

**LAPORAN PROYEK AKHIR  
MATA KULIAH  
KOM1327 - KECERDASAN BUATAN**

**PREDIKSI KEBUTUHAN IRIGASI PADA LAHAN PERTANIAN**

G6401221007	Rizkika Deviyanti
G6401221059	Muhammad Bagir Shahab
G6401221071	Muhammad Zaky Ghoetti Ananda
G6401221089	Aleeka Kiana Nakeisha Susanto



**DEPARTEMEN ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2024**

## **Latar Belakang**

Air merupakan salah satu sumber daya alam yang sangat penting dalam sektor pertanian. Namun, seiring dengan meningkatnya populasi dan perubahan iklim, ketersediaan air untuk irigasi menjadi semakin terbatas. Tantangan ini diperparah oleh pola cuaca yang tidak menentu, seperti curah hujan yang tidak stabil dan suhu yang ekstrem, yang memengaruhi produktivitas tanaman. Pada saat yang sama, lahan pertanian yang dikelola secara konvensional sering kali menggunakan air secara tidak efisien, yang dapat menyebabkan pemborosan dan penurunan kualitas lingkungan.

Dalam Proyek ini, kami ingin memanfaatkan kecerdasan buatan (AI) yang memiliki potensi besar untuk membantu petani mengoptimalkan penggunaan air. Dengan memanfaatkan data cuaca, kelembapan atau jenis tanah, jenis tanaman, dan faktor lainnya, AI dapat memberikan prediksi kebutuhan irigasi yang lebih akurat dan berbasis data. Hal ini tidak hanya membantu petani dalam mengelola sumber daya air secara efisien, tetapi juga mendukung praktik pertanian yang lebih berkelanjutan.

## **Rumusan Masalah**

1. Bagaimana cara memanfaatkan data cuaca, kelembapan tanah, dan jenis tanaman untuk memprediksi kebutuhan irigasi secara akurat?
2. Bagaimana model AI dapat diintegrasikan ke dalam sistem irigasi untuk mengoptimalkan penggunaan air di lahan pertanian?
3. Bagaimana model ini dapat membantu petani menghadapi tantangan perubahan iklim dan keterbatasan sumber daya air?
4. Bagaimana model ini dapat meningkatkan efisiensi penggunaan air tanpa mengorbankan produktivitas tanaman?

## **Tujuan**

Tujuan utama dari proyek ini adalah membangun model AI untuk memprediksi kebutuhan irigasi pada lahan pertanian dengan mempertimbangkan berbagai faktor penting seperti data cuaca seperti, curah hujan dan suhu, jenis tanah, serta jenis tanaman yang dibudidayakan. Model ini dirancang untuk memberikan prediksi yang akurat mengenai jumlah air yang diperlukan untuk irigasi, sehingga petani dapat mengatur irigasi secara efisien sesuai dengan kebutuhan spesifik tanaman dan kondisi lingkungan.

## **Manfaat**

Manfaat dari proyek ini dengan model AI yang digunakan adalah penggunaan air dapat dioptimalkan, mengurangi pemborosan, dan mendukung pengelolaan sumber daya yang lebih berkelanjutan. Selain itu, implementasi model ini juga dapat meningkatkan hasil panen dan produktivitas pertanian, sekaligus membantu petani dalam menghadapi tantangan perubahan iklim dan keterbatasan sumber daya air.

## **Tinjauan Pustaka**

Berdasarkan tinjauan pustaka, penelitian terdahulu menunjukkan bahwa teknologi berbasis kecerdasan buatan (AI) memiliki potensi yang signifikan dalam mendukung efisiensi irigasi pada sektor pertanian. Irigasi merupakan kegiatan yang melibatkan usaha untuk menyediakan air bagi berbagai kebutuhan pertanian, seperti sawah, ladang, dan perkebunan, yang mencakup pembangunan sarana dan prasarana guna mendistribusikan air secara teratur serta membuang kelebihan air yang tidak diperlukan untuk mencapai tujuan agronomis tertentu (Sudjarwadi, 1979).

