dinas Pertanian Kabupaten Mandailing Natal. Data sekunder yaitu yang diperoleh dari beberapa kajian Pustaka yang berhubungan dengan penelitian dengan mencari bahan dari jurnal, buku dan artikel ilmiah.

2.1 Langkah Penelitian



Gambar 1. Langkah Penelitian

Penelitian ini dimulai dengan studi literatur dari berbagai sumber untuk memberikan landasan teori yang relevan, kemudian dilanjutkan dengan data *collecting* (pengumpulan data), mengumpulkan data-data yang akan digunakan Dinas Pertanian Kabupaten Mandailing Natal. Selanjutnya melakukan *cleaning* data, yaitu memilih variable-variabel yang akan digunakan sebelum dilakukan proses pengujian klasifikasi dengan metode yang diusulkan. Selanjutnya melakukan analisis dan modelling terhadap seluruh data yang telah dikumpulkan untuk kemudian dianalisis. Tahapan berikutnya adalah mengimplemetasikan nya kedalam bentuk system yang sederhana.

2.2 Klasifikasi

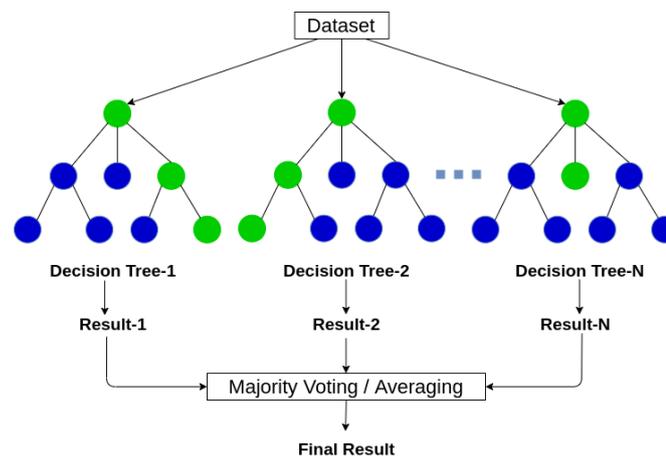
Klasifikasi menetapkan setiap catatan ke salah satukategori yang telah ditentukan sebelumnya. Set pelatihan adalah datanya yang diketahui diberi label kelas oleh algoritma pembelajaran yang digunakan untuk membangun klasifikasi model. Klasifikasi menggunakan algoritma pembelajaran untuk mengidentifikasi model di antara sekumpulan atribut dan label kelas yang akan dilatih [2]. Klasifikasi merupakan salah satu teknik dalam data mining. Klasifikasi (taksonomi) merupakan proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan objek yang digunakan. Salahsatu teknik klasifikasi yang populer digunakan adalah decision tree [3]. Proses penemuan model (atau fungsi) yang menggambarkan dan membedakankelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelasdari objek yang label kelasnya tidak diketahui. Algoritma klasifikasi yang banyak digunakan secara luas, yaitu classification trees, Random Forest, Bayesian classifiers/ Naïve Bayes classifiers dan Neural networks. Klasifikasi merupakan cara pengelompokkanbenda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya,klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun denganbantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer.Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, DecisionTree, Fuzzy dan Jaringan Saraf Tiruan [4].

2.3 CART

Algoritma *Classification And Regression Trees* (CART) adalah satu metode atau algoritma dari salah satu teknik eksplorasi data yaitu teknik pohon keputusan. Metode ini dikembangkan oleh Leo Breiman, Jerome H.Friedman, Richard A. Olshen dan Charles J. Stone sekitar tahun 1980. CART merupakan metodologi statistik non-parametrik yang dikembangkan untuk topik analisis klasifikasi, baik untuk variabel respon kategorik maupun kontinu [5]. Metode ini merupakan metode yang biasa diterapkan suatu himpunan data yang mempunyai jumlah besar, variabel yang sangat banyak dan dengan skala variabel campuran melalui prosedur pemilahan biner.

