

Model Decision Tree Forecasting Berbasis DHT22 pada Smart Hydroponic Microgreen

DHT22 Based Decision Tree Forecasting Model on Smart Hydroponic Microgreen

Charis Fathul Hadi¹, Ratna Mustika Yasi², Andiko Prasetyo³

^{1,2,3}*Teknik Elektro, Fakultas Teknik, Universitas PGRI Banyuwangi
Jl. Ikan Tongkol No.22, Kertosari, Banyuwangi, Jawa Timur*

*.¹Penulis korespondensi: charis@unibabwi.ac.id

²ratna.mustika@unibabwi.ac.id, ³andikoprasetyo0@gmail.com

Received on 20-07-2023, accepted on 27-09-2023, published on 31-01-2024

Abstrak

Sensor DHT22 diharuskan aktif selama 4 hari penuh, jika dilihat dari waktu rata-rata budidaya *microgreen* kacang hijau menggunakan *smart hydroponic*. Penggunaan sensor DHT22 dalam jangka waktu tertentu dapat mempengaruhi kestabilan jangka panjang dalam hal pembacaan suhu. Jika pembacaan DHT22 mulai tidak akurat, *smart system* tidak mampu bekerja maksimal. Peneliti mengusulkan sebuah sistem untuk menggantikan peran sensor tanpa mengurangi atau mengganggu kinerja dari *smart system* untuk mengurangi ketergantungan komponen elektronika seperti sensor suhu. Seiring kemajuan teknologi, terdapat salah satu *model machine learning* yang dapat diterapkan untuk menggantikan peran sensor yaitu prediksi suhu melalui *forecasting*. Algoritma *forecasting* yang digunakan adalah *decision tree*. Algoritma ini dipilih karena mampu memprediksi data hanya dengan satu jenis input data berupa waktu dengan proses pelatihan yang cukup cepat. Data latih dihasilkan dari perekaman data selama 4 hari pada *smart hydroponic microgreen* kacang hijau. *Model* akan dibuat menggunakan *grid search cross validation* dan *feature scaling*. Hasil penelitian menunjukkan *channel 80* layak dipilih menjadi *model* prediksi. Prediksi suhu *model decision tree forecasting* menghasilkan nilai *R-squared* sebesar 0,870955 dan *mean square error* (MSE) sebesar 0,074171. Kedua nilai tersebut menunjukkan bahwa *model* cukup kuat dalam memprediksi suhu dan layak untuk diterapkan dalam memantau suhu *smart hydroponic microgreen* kacang hijau.

Kata kunci: *Decision Tree, Forecasting, Microgreen Kacang Hijau, Smart Hydroponic.*

Abstract

The DHT22 sensor is required to be active for 4 full days when viewed from the average time for cultivating mung bean microgreens using smart hydroponics. The use of the DHT22 sensor for a certain period can affect long-term stability in terms of temperature readings. If the DHT22 reading becomes inaccurate, the smart system is unable to function optimally. To reduce dependence on electronic components such as temperature sensors, researchers propose a system to replace the role of sensors without reducing or disrupting the performance of the smart system. As technology advances, a machine learning model can be used to replace the role of sensors, namely temperature prediction through forecasting. The forecasting algorithm used is a decision tree. This algorithm is chosen because it is able to predict data with only one type of data input in the form of time with a fairly fast training process. The training data was generated from data recording for 4 days on smart hydroponic microgreen mung bean. The model was created using grid search cross-validation and feature scaling. The results of this research show that channel 80 deserves to be chosen as a prediction model. The temperature prediction of the decision tree forecasting model produces an R-squared value of 0.870955 and a mean square error (MSE) of 0.074171. Both of these values indicate that the model is strong enough in predicting temperature and is feasible to be applied in monitoring the temperature of mung bean microgreen smart hydroponics.

Keywords: Decision Tree, Forecasting, Mung Beans Microgreen, Smart Hydroponic

I. PENDAHULUAN

Smart hydroponic yang marak diteliti saat ini, fokus pada sistem kendali dan *monitoring*. Seperti perancangan dan pembuatan sistem pengendalian pengering cabai berbasis arduino uno menggunakan sensor DHT22 [1]. Merancang dan membangun sistem penyejuk lingkungan *smart growth box* untuk *microgreens* [2]. Penerapan *smart hydroponic system* pakan ternak jagung menggunakan ESP32 [3]. Penelitian - penelitian tersebut menegaskan bahwa sensor DHT22 dapat diterapkan pada bidang agrikultur. Namun berdasarkan *datasheet*, DHT22 merupakan sensor ruangan yang tidak dirancang khusus untuk agrikultur, khususnya hidroponik. Sehingga perlu dipertanyakan mengenai kestabilan sensor DHT22 jika diterapkan pada jangka waktu tertentu.