Integrasi data real-time dengan model berbasis AI terbukti mampu meningkatkan efisiensi penggunaan air. Selain itu, data historis cuaca memiliki peran penting dalam pengembangan model prediksi irigasi yang adaptif terhadap anomali iklim. Salah satu implementasi yang relevan adalah penerapan sistem irigasi pintar di wilayah lahan kering, di mana AI digunakan untuk menentukan jadwal irigasi yang optimal berdasarkan kebutuhan spesifik tanaman dan karakteristik tanah. Penerapan model ini tidak hanya berkontribusi pada pengurangan pemborosan air, tetapi juga secara signifikan meningkatkan produktivitas hasil panen. Penggabungan teknologi AI dan analisis data lingkungan memberikan peluang strategis dalam mengoptimalkan penggunaan sumber daya air secara berkelanjutan sekaligus mendukung adaptasi sektor pertanian terhadap tantangan perubahan iklim.

Beberapa penelitian mengenai sistem irigasi telah dilakukan, salah satunya pada tahun 2017 yang mengembangkan sistem irigasi berbasis kecerdasan buatan. Sistem tersebut menggunakan metode Fuzzy-NN untuk mengontrol pergerakan servo berdasarkan data yang diperoleh dari sensor kelembaban tanah. Penelitian ini juga memperhatikan aspek teknologi komunikasi, dengan mengembangkan dua jenis teknologi utama. Pertama, teknologi komunikasi menggunakan 3G UART Transponder yang diintegrasikan dengan mikrokontroler untuk menghubungkan sistem aplikasi

dengan mikrokontroler melalui jaringan internet. Kedua, teknologi komunikasi berbasis RFM96 LoRa Transponder, yang merupakan teknologi saluran tunggal (single channel). Oleh karena itu, diperlukan dua perangkat RFM96 LoRa yang masing-masing bertugas untuk mengirimkan dan menerima data (Kontogiannis et al., 2017).

Teknologi Internet of Things (IoT) merupakan era baru dalam dunia internet yang dapat digambarkan dengan menghubungkan peralatan elektronik dengan jaringan komputer untuk berinteraksi dengan embedded sistem (Adriantantri & Dedy irawan, 2019). Teknologi AI telah menjadi alat yang efektif dalam mendukung pertanian modern, terutama dalam mengoptimalkan penggunaan sumber daya seperti air dan pupuk. Metode pengembangan berbasis AI menggunakan Prototyping, dengan pendekatan ini dapat mengetahui dengan baik kebutuhan pengguna dan analisis hasil pengembangan sistem dengan cepat. Selain itu, sistem irigasi otomatis dapat diimplementasikan untuk memantau kondisi tanah secara real-time dan mengatur aliran air secara tepat.

## **Metodologi**

Penelitian ini bertujuan untuk memprediksi kebutuhan irigasi pada lahan pertanian dengan mempertimbangkan berbagai faktor penting seperti data cuaca seperti, curah hujan dan suhu, jenis tanah, serta jenis tanaman yang dibudidayakan dengan mengambil *sample dataset Smart Agriculture Dataset* pada Kaggle yang mencakup informasi terkait parameter agrikultur, seperti jenis tanaman, luas lahan, kadar kelembapan tanah, curah hujan, suhu, dan hasil panen. Penelitian dilakukan melalui tahapan yang sistematis, yang dirancang untuk menghasilkan model prediksi yang akurat.