2.4 Random Forest

Random Forest merupakan metode klasifikasi yang dilakukan dengan mengembangkan metode Decision Tree berdasarkan pemilihan atribut acak pada setiap node untuk menentukan klasifikasi [6]. *Random Forest* merupakan metode pembelajaran ensemble yang digunakan untuk klasifikasi dan regresi. Dikembangkan oleh Breiman (2001), metode ini menggabungkan pendekatan pengambilan sampel bagging Breiman (1996a), dan pemilihan fitur acak, yang diperkenalkan secara independen oleh Ho (1995; 1998) dan Amit dan Geman (1997), untuk membangun kumpulan pohon keputusan dengan variasi terkendali. Dengan menggunakan bagging, setiap pohon keputusan dalam ansambel dibangun menggunakan sampel dengan penggantian dari data pelatihan [7].



Gambar 2. pohon *Random Forest*

2.5 Machine Learning

Python adalah bahasa yang ditafsirkan dan tidak memerlukan kompilasi ataupun autotools, yang dapat menghemat banyak waktu dalam pengembangan program. Penerjemah dapat digunakan secara interaktif, sehingga memudahkan untuk bereksperimen dengan fitur bahasa, membuat program satu kali pakai, atau menguji fitur selama pengembangan program dari bawah ke atas [8]. *Python* mudah digunakan, tetapi merupakan bahasa pemrograman nyata yang menyediakan lebih banyak struktur dan dukungan untuk program besar daripada *skrip shell* atau *file batch*. Di sisi lain, *Python* memiliki lebih banyak pengecekan kesalahan dibandingkan C, dan karena merupakan bahasa tingkat paling tinggi, sehingga memiliki tipe data tingkat tinggi, seperti *array* dan kamus yang fleksibel. Tipe data yang lebih umum membuatnya dapat diterapkan pada permasalahan yang jauh lebih luas daripada *Awk* atau *Perl*, namun banyak hal yang setidaknya sesederhana bahasa tersebut.

2.6 Jenis Tanah

Tanah merupakan salah satu media tanam. Tanah memiliki beberapa faktor- faktor utama untuk pertumbuhan tanaman, diantaranya adalah kelembapan dan tingkat keasaman tanah. Tingkat kelembapan tanah dan keasaman tanah dapat menentukan kesuburan tanah dan juga tanaman yang cocok ditanami pada tanah tersebut. Beberapa alat yang telah ada kebanyakan hanya dapat menentukan tingkat kelembapan tanah atau tingkat keasaman saja, tanpa memberitahukan tingkat kesuburan tanah dan juga tanaman yang cocok ditanami pada tanah tersebut [9]. Dalam pemilihan tanaman pangan berdasarkan kondisi tanah ini adalah salah satu cara untuk meningkatkan produktivitas dengan mengetahui pengetahuan tentang kesesuaian tanaman pada kondisi lahan. Dengan mengetahui faktor –faktor tentang kesesuaian jenis tanaman pangan dengan kondisi lahan maka akan membantu dalam meningkatkan produktivitas saat panen dan mengurangi masalah gagal panen [10].

3. HASIL DAN PEMBAHASAN

3.1 Analisa Hasil

Pada penelitian ini peneliti menggunakan data dari Dinas Pertanian di Kabupaten Mandailing Natal. Data yang telah didapatkan dari Dinas pertanian Kabupaten Mandailing Natal tersebut akan diolah pada aplikasi google collaboratory menggunakan bahasa pemrograman python. Jumlah keseluruhan data yang digunakan adalah 310 data.

3.2 Hasil Pengumpulan Dataset

Data yang terdiri dari 6 variabel yaitu Drainase, Nilai kelembapan, Kelembapan, Tekstur tanah, Unsur hara dan Jenis tanaman. Variabel drainase memiliki 4 data yaitu sangat baik, baik, kurang baik dan buruk. Variabel nilai kelembapan memiliki nilai dari 0-100. Kelembapan memiliki 2 data yaitu normal dan tidak normal. Variabel tekstur tanah memiliki 5 data yaitu lempung berpasir, lempung liat berpasir, liat berliat, pasir berlempung dan liat berpasir. Variabel unsur hara memiliki 3 data yaitu bagus, cukup dan buruk. Variabel jenis tanaman memiliki 4 data yaitu padi, jagung, ubi kayu dan ubi jalar.