Pada penelitian ini *smart hydroponic* akan diterapkan untuk budidaya *microgreen* kacang hijau. Sensor DHT22 diharuskan aktif selama 4 hari penuh, jika dilihat dari waktu rata-rata budidaya *microgreen* kacang hijau menggunakan *smart hydroponic*. Penggunaan sensor DHT22 dalam jangka waktu tertentu dapat mempengaruhi *long-term stability* (kestabilan jangka panjang) [4]. Jika pembacaan DHT22 mulai tidak stabil atau tingkat *error* lebih dari 3% [5], hal ini akan mempengaruhi *smart system* sehingga menghambat pertumbuhan *hydroponic microgreen*. Untuk mengurangi ketergantungan komponen elektronika seperti sensor suhu, peneliti mengusulkan sebuah sistem untuk menggantikan peran sensor tanpa mengurangi atau mengganggu kinerja dari *smart system*. Sebuah sistem yang mampu meramalkan kondisi suhu lingkungan dapat menjadi solusi untuk masalah ini [6].

Model machine learning yang dapat diterapkan untuk meramalkan kondisi suhu lingkungan yaitu prediksi suhu melalui *forecasting*. *Forecasting* merupakan jenis mekanisme prediksi yang lebih banyak diimplementasikan dalam dunia nyata dibandingkan dengan aplikasi data *mining* lainnya [7]. *Time series forecasting* ini mengacu pada proses memprediksi masa depan dengan melihat perilaku data masa lalu. Data berupa setiap data yang menyertakan komponen waktu, seperti suhu pada setiap hari atau harga saham setiap jam [8]. Algoritma ini memiliki keunggulan mampu meramalkan data masa depan hanya dengan input data waktu. Berbeda dengan algoritma data *mining* lainnya yang membutuhkan beberapa jenis input data. Salah satu algoritma *forecasting* yang telah banyak diterapkan untuk memecahkan berbagai permasalahan di bidang kesehatan, industri, *agriculture*, dan lain-lain adalah *decision tree*. *Decision tree* merupakan metode untuk klasifikasi dan regresi menggunakan struktur hirarki, yang berguna untuk menemukan prediksi data dan menemukan keterkaitan tersembunyi dari sejumlah variabel input pada variabel target, *decision tree* umumnya mengikuti pendekatan struktur *flowchart* [9]. Metode ini mudah dipahami karena mengikuti proses yang sama dengan cara manusia saat membuat keputusan dalam kehidupan nyata [10]. Selain itu, metode ini memiliki waktu pelatihan yang cukup singkat [11]. Pada penelitian terdahulu, konsep *forecasting* telah diterapkan pada penelitian mengenai sebuah sistem yang dapat meramalkan kondisi suhu air dapat menjadi solusi untuk masalah perubahan suhu air yang ekstrem, hasilnya dapat disimpulkan bahwa *model decision tree forecasting* dengan nilai *mean squared error* (MSE) sebesar 0,01211 dan nilai *R-Squared* 0,80920 [12]. Hal ini menunjukkan algoritma *forecasting* mampu meramalkan kondisi suhu air pada *aquaponic*. Oleh sebab itu, algoritma *decision tree forecasting* akan diujikan pada *smart hydroponic microgreen* kacang hijau. Dan diharapkan mampu juga untuk meramalkan kondisi suhu hidroponik dengan akurat.

Berdasarkan permasalahan dan data-data yang disebutkan di atas, penulis bermaksud untuk meneliti performa prediksi *model decision tree forecasting* pada *smart hydroponic microgreen* kacang hijau. Sistem ini akan diaplikasikan selama 4 hari guna menjadi data latih bagi *model* dan akan dianalisis performanya. Selanjutnya digunakan dua teknik pengukuran yaitu *mean squared error* (MSE) dan *R-squared* yang biasa digunakan untuk mengukur kinerja suatu metode regresi [12] [13].

II. KAJIAN PUSTAKA

Tahap awal adalah melakukan studi literatur berupa mempelajari dan mencari sumber referensi tentang jurnal ilmiah penelitian. Konsep dikembangkan dengan cara merekam data suhu sensor DHT22 selama 4 hari dan membuat *model* prediksi berdasarkan data tersebut. Hasilnya akan ditelaah lebih lanjut menggunakan beberapa teknik pengukuran. Beberapa komponen yang diperlukan dalam penelitian ini meliputi:

A. Smart Hydroponic Microgreen Kacang Hijau

Smart hydroponic merupakan sistem hidroponik dengan *monitoring* secara *realtime* berbasis *Internet of Things* untuk memantau suhu, kelembapan dan penyiraman otomatis. *Smart Hidroponic* menggunakan

ESP32 sebagai pemroses data yang terkoneksi dengan internet. *Microgreen* kacang hijau adalah tumbuhan kacang hijau muda yang dapat dikonsumsi dengan tekstur lunak. Pemilihan sampel penelitian berupa *microgreen* kacang hijau karena biji kacang hijau mudah tumbuh di berbagai macam kondisi lingkungan, rentang waktu pertumbuhan yang cukup cepat, memiliki harga yang cukup murah dan mudah didapatkan di pasar tanpa harus membeli melalui *online shop*.



Gambar 1. *Hydroponic Microgreen*

Pemanenan *microgreen* dapat dipanen dengan memotong tepat di atas permukaan medium pertumbuhannya dengan panjang sekitar 3-9 cm tanpa akar [14]. Suhu dan kelembaban ideal untuk *microgreen* kacang hijau adalah 23-30°C dan 60%-70% RH dengan frekuensi penyiraman 3 kali sehari [15]. Sistem ini juga mampu meningkatkan pertumbuhan *microgreens* dengan hasil *microgreens* tumbuh lebih merata, daun lebih lebar, batang lebih tinggi, dan hasil panen meningkat [16].

