

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

1. “*Optimizing Lettuce Crop Yield Prediction in an Indoor Aeroponic Vertical Farming System Using IoT-Integrated Machine Learning Regression Models*”

Rajendiran & Rethnaraj, (2024) menekankan pentingnya prediksi hasil tanaman yang akurat dalam pertanian aeroponik modern. Pemanfaatan *machine learning* untuk memberikan prediksi hasil yang tepat dalam sistem aeroponik sangat penting untuk produksi pangan dan manajemen sumber daya. Hal ini mengoptimalkan faktor-faktor seperti hasil, penampilan tanaman, kandungan nutrisi, kualitas, dan rasa, sambil meminimalkan penggunaan sumber daya seperti nutrisi, air, dan energi, yang pada akhirnya meningkatkan efisiensi biaya dan pemanfaatan sumber daya secara optimal.

Pengumpulan data dilakukan menggunakan berbagai sensor seperti sensor TDS untuk mengukur kepekatan larutan nutrisi, sensor kekeruhan, sensor kelembapan, dan sensor cahaya. Data dikumpulkan dalam *growth tower* (alat untuk media tumbuh pada sistem aeroponik) pada *interval* waktu tertentu. Setelah data berhasil dikumpulkan, maka dilakukan *pre-processing* untuk menghilangkan *outliers*. Data dimanfaatkan untuk *training* dan mengevaluasi model *machine learning*. Penelitian ini menggunakan berbagai model *regresi* – *linear*, *support vector*, *random forest*, dan *XGBoost* – untuk memprediksi hasil penanaman aeroponik untuk tanaman selada.

Hasilnya menunjukkan bahwa *XGBoost* memiliki akurasi dan kemampuan prediksi paling akurat untuk memprediksi hasil pertumbuhan selada pada sistem aeroponik. Saran perbaikan yang berkaitan dengan penelitian adalah perlunya mempertimbangkan kondisi lingkungan, pola pertumbuhan tanaman, dan sebagainya.

2. ***“Cost Effective Smart Hydroponic Monitoring and Controlling System Using IoT”***

Penelitian oleh Ullah *et al.*, (2019) bertujuan untuk mengembangkan sistem hidroponik cerdas berbasis IoT yang hemat biaya. Sistem ini dirancang untuk memantau dan mengontrol parameter-parameter penting dalam pertanian hidroponik seperti tingkat air, pH, kelembaban, dan suhu melalui aplikasi seluler. Dengan sistem ini, diharapkan dapat meningkatkan efisiensi pertanian dengan memanfaatkan sumber daya minimal dan mengurangi dampak lingkungan.

Penelitian ini menggunakan pendekatan berbasis perangkat keras dan perangkat lunak. Sistem dirancang menggunakan mikrokontroler ESP32 yang diintegrasikan dengan berbagai sensor seperti sensor pH, kelembaban, suhu, dan tingkat air. Sensor ini terhubung dengan *microcontroller* untuk memantau data secara *real-time*. Data yang dikumpulkan diteruskan ke *server cloud* melalui *gateway router*, sehingga pengguna dapat mengakses informasi tersebut melalui aplikasi seluler. Sistem juga dilengkapi dengan komponen perangkat keras seperti pompa nutrisi, kipas, dan sistem pencahayaan yang dapat dikontrol melalui aplikasi.

Hasil penelitian menunjukkan bahwa sistem hidroponik IoT yang diusulkan mampu berfungsi dengan baik dalam kondisi lingkungan yang berbeda. Sistem ini berhasil memantau dan mengontrol parameter-parameter lingkungan secara otomatis. Data *real-time* yang dikumpulkan menunjukkan konsistensi dalam pengukuran suhu, kelembaban, pH, dan tingkat air. Sistem juga memiliki tampilan aplikasi seluler yang memungkinkan pengguna untuk mengontrol perangkat secara efisien. Prototipe perangkat keras menunjukkan kinerja yang stabil dengan implementasi komponen yang dirancang.

3. ***“Forecasting Area, Production and Yield of Onion in Bangladesh by Using ARIMA Model”***

Penelitian oleh Mila & Parvin, (2019) ini bertujuan untuk memprediksi perubahan area, produksi, dan hasil bawang merah di Bangladesh menggunakan model *Auto Regressive Integrated Moving Average* (ARIMA).