Tabel 1. Kumpulan Dataset

Drainase	Nilai kelembapan	Kelembapan	Tekstur tanah	Unsur hara	Jenis tanaman
Sangat baik	50	Normal	Lempung berpasir	Bagus	Padi
Buruk	70	Normal	Lempung berpasir	Cukup	Ubi jalar
Kurang baik	80	Tidak normal	Lempung berpasir	Cukup	Ubi kayu
Buruk	90	Tidak normal	Lempung liat berpasir	Buruk	Jagung
Baik	88	Tidak normal	Lempung berpasir	Bagus	Padi

3.3 Analisa Hubungan Kondisi Jenis Tanaman Dan Kelembapan Terhadap Jumlah Dataset

Tabel 2. hubungan jenis tanaman dan kelembapan terhadap jumlah dataset

Jenis tanaman	Kelembapan	Jumlah data
Padi	Normal	61
Padi	Tidak normal	55
Jagung	Normal	42
Ubi jalar	Normal	40
Jagung	Tidak normal	39
Ubi kayu	Normal	30

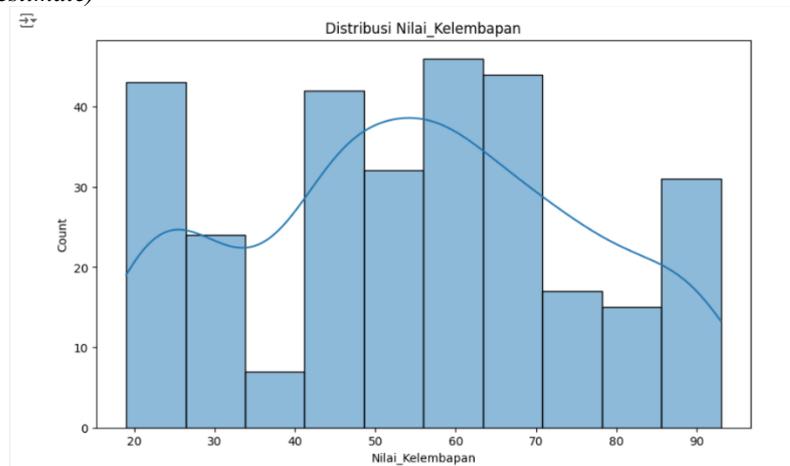
Pada padi kelembapan normal Jumlah data sampel untuk padi mencapai 61, yang merupakan angka tertinggi di antara semua kombinasi kondisi drainase dan kelembapan. Ini menunjukkan bahwa padi tumbuh paling baik atau paling sering dijumpai dalam kondisi kelembapan normal. Pada kondisi kelembapan tidak normal, jumlah data sampel padi menurun menjadi 55. Meski terjadi penurunan, jumlah ini masih relatif tinggi, menunjukkan bahwa padi tetap memiliki toleransi yang baik terhadap kelembapan yang tidak normal. Pada jagung kelembapan normal Jagung memiliki 42 sampel pada kondisi kelembapan normal, lebih rendah dibandingkan padi. Ini menunjukkan bahwa jagung mungkin kurang tahan atau kurang optimal pada kondisi kelembapan normal dibandingkan padi. Pada kelembapan tidak normal,

jumlah sampel jagung menurun sedikit menjadi 39, menunjukkan penurunan yang tidak terlalu signifikan, yang mungkin mengindikasikan jagung memiliki toleransi yang cukup baik terhadap perubahan kelembapan.

Pada ubi jalar memiliki 40 sampel pada kondisi kelembapan normal. Ini menunjukkan bahwa ubi jalar cenderung tumbuh dalam kondisi yang serupa dengan jagung, meskipun sedikit lebih rendah dalam jumlah sampel. Pada ubi kayu menunjukkan jumlah sampel yang paling rendah pada kondisi kelembapan normal dengan 30 sampel. Ini bisa mengindikasikan bahwa ubi kayu kurang optimal dalam kondisi kelembapan normal dibandingkan tanaman lain.