B. Sensor Suhu DHT22

Sensor DHT22 adalah sensor yang berfungsi mengukur suhu dan kelembaban. Sensor ini biasa digunakan untuk mengukur suhu ruangan dan lingkungan. Sensor ini dapat mendeteksi suhu dari -40°C hingga 125°C dengan *error* suhu $\pm 0,5^\circ\text{C}$ [17]. Dan untuk kelembaban, tingkat kesalahan sekitar $\pm 2\%$ hingga $\pm 5\%$ [17]. DHT22 adalah sensor digital yang sudah dilengkapi dengan konverter *analog to digital* bawaan [18]. Tegangan pengoperasian modul ini merupakan 3-5V [19].



Gambar 2. Sensor DHT22

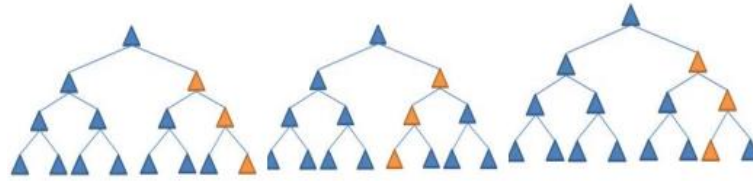
Pin DHT22 sensor suhu dan kelembaban terdiri dari 4 pin seperti DHT11 tetapi hanya 3 pin yang digunakan yaitu VCC, GND, dan data. Sensor ini dipilih karena telah sering diterapkan pada sektor pertanian, seperti hidroponik. Akurasi pembacaan yang memenuhi standar laboratorium dan murah juga menjadi kelebihan dari sensor DHT22.

C. Decision Tree Forecasting

Forecasting merupakan proses memprediksi nilai masa yang akan datang berdasarkan data masa lalu. Disiplin ilmu ini telah banyak digunakan di berbagai bidang termasuk energi, ekonomi dan lingkungan [20]. *Decision Tree* (DT) adalah salah satu metode *machine learning*, berdasarkan pelatihan, algoritma seperti pohon yang mampu membuat keputusan yang akurat [12]. Proses pembelajaran *model* menggunakan fungsi *grid search*. *Grid search* adalah suatu cara untuk menemukan parameter terbaik yang digunakan untuk *modeling* dalam *machine learning* [21]. Fungsi ini dipilih karena dapat memilih *model* terbaik berdasarkan *hyperparameter* yang ditentukan. Selain itu, fungsi ini dilengkapi *cross validation* (CV) yaitu metode statistik untuk mengevaluasi kinerja *model/* algoritma dengan memisahkan data menjadi dua subset secara diagonal berupa data proses pembelajaran dan validasi [22].

Decision tree merupakan metode untuk regresi menggunakan struktur hirarki, yang berguna untuk menemukan prediksi data dan menemukan keterkaitan tersembunyi dari sejumlah variabel input pada

variabel target, DT umumnya mengikuti pendekatan struktur flowchart [9], dimulai dengan simpul akar yang mencakup seluruh kumpulan data pelatihan dan indikator diskriminatif, menciptakan simpul dengan dua cabang di sebagian besar (partisi biner) yang membagi kumpulan data menjadi himpunan bagian yang semakin homogen [23], keterkaitan antar variabel dari *decision tree* digambarkan sebagai berikut.



Gambar 3. Keterkaitan antar variabel *decision tree*

Standar deviasi untuk 1 atribut [24]:

$$\bar{x} = \frac{\sum x}{n} \quad (1)$$

Dengan:

\bar{x} = rata-rata

x = nilai data

n = jumlah data

$$S = \sqrt{\frac{\sum(x - \bar{x})^2}{n}} \quad (2)$$

Dengan:

S = standar deviasi

$$CV = \frac{S}{\bar{x}} \times 100\% \quad (3)$$

Dengan:

CV = koefisien varians

Standar deviasi untuk 2 atribut:

$$S(T, X) = \sum_{c \in X} P(c)S(c) \quad (4)$$

Dengan:

$P(c)$ = peluang atribut

$S(c)$ = standar deviasi atribut

Standar deviasi reduksi:

$$SDR(T, X) = S(T) - S(T, X) \quad (5)$$

Dengan:

S = standar deviasi reduksi

Dalam penelitian ini, *decision tree forecasting* akan diuji keandalannya menggunakan 2 teknik pengukuran yaitu:

R-Square / R^2 adalah kuadrat dari korelasi sampel koefisien (r) antara nilai aktual dan prediksi nilai [12]. Nilai prediksi *model* akan dibandingkan dengan nilai suhu sensor. Jika nilai *R-Squared* $< 0,3$, maka *model* dianggap lemah, jika nilai *R-Squared* = $0,3 - 0,5$, maka *model* dianggap sedang, dan jika *R-Squared* $> 0,5$, maka *model* dianggap kuat [25].