Dengan meningkatnya permintaan bawang merah, penelitian ini berusaha menyediakan proyeksi yang akurat untuk mendukung kebijakan terkait produksi, perdagangan, dan keseimbangan pasokan dan permintaan di pasar domestik.

Data deret waktu (*time series*) tahunan mengenai area, produksi, dan hasil bawang merah di Bangladesh selama periode 1961-2017 dikumpulkan dari situs *Food and Agricultural Organization* (FAO) dan Biro Statistik Bangladesh. Penelitian ini menggunakan metodologi *Box-Jenkins* untuk mengembangkan model ARIMA yang optimal, dengan menguji keakuratan model menggunakan kriteria informasi seperti *AIC*, *AICc*, dan *BIC*. Analisis data dilakukan menggunakan perangkat lunak R. Langkah-langkah dalam metodologi ini meliputi pengujian kestasioneran data, identifikasi model yang sesuai, estimasi parameter, dan validasi model. Setelah proses tersebut, prediksi untuk periode 2018-2027 dilakukan.

Hasil prediksi menunjukkan tren peningkatan pada semua variabel yang dianalisis. Area budidaya bawang merah diperkirakan meningkat dari 193.932,6 hektar pada 2018 menjadi 265.770,9 hektar pada 2027. Produksi bawang merah akan meningkat dari 2.073,61 ribu ton menjadi 3.574,06 ribu ton, sementara hasil bawang merah diproyeksikan naik dari 10.343,17 kg/ha menjadi 12.988,02 kg/ha dalam periode yang sama. Temuan ini menyoroti kebutuhan untuk meningkatkan hasil melalui teknologi yang lebih baik dan dukungan pemerintah, seperti subsidi atau penyediaan benih unggul, untuk memenuhi permintaan yang terus meningkat.

4. ***“Predictive models for Lettuce quality from Internet of Things-based hydroponic farm”***

Gertphol *et al.*, (2018) menyebutkan bahwa pemanfaatan *machine learning* dan IoT untuk pertanian memiliki potensi untuk meningkatkan efisiensi sumber daya. Potensi ini dapat tercapai karena adanya pengumpulan informasi yang akurat tentang kondisi pertumbuhan tanaman.

Pada penelitian ini, data dikumpulkan menggunakan sensor. Setiap hari, sistem akan merekam data intensitas cahaya, kelembapan, suhu udara, suhu air,

EC, dan tingkat keasaman larutan nutrisi. Setiap parameter akan disimpulkan menjadi nilai tertinggi, terendah, dan rata-rata mingguan. Sedangkan data pertumbuhan yang dipantau dan dikumpulkan secara mingguan adalah jumlah daun, ketinggian tanaman, lebar tanaman, diameter batang, dan lebar daun. Data yang dikumpulkan saat panen adalah jumlah daun, ketinggian tanaman, lebar tanaman, diameter batang, lebar daun, tunas tanaman, berat total dalam keadaan basah, berat total dalam keadaan kering, dan kandungan nitrat. Data dinormalisasi lalu dipecah menjadi dua bagian yaitu data *training* dan data *testing*. Model yang digunakan adalah *SGDRegressor*, *BayesianRidge*, *LassoLars*, *ARDRegression*, *PassiveAggressiveRegressor*, *TheilSenRegressor*, *LinearRegression*, serta model yang mendukung banyak variabel seperti SVR (*Support Vector Regression*), MLR (*Multiple Linear Regression*), dan ANN (*Artificial Neural Network*).

Hasil penelitian menunjukkan bahwa peneliti tidak dapat mengandalkan pada satu model saja untuk melakukan prediksi. Prediksi yang direkomendasikan adalah setelah pertumbuhan hingga minggu ketiga karena menghasilkan akurasi yang baik. Saran perbaikan penelitian menyebutkan bahwa untuk memprediksi hasil maka kontrol terhadap variabel, seperti sinar matahari dan suhu udara, harus sangat dikendalikan. Variabel tersebut sangat memengaruhi akurasi dari prediksi.