3.4 Analisa Distribusi Nilai Kelembapan Pada Dataset

Pada gambar dibawah ini menunjukkan histogram nilai kelembapan pada dataset yang diikuti oleh kurva KDE (*kernel density estimate*)



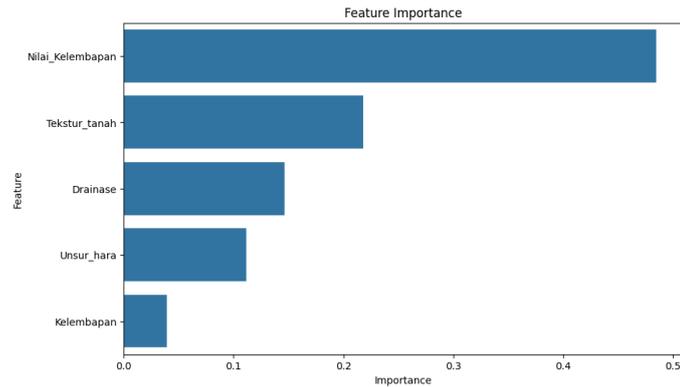
Gambar 3. grafik nilai kelembapan

Histogram ini memperlihatkan distribusi data yang cukup bervariasi, di mana terdapat beberapa puncak dan lembah yang menunjukkan konsentrasi dan kelangkaan data pada nilai-nilai kelembapan tertentu. Terlihat bahwa nilai kelembapan antara 20-30 dan 50-60 memiliki frekuensi sampel yang lebih tinggi dibandingkan dengan nilai kelembapan lainnya. Ini menunjukkan bahwa kelembapan pada rentang tersebut lebih sering terjadi dalam dataset. Gambar ini membantu memahami pola distribusi kelembapan dalam sampel data yang digunakan. Distribusi yang tidak merata dengan beberapa puncak ini bisa menunjukkan adanya pengelompokan data pada nilai kelembapan tertentu, yang mungkin terkait dengan kondisi lingkungan atau jenis tanaman tertentu.

Dengan melihat distribusi ini, kita bisa mengidentifikasi kecenderungan nilai kelembapan yang lebih umum terjadi dalam dataset dan menggunakannya untuk analisis lebih lanjut terkait pengaruh kelembapan terhadap hasil atau kondisi tanaman yang ada di dalam dataset tersebut.

3.5 Variable Importance (variabel penting)

Salah satu output dalam *Random Forest* adalah *variable importance* (variabel terpenting). Variable importance ini didapat dari rata-rata penurunan indeks gini yang diperoleh selama pembentukan *forest*. Grafik ini berguna untuk memahami mana fitur yang paling berpengaruh dan mana yang kurang penting dalam model analitik atau prediktif yang sedang dianalisis.



Gambar 4. grafik variable importance

Gambar diatas adalah grafik batang horizontal yang menunjukkan "Feature Importance" atau pentingnya fitur dalam suatu model atau analisis. Nilai_Kelembapan memiliki kontribusi terbesar dengan nilai mendekati 0.5. Ini berarti bahwa fitur ini memiliki pengaruh paling signifikan terhadap hasil model atau analisis. Tekstur_tanah adalah fitur penting kedua, dengan kontribusi sekitar 0.3. Drainase berada di urutan ketiga, dengan pentingnya sekitar 0.2. Unsur_hara memiliki kontribusi yang sedikit lebih kecil, berada di urutan keempat. Kelembapan memiliki nilai terendah dalam hal pentingnya fitur, menunjukkan pengaruh yang lebih kecil terhadap hasil model.

3.6 Classification Report

Tabel 3. Classification report

Class	Precision	recall	F1-score	support
0	0,39	0,35	0,37	20
1	0,46	0,29	0,35	21
2	0,12	0,15	0,13	13
3	0,15	0,29	0,20	7
Accuracy	0.2786885245901639			

Pada data diatas adalah hasil dari *accuracy*, *recall*, *f-1 score* dan *support* dari data jenis tanaman. Hasil akurasi yang di dapatkan 0,27% dengan arti hasil akurasi yang sangat rendah

3.7 Implementasi Hasil

Pada sub-bab ini, peneliti akan melanjutkan pengolahan data menggunakan *Google Colab*, proses implementasi sistem untuk klasifikasi jenis tanah untuk tanaman pangan akan diolah dengan efisien dan efektif

```

jenis_tanaman_counts = data['Jenis_tanaman'].value_counts()
plt.figure(figsize=(10, 6))
jenis_tanaman_counts.plot(kind='bar')
plt.title('Frekuensi Jenis Tanaman')
plt.ylabel('Jumlah')
plt.xticks(rotation=45)
plt.show()