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (6)$$

Dengan:

Y_i = item dari sumbu Y

n = jumlah data

Mean squared error (MSE) adalah selisih rata-rata kuadrat selisih antara nilai prediksi dan aktual. Metode *mean squared error* biasa digunakan untuk mengecek nilai kesalahan pada *forecasting*. Jika nilai *mean squared error* mendekati nol menunjukkan hasil *forecasting* sesuai dengan data aktual dan bisa digunakan untuk perhitungan *forecasting* di periode mendatang [13].

$$MSE = \frac{1}{n} \left(\sum_{i=1}^n Y_i^2 - \hat{Y}_i^2 \right) \quad (7)$$

Dengan:

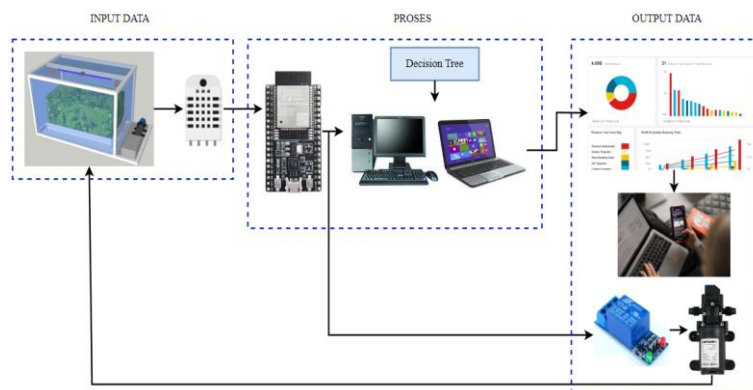
Y_i = item dari sumbu Y

n = jumlah data

III. METODE PENELITIAN

A. Perancangan Alat

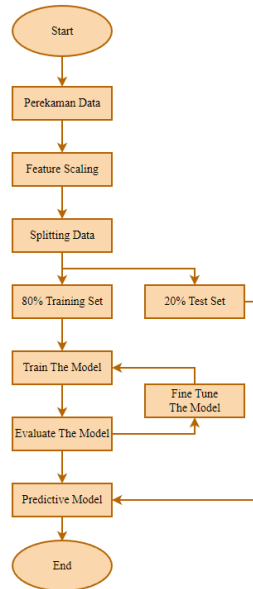
Pada perancangan alat, *smart hydroponic microgreen* terdiri dari sebuah *box* akrilik yang dilengkapi dengan *led grow light*, kipas DC, motor pompa 5V, sensor DHT22, dan rangkaian PCB ESP32. Dimulai dari menghubungkan *power supply* 12V 5A pada PCB ESP32 sebagai sumber tegangan untuk mikrokontroler ESP32 dan *output* seperti, motor pompa 5V dan *led grow light* yang akan dikontrol melalui *relay*.



Gambar 4. Skema Alat

B. Alur Pengujian Decision Tree Forecasting

Penelitian ini menerapkan lingkungan hidroponik *microgreen* dimana perangkat akhir *smart system* akan dipasang pada *box* hidroponik *microgreen*. Perangkat *smart system* akan terhubung dengan *platform internet of things* dalam bentuk server *python* berbasis *web*. Biji kacang hijau akan disemai di dalam *box* hingga 4 hari. Data selama 4 hari berupa suhu dan waktu akan direkam secara otomatis menggunakan aplikasi *VS Code*. Data suhu DHT22 akan dikirim dari ESP32 menuju *laptop* melalui komunikasi serial. Data akan berbentuk tabel dengan format *comma separated values* (csv). Data tersebut akan diolah menggunakan bantuan aplikasi Kaggle hingga didapatkan sebuah *model* berbentuk *file .pkl*. *Model* yang sudah dihasilkan akan ditanamkan pada PC untuk *monitoring* suhu prediksi. Pada berkas (*file*) tersebut terdiri dari tiga kolom untuk prediktor dan sebuah target. Kolom sebagai prediktor adalah jam, menit, dan detik. Kolom sebagai target adalah suhu. Format *dataset* tersebut dapat dilihat pada gambar 6.



Gambar 5. Diagram Alir *Decision Tree Forecasting*

Dataset dihasilkan dari perekaman suhu selama 4 hari pada *box* hidroponik *microgreen*, sehingga dapat disebut peneliti melakukan pengambilan data sendiri. Sama seperti manusia, mesin juga membutuhkan data yang mudah dipahami agar dapat memahami pola sekumpulan data. Untuk itulah data fitur diubah skalanya dengan *feature scaling* seperti pada gambar 7. Metode dipilih karena mengubah skala data dengan rentang 0 hingga 1 sehingga mesin dapat mempelajari data lebih cepat dan akurat. Data target tidak perlu dilakukan *feature scaling* karena data suhu merupakan data *output*. *Feature scaling* dengan metode *z-score* dipilih karena mampu mengambil titik data dari berbagai rentang waktu dengan rata-rata dan standar deviasi yang berbeda, lalu mengubahnya menjadi nilai dengan skala yang sama. Dengan skala yang standar tersebut, analisis dan perbandingan data dengan berbagai jenis variabel jadi lebih mudah dilakukan. Selain itu, mampu mencegah adanya atribut pada dataset yang memiliki rentang nilai jauh berbeda dari atribut lainnya

| | Nomor | Suhu | Jam | Menit | Detik |
|-------|-------|------|-----|-------|-------|
| 0 | 0 | 27.2 | 6 | 59 | 38 |
| 1 | 1 | 27.2 | 6 | 59 | 39 |
| 2 | 2 | 27.0 | 6 | 59 | 40 |
| 3 | 3 | 27.0 | 6 | 59 | 41 |
| 4 | 4 | 27.1 | 6 | 59 | 42 |
| ... | ... | ... | ... | ... | ... |
| 23030 | 23030 | 27.3 | 6 | 35 | 14 |
| 23031 | 23031 | 27.3 | 6 | 35 | 15 |
| 23032 | 23032 | 27.3 | 6 | 35 | 16 |
| 23033 | 23033 | 27.3 | 6 | 35 | 17 |
| 23034 | 23034 | 27.3 | 6 | 35 | 18 |