5. “ARIMA Model-Based Prediction of pH for Hydroponic Vertical Farm”

Penelitian oleh Kaur *et al.*, (2024) ini bertujuan untuk mengembangkan model prediksi berbasis ARIMA untuk memprediksi nilai pH dalam sistem hidroponik vertikal. Mengingat pentingnya stabilitas pH dalam solusi nutrisi hidroponik untuk menunjang pertumbuhan tanaman, studi ini berfokus pada analisis data historis pH untuk memahami pola dan tren, yang selanjutnya digunakan untuk peramalan. Tujuan akhirnya adalah memberikan alat yang efektif bagi petani hidroponik untuk mengelola solusi nutrisi secara proaktif, sehingga meningkatkan efisiensi sumber daya dan hasil tanaman.

Penelitian ini menggunakan model ARIMA untuk analisis deret waktu data pH yang dikumpulkan dari sistem hidroponik vertikal berbasis IoT. Sistem

eksperimental terdiri dari tiga tingkat rak vertikal dengan total kapasitas 81 tanaman, di mana parameter seperti pH dan TDS direkam menggunakan sensor yang terhubung ke platform *cloud ThingSpeak* selama 30 hari. Data ini diuji menggunakan *Augmented Dickey Fuller* (ADF), dan pola-pola deret waktu dianalisis melalui grafik *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Berbagai konfigurasi parameter ARIMA diuji, termasuk penggunaan *auto-SARIMA* untuk menentukan model optimal. Kinerja model diukur dengan metrik *root mean square error* (RMSE) dan nilai *Akaike Information Criterion* (AIC).

6. “IoT Based Hydroponics System Using Deep Neural Networks”

Penelitian oleh Mehra *et al.*, (2018) bertujuan untuk mengembangkan sistem hidroponik berbasis IoT yang cerdas menggunakan *Deep Neural Networks* (DNN). Studi ini difokuskan pada pengendalian otomatis lingkungan hidroponik untuk pertumbuhan tanaman tomat. Penelitian ini menekankan pentingnya sistem cerdas yang dapat mengambil tindakan pengendalian yang sesuai berdasarkan parameter lingkungan yang dikumpulkan secara *real-time*. Tujuan utama dari penelitian ini adalah memberikan sistem yang lebih akurat dan efektif dalam mendukung pertumbuhan tanaman hidroponik tanpa memerlukan intervensi manual yang intensif.

Sistem yang dikembangkan terdiri dari tiga komponen utama: sensor yang dihubungkan ke Arduino untuk mengumpulkan data, Raspberry Pi3 yang menjalankan algoritma *Deep Neural Network* sebagai pusat pengambilan keputusan, dan layanan *cloud* untuk penyimpanan dan visualisasi data. Parameter yang dikumpulkan meliputi pH, suhu, kelembapan, intensitas cahaya, dan tingkat air. Data ini dilatih menggunakan model DNN dengan 10.000 iterasi (*epochs*) untuk menghasilkan delapan tindakan pengendalian, seperti pengoperasian pompa air, kipas angin, atau lampu LED. Algoritma ini berhasil mencapai akurasi prediksi sebesar 88,5%, menunjukkan potensi signifikan untuk implementasi di skala lebih besar.

Sistem ini juga dilengkapi dengan prototipe perangkat keras yang mencakup berbagai sensor, Arduino, dan Raspberry Pi3. Data yang dihasilkan

disimpan di layanan *cloud Firebase*, yang memungkinkan pemantauan dan pengendalian jarak jauh. Penelitian ini memanfaatkan keunggulan DNN untuk menghasilkan prediksi yang lebih akurat dibandingkan pendekatan *Artificial Neural Networks* (ANN) tradisional. Dengan pelabelan data kontrol menjadi delapan kategori, sistem ini mampu mengotomatisasi tindakan berdasarkan kondisi lingkungan yang terdeteksi.

Kesimpulan dari penelitian ini adalah bahwa penggunaan *Deep Neural Networks* dalam sistem hidroponik berbasis IoT dapat meningkatkan efisiensi dan akurasi dalam pengelolaan variabel lingkungan. Peneliti juga merekomendasikan perluasan penelitian di masa depan, termasuk penerapan pada jenis tanaman lain, peningkatan jumlah dataset, dan integrasi dengan lebih banyak tanki hidroponik untuk meningkatkan akurasi dan skala implementasi.