Metodologi penelitian dimulai dengan pengumpulan data yang mencakup berbagai parameter agrikultur, seperti jenis tanaman, luas lahan, kadar kelembapan tanah, curah hujan, suhu, dan hasil panen. Dataset yang diperoleh kemudian melalui proses pra-pemrosesan, termasuk pembersihan data untuk menghapus nilai yang hilang atau duplikat, normalisasi data numerik, transformasi data kategorik menggunakan encoding, dan pembagian dataset menjadi data pelatihan (80%) dan data pengujian (20%). Algoritma Decision Tree. Decision Tree membangun model prediksi berdasarkan pembagian dataset secara iteratif. Metode ini dipilih karena kemampuannya dalam menangani data numerik dan kategorik, serta memberikan interpretasi yang mudah dipahami melalui struktur pohon keputusan. Setelah pelatihan model dilakukan pada data pelatihan, performa model dievaluasi menggunakan data pengujian. Model terbaik

kemudian diimplementasikan untuk membangun sistem prediksi irigasi yang didukung oleh visualisasi data berupa diagram pohon keputusan, grafik pentingnya fitur, dan perbandingan hasil prediksi dengan data aktual.

Dalam pembangunan model decision tree, langkah pertama yang dilakukan adalah pemilihan fitur sebagai root. Fitur yang dipilih untuk root adalah fitur yang memiliki Information Gain tertinggi atau Entropy terendah, yang menunjukkan fitur tersebut paling efektif dalam memisahkan data dengan menggunakan rumus entropy dan gain sebagai berikut:

$$Entropy(S) = - \sum_{i=1}^m p_i \log_2(p_i)$$

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Proses ini kemudian diulang secara iteratif untuk memilih fitur pada parent node dan child node menggunakan kriteria yang sama, hingga kondisi penghentian tercapai, seperti semua data dalam node memiliki kelas yang sama atau kedalaman pohon mencapai batas. Pada proyek ini, output yang dihasilkan oleh decision tree adalah dua kelas, yaitu Used dan Not Used, yang ditentukan berdasarkan fitur-fitur prioritas yang membantu dalam menentukan apakah data tersebut termasuk dalam kategori Used atau Not Used.

## Pembahasan

Data yang digunakan untuk membuat model diperoleh dari website Kaggle yang berisikan data terkait jenis tanaman (*crop ID*), jenis tanah (*soil\_type*), kelembapan tanah (*MOI*), suhu (*temp*), dan penggunaan irigasi (*return*) pada tanaman. Terdapat 16.411 catatan data yang diolah untuk melatih dan menguji model yang dibuat.

Setelah data ditelusuri, ditemukan dua jenis data, yaitu data kategorik pada fitur jenis tanah dan jenis tanaman dan data numerik pada fitur kelembapan tanah, suhu, dan kebutuhan irigasi. Data kategorik tersebut ditransformasi menjadi data numerik dalam mempermudah pengolahan data.

```
d = {'Black Soil' : 0, 'Alluvial Soil' : 1, 'Sandy Soil' : 2, 'Red Soil' : 3, 'Clay Soil' : 4, 'Loam Soil' : 5, 'Chalky Soil' : 6}
df['soil_type'] = df['soil_type'].map(d)

d = {'Wheat' : 0, 'Potato' : 1, 'Carrot' : 2, 'Tomato' : 3, 'Chilli' : 4}
df['crop ID'] = df['crop ID'].map(d)

if df[['soil_type', 'crop ID']].isnull().any().any():
    print("Warning: NaN values detected in 'soil_type' or 'crop ID' after mapping.")
    print(df[['soil_type', 'crop ID']].isnull().sum())
    df = df.dropna(subset=['soil_type', 'crop ID'])
```

Setelah melakukan penelusuran pada data, terdapat beberapa data yang tidak sesuai sehingga sebelum dilakukan *sampling* terhadap data, data yang tidak sesuai dihilangkan terlebih dahulu. Dua ribu sampel diambil dengan metode *random sampling* dengan 1000 data mewakili masing-masing tanaman yang diperlakukan irigasi dan tidak. Beberapa fitur juga dapat diwakili oleh fitur lainnya sehingga dalam pengolahan data, fitur yang diambil hanya jenis tanaman, suhu, dan kelembapan tanah.

```
features = ['crop ID', 'MOI', 'temp']
dt_majority = df[df['result'] == 0]
dt_minority = df[df['result'] == 1]
sampA = dt_majority.sample(n=1000, random_state=42)
sampB = dt_minority.sample(n=1000, random_state=42)
samp = pandas.concat([sampA, sampB])
```

Setiap fitur memiliki nilai yang beragam, sehingga perlu dilakukan normalisasi data untuk meminimalisir kesalahan dan mengurangi duplikasi data.