```

Gambar 5. Frekuensi jenis tanaman

Tujuan dari tahapan kode diatas adalah membuat grafik jenis tanaman dari *dataset* menggunakan pustaka *Matploilib* (plt). Kode ini akan menghasilkan sebuah grafik batang yang menunjukkan frekuensi atau jumlah dari setiap jenis tanaman dalam dataset. Grafik tersebut akan memiliki ukuran 10x6 inci, judul "Frekuensi Jenis Tanaman", label sumbu y "Jumlah", dan label sumbu x yang diputar sebesar 45 derajat untuk kemudahan membaca.

```
# Split dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Initialize RandomForestClassifier using CART (default criterion is 'gini')
rf = RandomForestClassifier(n_estimators=10, max_depth=5, random_state=42, criterion='gini')

# Train the model
rf.fit(X_train, y_train)

# Make predictions
predictions = rf.predict(X_test)
```

Gambar 6. Split data

RandomForestClassifier dengan Kriteria Gini Pertama, dataset dibagi menjadi dua bagian, yaitu data pelatihan (training set) dan data pengujian (testing set) menggunakan fungsi `train_test_split`. Sebanyak 80% data digunakan untuk pelatihan, sementara 20% sisanya digunakan untuk pengujian, dengan parameter `random_state=42` untuk memastikan bahwa pembagian dataset ini konsisten setiap kali dijalankan

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import GridSearchCV

rf_model = RandomForestClassifier()

param_grid = {
    'n_estimators': [10, 50, 100],
    'max_depth': [None, 10, 20],
    'min_samples_split': [2, 5]
}

grid_search = GridSearchCV(estimator=rf_model, param_grid=param_grid, cv=3, n_jobs=-1, verbose=2)

grid_search.fit(X_train_smote, y_train_smote)