Penelitian ini memberikan kontribusi signifikan terhadap otomatisasi pertanian hidroponik dengan memanfaatkan teknologi IoT dan *machine learning*, serta membuktikan efektivitas DNN dalam mendukung pengelolaan pertanian berbasis data secara *real-time*. Hal ini relevan untuk menciptakan sistem pertanian yang lebih efisien dan adaptif di era modern.

7. “An Automated Hydroponics System Based on Mobile Application”

Penelitian oleh *Kularbphetong et al.*, (2019) bertujuan untuk merancang dan mengembangkan sistem hidroponik otomatis berbasis aplikasi mobile yang didukung teknologi IoT. Penelitian ini menekankan pada pengendalian dan pemantauan otomatis faktor lingkungan penting seperti suhu, kelembapan, intensitas cahaya, dan tingkat air untuk mendukung pertumbuhan tanaman hidroponik. Tujuan utama dari penelitian ini adalah menciptakan sistem yang dapat membantu petani dalam mengelola tanaman hidroponik secara efisien dan produktif, dengan memanfaatkan data *real-time* yang dikumpulkan melalui sensor dan diolah oleh sistem berbasis IoT.

Sistem ini mengintegrasikan berbagai perangkat keras seperti Arduino WeMos D1, modul sensor pH, sensor kelembapan, sensor intensitas cahaya, dan modul relay untuk mengontrol tindakan otomatis seperti pengaturan air, pencahayaan, dan tingkat pH larutan. Data yang dikumpulkan dikirim ke

database melalui koneksi *Wi-Fi* untuk penyimpanan dan analisis. Selain itu, aplikasi *mobile* memungkinkan pengguna untuk memantau kondisi lingkungan tanaman dan mengambil tindakan manual atau otomatis berdasarkan informasi yang diberikan oleh sistem.

Metodologi penelitian ini mencakup pengembangan prototipe sistem yang mampu merekam data sensor setiap 10 menit dan menganalisisnya untuk mengontrol kondisi lingkungan. Prototipe ini mengadopsi protokol *Message Queuing Telemetry Transport* (MQTT) untuk komunikasi antar perangkat, dengan data disinkronkan secara *real-time* menggunakan *database* Firebase. Prototipe diuji untuk mengukur efektivitas pengendalian suhu, pencahayaan, kelembapan, dan tingkat air. Hasil pengujian menunjukkan bahwa sistem ini mampu meningkatkan kualitas tanaman hidroponik dibandingkan dengan metode tradisional.

Hasil evaluasi menunjukkan bahwa sistem hidroponik otomatis berbasis IoT ini mampu meningkatkan efisiensi pengelolaan dan kualitas hasil tanaman. Sistem ini juga dinilai sangat memuaskan oleh pengguna dalam aspek fungsionalitas, keamanan, dan kemudahan penggunaan. Keunggulan utama sistem ini adalah kemampuannya untuk mengotomatisasi proses pengendalian lingkungan secara presisi, dengan hasil yang lebih baik dibandingkan metode konvensional.

Kesimpulan dari penelitian ini adalah bahwa sistem hidroponik berbasis aplikasi *mobile* yang dikembangkan berhasil memberikan solusi yang efisien dan efektif untuk mendukung pertanian modern. Para peneliti merekomendasikan pengembangan lebih lanjut, termasuk integrasi dengan perangkat terkait lainnya dan penerapan teknik *data mining* untuk analisis prediktif. Penelitian ini memberikan kontribusi signifikan terhadap otomatisasi hidroponik, terutama dalam mendukung *smart farming* dan ketahanan pangan global.

8. “Automated Smart Hydroponics System Using Internet of Things”

Penelitian oleh Lakshmanan *et al.*, (2020) bertujuan untuk merancang dan mengimplementasikan sistem hidroponik pintar otomatis berbasis *Internet of Things*. Penelitian ini bertujuan untuk menjawab tantangan global terhadap peningkatan kebutuhan pangan serta kebutuhan akan metode pertanian berkelanjutan yang baru. Sistem ini dirancang untuk memantau dan mengontrol parameter lingkungan hidroponik secara *real-time* dengan memanfaatkan teknologi IoT, seperti NodeMCU, MQTT, dan platform cloud.