[251227 rows x 5 columns]
 Jumlah Dimensi data gabungan : (251227, 5)

Gambar 6. Format *Dataset*

Proses pelatihan pada *machine learning* merupakan mekanisme pembentukan *model* dari fitur untuk mendapatkan koefisien atau bobot. Pemilihan persentase dalam pembagian data/ *splitting* data pada penelitian ini didasari oleh Prinsip Pareto, dimana prinsip ini juga dikenal sebagai aturan 80/20. Ini pada dasarnya adalah teori yang menetapkan bahwa 80% dari *output* atau hasil berasal dari 20% efek atau aliansi yang tidak proporsional antara *input* dengan *output* [26]. 80% data dari fitur yang dilatih menggunakan *model decision tree*. Proses pemakaian *model* menggunakan rekaman bobot terbaik pada iterasi yang dihasilkan terhadap fitur latih.

| | 0 | 1 | 2 |
|--------|-----------|----------|-----------|
| 0 | -0.753213 | 1.719774 | 0.490982 |
| 1 | -0.753213 | 1.719774 | 0.548730 |
| 2 | -0.753213 | 1.719774 | 0.606478 |
| 3 | -0.753213 | 1.719774 | 0.664226 |
| 4 | -0.753213 | 1.719774 | 0.721974 |
| -- | -- | -- | -- |
| 251222 | -0.753213 | 0.335414 | -0.894974 |
| 251223 | -0.753213 | 0.335414 | -0.837226 |
| 251224 | -0.753213 | 0.335414 | -0.779477 |
| 251225 | -0.753213 | 0.335414 | -0.721729 |
| 251226 | -0.753213 | 0.335414 | -0.663981 |

Gambar 7. Format Dataset setelah Feature Scaling

Machine learning model adalah program komputer yang biasa digunakan untuk mengenali pola dalam data ataupun membuat sebuah prediksi. Pada pelatihan *model decision tree*, akurasi dari tiap *channel* akan dibandingkan dengan proses pelatihan dan pengujian dalam *pipeline* yang dirancang dalam *grid search*. *Hyperparameter* terdiri dari *max_depth* merupakan kedalaman maksimum pohon, *min_sample_split* merupakan jumlah sampel minimum yang diperlukan untuk membelah *internal node* (simpul dalam), *min_sample_leaf* merupakan jumlah minimum sampel yang diperlukan berada pada *leaf node* (cabang dari simpul dalam). Data akurasi tertinggi akan dianalisis dan digunakan pada pengujian berikutnya, pemakaian fitur berdasarkan skala akurasi tertinggi diharapkan dapat memperbaiki atau meningkatkan akurasi pada *model decision tree*. Setelah itu sisa data 20% digunakan untuk menguji *model* sehingga muncullah data prediksi terhadap waktu berupa suhu. Selanjutnya *R-Squared* dan *mean square error* (MSE) untuk menunjukkan performa dari penerapan *model* terhadap implementasi sebenarnya.

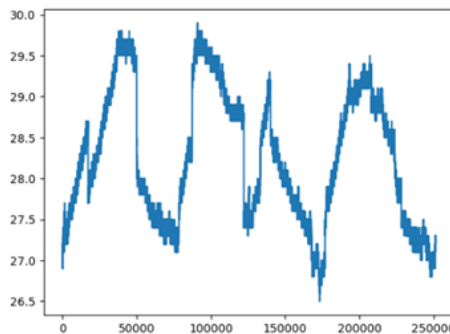
Tabel 1. *Hyperparameter Decision Tree*

| <i>Model Forecasting</i> | <i>Hyperparameter</i> |
|--------------------------|--|
| <i>Decision Tree</i> | <i>max_depth: 1,2,3,4,5</i> <i>min_samples_split: 2,3,4,5</i> <i>min_samples_leaf: 1,2,3,4,5</i> |

IV. HASIL DAN PEMBAHASAN

Penelitian ini menjelaskan tentang hasil dan pembahasan analisis *model decision tree forecasting* berbasis DHT22 pada *smart hydroponic microgreen*. Penelitian *microgreen* ini menggunakan biji kacang hijau. Penelitian dilakukan selama 4 hari untuk perekaman data *forecasting*. Data tersebut akan ditujukan sebagai data latih *model* dan akan dianalisis hasilnya. Sehingga pada penelitian ini didapatkan data berikut:

A. Dataset



Gambar 8. *Dataset*

Keterangan:

Sumbu X = Urutan data waktu

Sumbu Y = Rentang Suhu

Dataset diperoleh dari hasil perekaman data waktu dan suhu selama 4 hari pada *box smart hydroponic microgreen*. *Dataset* terdiri dari 251226 data yang terdiri dari perekaman data suhu setiap detik. data suhu meningkat dan menurun seiring berjalannya waktu. Selain faktor lingkungan, suhu meningkat karena pada

uji coba ini belum diterapkan sistem kendali untuk mengatur suhu agar stabil pada rentang tertentu. Kipas DC dan *led grow light* yang berada di dekat sensor dan alat ukur membuat suhu semakin meningkat dan suhu akan cenderung turun ketika pompa aktif pada proses penyiraman. Data cenderung meningkat pada rentang 40000, 100000, 150000, dan 200000 karena data ini merekam pada saat siang hari dan berkebalikan dengan rentang 60000, 120000, 180000, dan 250000 karena data ini merekam pada saat malam hari. Pada perekaman data, data sempat hilang beberapa kali akibat *force close* sehingga data per detik selama 4 hari yang seharusnya terkumpul sekitar 345.600 menjadi 251.226. Tidak hanya itu, sumber listrik yang mati saat kegiatan maintenance berkala dari PLN, membuat data hilang karena tidak adanya sumber daya. Penyebab hilangnya data juga disebabkan komunikasi antara ESP32 dan laptop terjeda karena mengeksekusi program penjadwalan *led grow light* dan motor pompa secara bersamaan. Hal ini disebabkan komunikasi serial merupakan metode komunikasi mikrokontroler dengan cara mengirim data per bit secara berurutan dan bergantian. Selain itu, semakin cepat dan banyaknya program membuat paket-paket bit data menjadi terdistorsi sehingga data yang dikirim atau diterima bisa mengalami *error*.

B. Model Decision Tree Forecasting

Proses *Grid search* akan menghasilkan sekitar 500 *channel* dan akan ditampilkan 5 hasil terbaik berdasarkan skor seperti tabel 2.

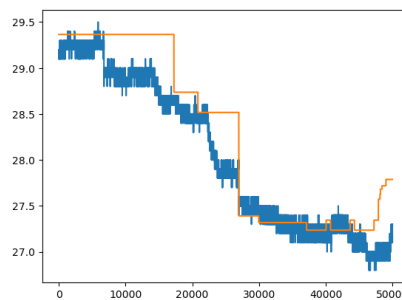
Tabel 2. Hasil Uji Channel pada Model Decision Tree

| Channel | Mean Fit Time | Std Fit Time | Mean Score Time | Std Score Time | Mean Test Score |
|---------|---------------|--------------|-----------------|----------------|-----------------|
| 80 | 0.094109 | 0.001304 | 0.002528 | 0.000124 | 0.886871 |
| 82 | 0.094958 | 0.001636 | 0.002656 | 0.000160 | 0.886871 |
| 92 | 0.095894 | 0.003955 | 0.002556 | 0.000096 | 0.886871 |
| 95 | 0.098145 | 0.002293 | 0.002705 | 0.000148 | 0.886871 |
| 83 | 0.095308 | 0.002577 | 0.002640 | 0.000356 | 0.886871 |

Berdasarkan tabel 2, terdapat 5 *channel* teratas dengan *mean test score* tertinggi. *Channel* 80 menempati posisi teratas dengan *mean test score* 0,886871. Pada urutan ke-2 hingga ke-4 nilai *mean test score* juga sama, namun *channel* 80 lebih unggul karena dari segi *mean fit time*, *std fit time*, *mean score time*. Walaupun nilai *std score time channel* 80 dibawah *channel* 92. Dengan hasil ini, *model* tersimpan akan otomatis menggunakan konfigurasi *hyperparameter channel* 80. *Channel* 80 terpilih dengan pengaturan *hyperparameter* max depth sebesar 50, min sample leaf sebesar 1, min sample split sebesar 2. Selain itu, *model channel* 80 menunjukkan bahwa *hyperparameter tuning* dengan *grid search cross validation* mampu membantu peneliti untuk mendapatkan *model* terbaik tanpa harus mengatur *hyperparameter* secara manual.

C. Pengujian Model Decision Tree Forecasting

Pengujian terakhir dilakukan menggunakan data acak sebanyak 50.000 atau 20% data awal. Setelah dilakukan *feature scaling*, data acak akan dipilah menjadi data fitur dan data target. Hasil prediksi ditunjukkan pada tabel 3 dan divisualisasikan pada gambar 6.



Gambar 9. Grafik Hasil Pengujian

Keterangan:
 Garis Biru = Suhu aktual
 Garis Jingga = Suhu prediksi
 Sumbu X = Urutan data waktu

Sumbu Y = Rentang Suhu

Tabel 3. Hasil Pengujian

| Waktu | Suhu Sensor (°C) | Suhu Prediksi (°C) | R ² | MSE |
|--------------|------------------|--------------------|----------------|----------|
| 20% data uji | 26.8-29,5 | 27,2-29,3 | 0,870955 | 0,074171 |