Kemudian data dibagi menjadi dua bagian, yaitu data train yang akan digunakan untuk membuat dan melatih model, dan data test untuk menguji akurasi model yang sudah dilatih. Kedua data tersebut diambil dari sampel data dengan perbandingan 8:2.

```
scaler = MinMaxScaler()
X[features] = scaler.fit_transform(X[features])
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
```

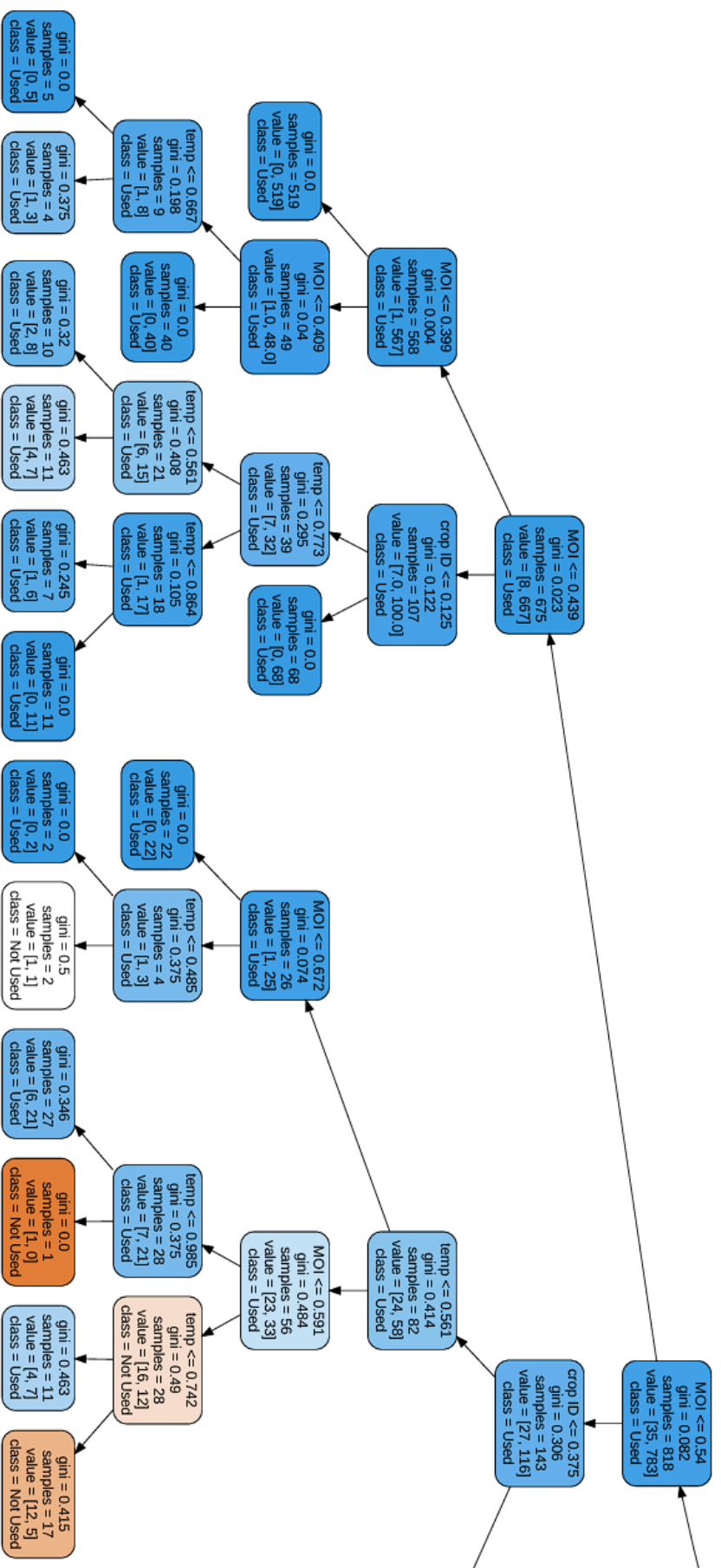
Data yang sudah dinormalisasi tersebut siap digunakan untuk melatih model yang dibuat. Model yang digunakan adalah *decision tree*.

```
dtree = DecisionTreeClassifier(max_depth=7)
dtree = dtree.fit(X_train, y_train)
plt.figure(figsize=(70, 50))
tree.plot_tree(dtree,
                feature_names=features,
                filled=False,
                rounded=True,
                label='all',
                class_names=['Not Used', "Used"],
                fontsize=14)
plt.savefig("decision_tree_plot.png")
```