Sistem ini terdiri dari beberapa komponen utama, termasuk sensor untuk mengukur parameter penting seperti suhu, kelembapan, pH, dan tingkat air. Data dari sensor ini dikumpulkan dan dikirimkan ke *server cloud* menggunakan protokol MQTT, di mana data dapat dianalisis dan dikontrol melalui antarmuka pengguna grafis (GUI). *NodeRed* digunakan sebagai alat penghubung perangkat keras, API, dan layanan *online* untuk memproses data secara efisien. Selain itu, sistem ini mendukung kontrol otomatis berdasarkan data *real-time*, dengan kemampuan pengguna untuk mengontrol manual melalui aplikasi *mobile*.

Metodologi yang diterapkan mencakup pembangunan prototipe sistem yang menggunakan NFT untuk menyediakan nutrisi ke tanaman. Sistem ini dirancang dalam tiga tahap: bagian klien (*client-side*) untuk pemrograman *microcontroller* dan sensor, bagian *server* untuk pengelolaan data menggunakan NodeRed dan MQTT, serta antarmuka pengguna grafis (GUI) untuk memantau dan mengontrol sistem secara *real-time*. Penggunaan MQTT dan *NodeRed* memberikan efisiensi dan keamanan data yang tinggi, sementara desain GUI memungkinkan pengguna untuk berinteraksi dengan sistem secara mudah.

Hasil evaluasi menunjukkan bahwa sistem ini berhasil meningkatkan efisiensi pengelolaan hidroponik, dengan kemampuan untuk memantau dan mengontrol parameter secara *real-time*. Pengujian pada berbagai parameter lingkungan menunjukkan bahwa sistem ini mampu menghasilkan tanaman hidroponik dengan kualitas yang lebih baik dibandingkan metode

konvensional. Evaluasi juga menyoroti performa sistem yang stabil, dengan latensi rendah dan kemampuan untuk memberikan notifikasi kepada pengguna melalui aplikasi *mobile* dan layanan SMS.

Kesimpulan dari penelitian ini adalah bahwa integrasi IoT dengan sistem hidroponik dapat memberikan solusi yang efektif untuk mendukung pertanian pintar dan memenuhi kebutuhan pangan global yang terus meningkat. Peneliti merekomendasikan pengembangan lebih lanjut dengan menambahkan analitik data dan algoritma kecerdasan buatan untuk meningkatkan akurasi prediksi pertumbuhan tanaman serta efisiensi sistem. Penelitian ini memberikan kontribusi signifikan terhadap otomatisasi sistem hidroponik dan menjadi langkah penting dalam mendukung keberlanjutan pertanian modern.

9. ***“Hydroponic Nutrient Control System Based on Internet of Things and K-Nearest Neighbors”***

Penelitian oleh Adidrana & Surantha (2019) bertujuan untuk mengembangkan sistem kontrol nutrisi hidroponik berbasis *Internet of Things* menggunakan algoritma *k-Nearest Neighbors* (KNN). Penelitian ini difokuskan pada pengelolaan nutrisi hidroponik secara otomatis untuk tanaman selada dalam sistem NFT. Sistem ini bertujuan untuk meningkatkan efisiensi pengelolaan nutrisi dengan memberikan rekomendasi tindakan berbasis klasifikasi kondisi nutrisi.

Sistem ini terdiri dari beberapa komponen utama, seperti sensor untuk mengukur pH, TDS, dan suhu air. Data sensor dikumpulkan menggunakan modul *Arduino* dan *NodeMCU*, kemudian dikirim ke platform IoT *ThingSpeak* melalui komunikasi nirkabel ESP8266. Data tersebut dianalisis menggunakan algoritma KNN untuk mengklasifikasikan kondisi nutrisi ke dalam 27 kemungkinan label berdasarkan kombinasi nilai pH, TDS, dan suhu. Setiap label ini mewakili kondisi nutrisi tertentu dan dikaitkan dengan tindakan pengendalian, seperti menghidupkan atau mematikan pompa nutrisi A/B, pompa pH, atau filter TDS.