Berdasarkan tabel 3, didapatkan data suhu sensor dengan rentang 26.8-29,5 °C dan data suhu prediksi dengan rentang 27,2-29,3°C. Didukung juga dengan hasil $R^2 = 0.870955$. R^2 disebut juga sebagai koefisien determinasi yang menjelaskan seberapa jauh data dependen (suhu prediksi) dapat dijelaskan oleh data independen (waktu). R^2 bernilai antar 0 – 1 dengan ketentuan semakin mendekati angka satu berarti semakin baik. R^2 pada penelitian ini bernilai 0.87, berarti 87% sebaran variabel dependen (suhu prediksi) dapat dijelaskan oleh variabel independen (waktu). Sedangkan MSE sebesar 0.074171 menunjukkan rata-rata kesalahan kuadrat antara prediksi dengan suhu aktual. Semakin kecil nilai MSE, semakin baik kualitas *model* tersebut. Hasil pada gambar 9 juga berimpit khususnya pada rentang 30000 sampai dengan 40000 dikarenakan suhu yang diprediksi bertepatan pada waktu malam hari yang cenderung stabil pada nilai 27-27,5°C. Data prediksi terlihat beberapa kali tidak tepat meramalkan suhu akibat kurangnya data perekaman, salah satu penyebabnya adalah hilangnya data saat proses perekaman data. Selain itu, terdapat beberapa hasil prediksi yang cenderung lurus, dan tidak berimpit dengan data suhu sensor sehingga terdapat selisih pembacaan yang cukup besar di beberapa area. Hal ini terjadi karena dampak *overfitting* yaitu kondisi dimana nilai akurasi pada data *training* (latih) tinggi dengan nilai akurasi data *testing* (uji) rendah. Dapat dilihat dari nilai *R-squared* saat *training* turun setelah melalui proses *testing* dari 0.886871 menjadi 0.870955. Namun, berdasarkan dari rentang hasil *R-squared* saat *testing*, *model* dianggap cukup kuat dalam memprediksi suhu sehingga layak untuk diterapkan pada *smart system*.

V. KESIMPULAN

Pada penelitian ini telah dihasilkan sebuah *model* yang ditanamkan pada *PC monitoring* untuk memprediksi suhu *box* hidroponik *microgreen* kacang hijau. Berdasarkan hasil *grid search cross validation*, *channel 80* lebih unggul karena dari segi *mean fit time*, *std fit time*, *mean score time*. Dan layak dipilih menjadi *model* prediksi. Prediksi suhu *model decision tree forecasting* menghasilkan nilai *R-squared* sebesar 0,870955 dan *mean square error* (MSE) sebesar 0,074171. Nilai *R-squared* mendekati nilai optimal dikarenakan data sebaran variabel dependen (suhu prediksi) dapat dijelaskan oleh variabel independen (waktu). Sedangkan nilai MSE sangat kecil akibat kecilnya nilai rata-rata kesalahan kuadrat antara prediksi dengan suhu aktual. Pada proses penelitian, terdapat fenomena *overfitting*, dapat dilihat dari nilai *R-squared* saat *training* turun setelah melalui proses *testing* dari 0.886871 menjadi 0.870955. Nilai *R-squared* dan MSE tersebut menunjukkan bahwa *model* cukup kuat dalam memprediksi suhu *smart hydroponic microgreen* kacang hijau. Sehingga didapatkan hasil bahwa penerapan *model decision tree forecasting* mampu memprediksi suhu dataset hasil perekaman data pada *box* hidroponik *microgreen* kacang hijau. Untuk penelitian selanjutnya, *model* ini perlu diujikan pada sistem *realtime* dengan media *microgreen* kacang hijau yang ditanama selama 4 hari serta dianalisis hasil prediksi dengan alat ukur suhu terkalibrasi.

REFERENSI

- [1] R. M. Sembiring, "Perancangan dan Pembuatan Sistem Pengendalian Pengerincian Cabai Menggunakan Sensor DHT22 Berbasis Arduino Uno," 2019.
- [2] Nurussa'adah, S. N. Sari, T. Rohman, M. Alfanny, and A. Wicaksono, "Design and Build an Environmental Conditioning System on a Smart Growth Box for Microgreens," in *2022 11th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS)*, 2022, pp. 123–128.
- [3] A. Prasetyo, "Implementasi Sistem Hidroponik Cerdas Pakan Ternak Jagung Menggunakan ESP32," *Transm. J. Ilm. Tek. Elektro*, vol. 25, no. 1, 2022.
- [4] T. Liu, *DHT22 Aosong Electronics Co.*, vol. 22. 2015.
- [5] A. Amalia, H. R. Fajrin, and A. S. Wibowo, "Thermohygrometer Dengan Penyimpanan Data Untuk Monitoring Kamar Bedah," *Med. Tek. J. Tek. Elektromedik Indones*, vol. 2, no. 1, 2020.
- [6] N. Papakonstantinou, J. Savolainen, J. Koistinen, A. Aikala, and V. Vyatkin, "District Heating Temperature Control Algorithm Based on Short Term Weather Forecast and Consumption Predictions," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2016, pp. 1–8.
- [7] P. Mahawar and V. Shrivastava, "Analysing And Forecasting Rajasthan's Agriculture Productivity Using Association Rule Mining And ARIMA Modelling," in *2018 4th International Conference on Computing Communication and Automation, ICCCA 2018*, 2018, pp. 1–4, doi: 10.1109/CCAA.2018.8777574.