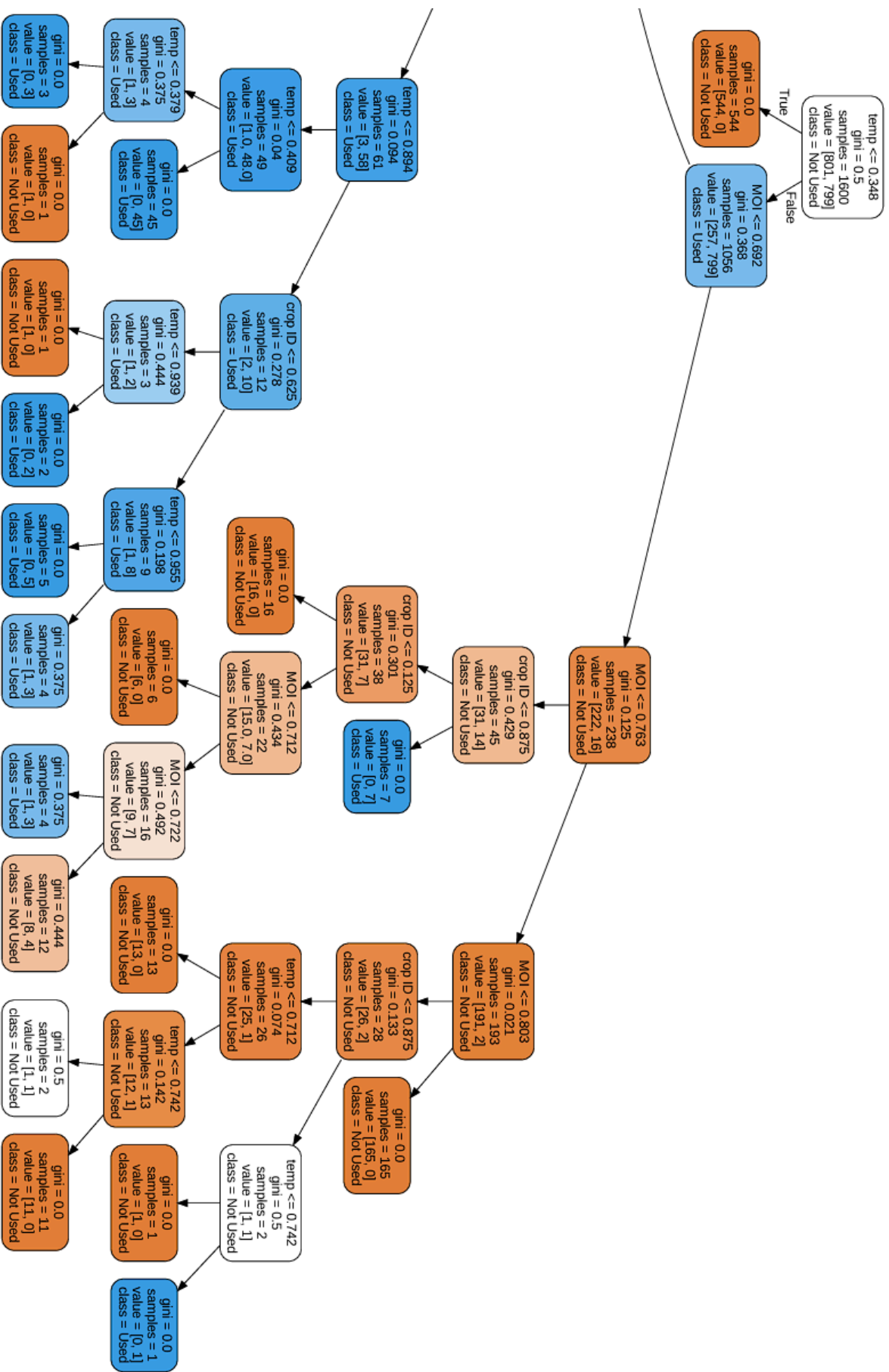
Kedalaman *tree* dari model *decision tree* yang dibentuk dibatasi dengan kedalaman 7 untuk menghindari terjadinya *overfitting*. Kemudian model tersebut diuji dengan menghitung akurasi dari hasil prediksi terhadap kedua data train dan data test. Diperoleh akurasi terhadap data train sebesar 98,0625% dan akurasi terhadap data test sebesar 96,75%.

```
y_pred = dtree.predict(X_test)
y_pred_train = dtree.predict(X_train)
print("Accuracy : {}".format(accuracy_score(y_test, y_pred)))
print("Train Accuracy : {}".format(accuracy_score(y_train,
y_pred_train)))
# Output
Accuracy : 0.9675
Train Accuracy : 0.980625
```