Metodologi penelitian ini mencakup tiga tahapan utama: evaluasi modul sensor, modul aktuator, dan *server* KNN. Modul sensor diuji untuk memastikan

akurasi pengukuran data nutrisi secara *real-time*. Modul aktuator dirancang untuk memberikan tindakan otomatis berdasarkan klasifikasi KNN. *Server* KNN menjalankan algoritma KNN untuk memprediksi klasifikasi nutrisi menggunakan data *real-time* dari ThingSpeak. Data ini kemudian digunakan untuk memberikan perintah pada aktuator.

Hasil evaluasi menunjukkan bahwa algoritma KNN mampu mencapai akurasi prediksi sebesar 93,3% dengan nilai $k=5$. Sistem ini berhasil mengotomatisasi kontrol nutrisi hidroponik, termasuk pengelolaan pH, TDS, dan suhu nutrisi secara simultan. Selain itu, pengujian *real-time* menunjukkan bahwa sistem ini dapat merespons perubahan kondisi nutrisi dengan cepat dan akurat, meningkatkan efisiensi pengelolaan hidroponik dibandingkan metode manual.

Kesimpulan dari penelitian ini adalah bahwa penggunaan algoritma KNN dalam sistem kontrol nutrisi hidroponik berbasis IoT memberikan solusi yang efektif untuk meningkatkan produktivitas dan efisiensi pertanian hidroponik. Peneliti merekomendasikan pengembangan lebih lanjut, termasuk peningkatan *dataset*, penggunaan sensor tambahan, dan integrasi dengan analitik data atau kecerdasan buatan untuk prediksi yang lebih akurat. Penelitian ini memberikan kontribusi signifikan terhadap otomatisasi pertanian pintar dan mendukung keberlanjutan sistem hidroponik di era modern.

10. “Automated Hydroponics Nutrition Plants Systems Using Arduino Uno Microcontroller Based on Android”

Penelitian oleh Sihombing *et al.*, (2018) bertujuan untuk mengembangkan sistem hidroponik otomatis berbasis *Arduino Uno microcontroller* yang dikontrol melalui aplikasi *Android*. Fokus utama penelitian ini adalah menyediakan sistem yang dapat mengelola nutrisi tanaman hidroponik secara efisien, dengan menggunakan sensor untuk mendeteksi parameter penting seperti ketinggian larutan dan suhu lingkungan sekitar tanaman.

Sistem ini menggunakan *Ultrasonic Sensor HC-SR04* untuk mendeteksi ketinggian larutan nutrisi dalam tabung hidroponik, serta *Temperature Sensor LM35* untuk memantau suhu sekitar. Data yang dihasilkan sensor dikirim ke

Arduino Uno microcontroller dan ditampilkan pada layar LCD. Selain itu, modul *Wi-Fi ESP8266* digunakan untuk mengirim data ke aplikasi *Android*, yang memungkinkan pengguna memantau dan mengontrol sistem secara real-time. Logika *if-else* diterapkan dalam pemrograman untuk mengatur aktuator seperti pompa nutrisi berdasarkan data dari sensor.