- [8] J. K. Sethi and M. Mittal, "Analysis Of Air Quality Using Univariate And Multivariate Time Series Models," in *Proceedings of the Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering*, 2020, pp. 823–827, doi: 10.1109/Confluence47617.2020.9058303.
- [9] K. Khalili-Damghani, F. Abdi, and S. Abolmakarem, "Hybrid Soft Computing Approach Based On Clustering, Rule Mining, And Decision Tree Analysis For Customer Segmentation Problem: Real Case Of Customer-Centric Industries," *Appl. Soft Comput.*, vol. 73, pp. 816–828, 2018.
- [10] M. RL and A. K. Mishra, "Measuring Financial Performance Of Indian Manufacturing Firms: Application Of Decision Tree Algorithms," *Meas. Bus. Excell.*, vol. 26, no. 3, pp. 288–307, 2022.
- [11] W. W. Sakti, K. Anam, S. B. Utomo, B. Marhaenanto, and S. Nahela, "Artificial Intelligence IoT Based EEG Application Using Deep Learning For Movement Classification," in *2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2021, pp. 192–196.
- [12] A. Taufiqurrahman, A. G. Putrada, and F. Dawani, "Decision Tree Regression with AdaBoost Ensemble Learning for Water Temperature Forecasting in Aquaponic Ecosystem," in *6th International Conference on Interactive Digital Media, ICIDM 2020*, 2020, no. Icidm, doi: 10.1109/ICIDM51048.2020.9339669.
- [13] A. R. Barron, "Predicted Squared Error: A Criterion For Automatic Model Selection," in *Self-organizing methods in modeling*, CRC Press, 2020, pp. 87–103.
- [14] D. A. S. Hartanti, Y. Puspiningrum, and A. I. Yuliana, "Improving The Skills Of The Residents Of The Abadi Megah Regency Jombang In Microgreens Cultivation Techniques As An Effort To Improve Family Food Security," *Community Empower.*, vol. 7, no. 9, pp. 1562–1567, 2022.
- [15] [Pertanian.go.id, "Budidaya Kacang Hijau," 2019. http://cybex.pertanian.go.id/mobile/artikel/75856/TEKNIK-BUDIDAYA-KACANG-HIJAU/ \(accessed May 12, 2023\).](http://cybex.pertanian.go.id/mobile/artikel/75856/TEKNIK-BUDIDAYA-KACANG-HIJAU/)
- [16] B. Marhaenanto, "Rancangan Sistem Pengendali Suhu Dan Kelembaban Berbasis Mikrokontroler Pada Ruang Pertumbuhan Fodder (Fodder Growth Chamber)," *J. Agroteknologi*, vol. 15, no. 02, 2021.
- [17] M. T. Islam Juel and M. S. Ahmed, "A Smart Auto-Balanced System For Incubation Process," in *1st International Conference on Robotics, Electrical and Signal Processing Techniques, ICREST 2019*, 2019, pp. 564–569, doi: 10.1109/ICREST.2019.8644295.
- [18] S. N. Khine and Z. Tun, "Mongodb On Cloud For Weather Data (Temperature And Humidity) In Sittway," 2020, doi: 10.1109/ICCA49400.2020.9022825.
- [19] A. Farooqi, M. R. Efendi, D. T. Ismail, and W. Darmalaksana, "Design Of Arduino Uno Based Duck Egg Hatching Machine With Sensor DHT22 And PIR Sensor," in *Proceedings - 2020 6th International Conference on Wireless and Telematics, ICWT 2020*, 2020, pp. 14–17, doi: 10.1109/ICWT50448.2020.9243640.
- [20] A. W. Bartlow *et al.*, "Forecasting Zoonotic Infectious Disease Response To Climate Change: Mosquito Vectors And A Changing Environment," *Vet. Sci.*, vol. 6, no. 2, p. 40, 2019.
- [21] A. Al-Fugara, M. Ahmadlou, A. R. Al-Shabeeb, S. AlAyyash, H. Al-Amoush, and R. Al-Adamat, "Spatial Mapping Of Groundwater Springs Potentiality Using Grid Search-Based And Genetic Algorithm-Based Support Vector Regression," *Geocarto Int.*, vol. 37, no. 1, pp. 284–303, 2022, doi: 10.1080/10106049.2020.1716396.
- [22] L. Yao, Z. Fang, Y. Xiao, J. Hou, and Z. Fu, "An intelligent fault diagnosis method for lithium battery systems based on grid search support vector machine," *Energy*, vol. 214, p. 118866, 2021.
- [23] R. M. da S. P. Vieira *et al.*, "Land Degradation Mapping In The MATOPIBA Region (Brazil) Using Remote Sensing Data And Decision-Tree Analysis," *Sci. Total Environ.*, vol. 782, p. 146900, 2021.
- [24] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, 2023.
- [25] T. Plieninger, M. Torralba, T. Hartel, and N. Fagerholm, "Perceived Ecosystem Services Synergies, Trade-Offs, And Bundles In European High Nature Value Farming Landscapes," *Landsc. Ecol.*, vol. 34, pp. 1565–1581, 2019.
- [26] A. Abyad, "The Pareto Principle: Applying the 80/20 Rule to Your Business," *Middle East J.*, 2020.