print("Best Parameters:", grid_search.best_params_)
print("Best Score:", grid_search.best_score_)
```

Gambar 7. Grid Search

GridSearchCV dari `sklearn.model_selection` untuk melakukan pencarian parameter grid yang optimal. Setelah itu, model *Random Forest* diinisialisasi tanpa parameter tambahan, sehingga menggunakan pengaturan default

3.8 Implementasi Sistem

Implementasi sistem proses untuk mengembangkan aplikasi yang mengimplementasikan model yang sudah dilatih sebelumnya. Dimana aplikasi ini juga sebagai data visual agar mempermudah *user* melihat hasil klasifikasi jenis tanah yang sesuai terhadap tanaman pangan. Sistem dibangun dengan bahasa python dan *framework* streamlit untuk mempermudah dalam penampilan data. Sistem yang telah dirancang untuk dibangun pada penelitian ini diimplementasikan sebagai berikut:

1. Halaman home

Gambar 8. Tampilan halaman home

Di halaman home terdapat *sidebar* tempat menu menu yaitu drainase, nilaikelembapan, kelembapan, tekstur tanah dan unsur hara.

2. Tampilan pilihan data drainase



Masukkan Karakteristik Tanah

Drainase

sangat baik

sangat baik

buruk

kurang baik

baik

normal

Tekstur tanah

lempung berpasir

Unsur hara

bagus

Prediksi Jenis Tanaman yang Cocok untuk Jenis Tanah

Prediksi

Gambar 9. Tampilan pilihan data drainase

4. DISKUSI

Penelitian ini berhasil mengembangkan model klasifikasi jenis tanah yang sesuai untuk tanaman pangan di Kabupaten Mandailing Natal menggunakan algoritma Random Forest. Hasil yang diperoleh menunjukkan bahwa variabel seperti nilai kelembapan, tekstur tanah, drainase, dan unsur hara berperan penting dalam menentukan kesesuaian tanah untuk berbagai jenis tanaman pangan. Nilai kelembapan terbukti menjadi variabel yang paling berpengaruh, diikuti oleh tekstur tanah dan drainase, sebagaimana yang ditunjukkan oleh grafik *variable importance*. Temuan ini konsisten dengan penelitian sebelumnya yang menyatakan bahwa kelembapan tanah merupakan faktor krusial dalam produktivitas tanaman pangan, terutama padi, jagung, ubi kayu, dan ubi jalar.

Ketika dibandingkan dengan studi serupa, model Random Forest menunjukkan tingkat akurasi yang lebih rendah (0,27%) dibandingkan metode lain yang biasa digunakan dalam klasifikasi tanah, seperti Decision Tree atau Support Vector Machine. Hasil akurasi yang rendah mungkin disebabkan oleh beberapa faktor, termasuk distribusi data yang tidak seimbang atau variabel yang kurang representatif dalam dataset. Beberapa penelitian lain yang menggunakan Random Forest untuk klasifikasi tanah telah melaporkan akurasi lebih tinggi, menunjukkan bahwa peningkatan model ini mungkin dilakukan melalui penambahan data atau pemilihan variabel yang lebih baik.

Hasil klasifikasi ini juga memberikan wawasan penting bagi petani di Mandailing Natal mengenai jenis tanah yang paling cocok untuk tanaman mereka. Misalnya, kelembapan normal terbukti mendukung pertumbuhan padi dengan jumlah sampel tertinggi, sementara ubi kayu kurang optimal dalam kondisi tersebut. Penemuan ini dapat menjadi panduan praktis bagi petani untuk memilih jenis tanah yang tepat berdasarkan kondisi kelembapan, drainase, dan tekstur tanah di lahan mereka, dengan harapan dapat meningkatkan hasil panen dan mengurangi risiko gagal panen.

5. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi untuk menentukan jenis tanah yang sesuai bagi tanaman pangan di Kabupaten Mandailing Natal menggunakan algoritma *Random Forest*. Model ini menganalisis variabel-variabel penting seperti nilai kelembapan, tekstur tanah, drainase, dan unsur hara, dan menemukan bahwa nilai kelembapan memiliki pengaruh terbesar terhadap kesesuaian tanah untuk tanaman pangan.

Meskipun demikian, tingkat akurasi yang diperoleh cukup rendah, yaitu sebesar 27%. Hasil akurasi yang rendah ini kemungkinan disebabkan oleh beberapa faktor, seperti distribusi data yang tidak seimbang atau kurang representatifnya variabel-variabel dalam dataset. Namun, penelitian ini tetap memberikan wawasan berharga mengenai jenis tanah yang paling cocok untuk beberapa tanaman pangan, khususnya padi, yang menunjukkan hasil terbaik dalam kondisi kelembapan normal.

Temuan ini dapat menjadi panduan praktis bagi petani di Mandailing Natal untuk memilih jenis tanah yang tepat sesuai dengan kondisi kelembapan, drainase, dan tekstur tanah. Dengan demikian, diharapkan hasil panen dapat meningkat dan risiko gagal panen dapat diminimalkan. Implementasi sistem berbasis *machine learning* ini juga menunjukkan potensi untuk dikembangkan lebih lanjut guna membantu petani dalam pengambilan keputusan terkait lahan pertanian mereka.

6. UCAPAN TERIMA KASIH

Terimakasih kepada dosen pembimbing serta orang tua dan keluarga yang membantu dan mendukung proses penelitian ini sehingga dapat berjalan lebih baik. Serta terimakasih kepada Universitas Malikussaleh yang telah memberikan fasilitas untuk menunjang penelitian ini.

7. DAFTAR PUSTAKA

- [1] Adhitya, R. R., Witanti, W., Yuniarti, R., Jenderal, U., & Yani, A. (2023). Perbandingan metode *CART* dan *Naive Bayes* untuk klasifikasi 9(2), 307–318.
- [2] Arifin, M., Putri, N. D., Sandrawati, A., & Harryanto, R. (2019). Pengaruh Posisi Lereng terhadap Sifat Fisika dan Kimia Tanah pada Inceptisols di Jatinangor. *SoilREns*, 16(2), 37–44. <https://doi.org/10.24198/soilrens.v16i2.20858>
- [3] Hidayat, H., Sunyoto, A., & Al Fatta, H. (2023). Klasifikasi Penyakit Jantung Menggunakan *Random Forest* Clasifier. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 7(1), 31–40. <https://doi.org/10.47970/siskom-kb.v7i1.464>
- [4] Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- [5] Subarkah, P. (2020). *Penerapan Algoritma Klasifikasi Classification And Regression Trees (CART) untuk Diagnosis Penyakit Diabetes Retinopathy*. 19(2), 294–301.
- [6] Sugiono, M. (2019). *Instal: Jurnal Komputer*. 15, 113–119. <https://scholar.archive.org/work/te4arwrevjhhdg47urjqwoha2y/access/wayback/https://journalinstal.cattleyadf.org/index.php/Instal/article/download/8/8>