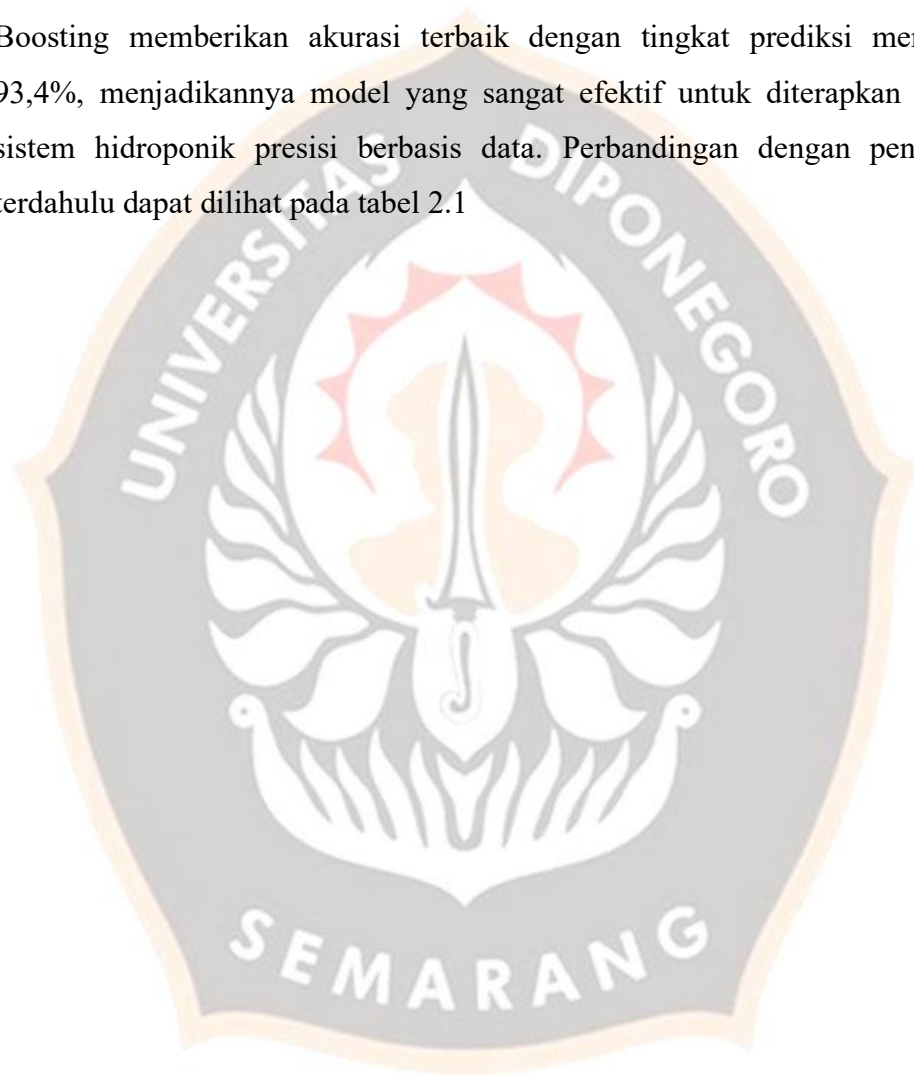
Metodologi penelitian ini mencakup pengembangan prototipe sistem yang diuji dalam pengelolaan nutrisi tanaman hidroponik. Sistem ini dirancang untuk secara otomatis mengatur ketinggian larutan nutrisi hingga batas tertentu. Ketika ketinggian larutan di bawah ambang batas, pompa nutrisi diaktifkan untuk menambah larutan, dan secara otomatis berhenti ketika ketinggian optimal tercapai. Data seperti suhu dan ketinggian larutan dipantau secara terus-menerus dan disimpan pada platform *Thingspeak* untuk keperluan analisis.

Hasil penelitian menunjukkan bahwa sistem ini mampu mempertahankan ketinggian larutan nutrisi hingga 6 cm dalam tabung hidroponik. Sensor mendeteksi perubahan ketinggian larutan secara akurat, dan pompa nutrisi berfungsi dengan baik sesuai perintah dari *microcontroller*. Eksperimen menunjukkan bahwa penggunaan sistem ini menghasilkan tanaman hidroponik dengan tingkat pertumbuhan yang lebih cepat dibandingkan metode konvensional. Sistem juga menawarkan kenyamanan bagi pengguna karena dapat dikontrol melalui perangkat mobile.

Kesimpulan dari penelitian ini adalah bahwa sistem hidroponik otomatis berbasis *Arduino Uno microcontroller* dan aplikasi *Android* berhasil menyediakan solusi yang efisien untuk pengelolaan nutrisi tanaman. Peneliti merekomendasikan pengembangan lebih lanjut, seperti penambahan fitur untuk mendeteksi tingkat keasaman (*pH*) dan viskositas larutan, serta perluasan skala sistem untuk aplikasi yang lebih luas dalam industri hidroponik. Penelitian ini memberikan kontribusi penting bagi otomatisasi pertanian dan mendukung pertanian berkelanjutan.