Berikut model *decision tree* yang dihasilkan dalam bentuk tree :







Lahan dengan nilai suhu hasil normalisasi dibawah 0,348 dapat dengan mudah untuk menetapkan bahwa lahan tersebut tidak membutuhkan irigasi. Jika tidak, maka kita perlu mempertimbangkan nilai fitur lainnya untuk menentukan apakah lahan tersebut memerlukan irigasi atau tidak, seperti lahan dengan nilai MOI yang lebih kecil atau sama dengan 0,54 pasti membutuhkan irigasi, dan sebaliknya nilai MOI yang lebih besar dari 0,54 cenderung tidak memerlukan irigasi, dan pertimbangan nilai-nilai lainnya.

### **Kesimpulan**

Kebutuhan irigasi pada lahan pertanian dapat diprediksi melalui model *decision tree* yang sudah dilatih menggunakan data yang ada dengan tingkat akurasi 96,75%. Prediksi dapat dilakukan hanya dengan mendapatkan nilai suhu, kelembapan tanah, dan jenis tanaman yang ditanam pada lahan.

Lahan dengan nilai suhu hasil normalisasi dibawah 0,348 dapat dengan mudah untuk menetapkan bahwa lahan tersebut tidak membutuhkan irigasi. Jika tidak, maka kita perlu mempertimbangkan nilai fitur lainnya untuk menentukan apakah lahan tersebut memerlukan irigasi atau tidak, seperti lahan dengan nilai MOI yang lebih kecil atau sama dengan 0.54 pasti membutuhkan irigasi, dan nilai-nilai lainnya.

## Daftar Pustaka

- Adriantantri, E., & Irawan, J. D. (2019). Implementasi IoT pada Remote Monitoring dan Controlling Green House. *Jurnal Mnemonic*, 1(1), 56–60.
- Chaitanya, G. Smart Agriculture Dataset. [Smart Agriculture Dataset](#) (Diakses 12 Desember 2024)
- Kontogiannis, S., Kokkonis, G., Ellinidou, S., & Valsamidis, S. (2017). Proposed Fuzzy-NN Algorithm with LoRa Communication Protocol for Clustered Irrigation Systems. *Future Internet*, 9(4), 78. <https://doi.org/10.3390/fi9040078>.
- Sudjarwadi. (1979). Pengantar Teknik Irigasi. Fakultas Teknik Universitas Gadjah Mada, Yogyakarta.