Penelitian terbaru menunjukkan bahwa pemanfaatan machine learning dalam sistem hidroponik semakin diarahkan untuk membangun model

prediktif pertumbuhan tanaman yang akurat dan efisien. Dalam penelitian yang dilakukan oleh Rahmadi *et al.*, (2025), empat algoritma machine learning (*Gradient Boosting*, *Random Forest*, SVM, dan K-NN) dibandingkan untuk memprediksi pola pertumbuhan selada dalam sistem NFT menggunakan data primer dari empat siklus panen. Hasilnya menunjukkan bahwa Gradient Boosting memberikan akurasi terbaik dengan tingkat prediksi mencapai 93,4%, menjadikannya model yang sangat efektif untuk diterapkan dalam sistem hidroponik presisi berbasis data. Perbandingan dengan penelitian terdahulu dapat dilihat pada tabel 2.1



SEKOLAH PASCASARJANA

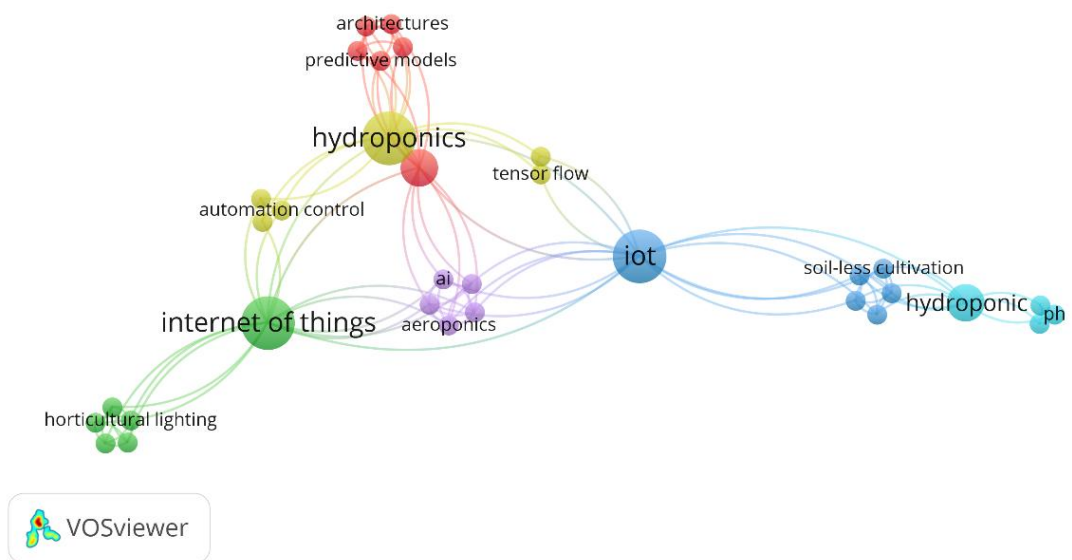
Tabel 2.1 Perbandingan penelitian terdahulu

Penelitian	Compare	Contrast	Model Machine Learning	Metode Evaluasi	Data Preprocessing	Variabel yang Dinilai	Tujuan Penelitian
(Rajendiran & Rethnaraj, 2024)	Fokus pada IoT dan machine learning; menggunakan beberapa model prediksi	Spesifik pada aeroponik; menggunakan model regresi seperti XGBoost; menargetkan prediksi hasil tanaman selada	<i>Linear, Support Vector, Random Forest, XGBoost</i>	Perbandingan akurasi model	Penghapusan outlier	TDS, kekeruhan, kelembapan, intensitas cahaya	Mengoptimalkan hasil panen tanaman selada dalam sistem aeroponik
(Ullah <i>et al.</i> , 2019)	Sistem berbasis IoT; pemantauan lingkungan	Fokus pada sistem hidroponik hemat biaya; pemantauan data secara real-time	Tidak menggunakan machine learning	Evaluasi kinerja prototipe	Pengumpulan data real-time	pH, kelembapan, suhu, tingkat ketinggian air	Mengembangkan sistem monitoring hidroponik berbasis IoT yang hemat biaya
(Mila & Parvin, 2019)	Menggunakan ARIMA untuk analisis deret waktu	Fokus pada bawang merah; data dari basis data nasional	ARIMA	AIC, AICc, BIC	ADF	Luas area, hasil produksi (yield)	Memprediksi tren produksi bawang merah
(Gertphol <i>et al.</i> , 2018)	Integrasi IoT; menggunakan berbagai model machine learning	Penekanan pada kualitas tanaman selada; menguji berbagai model ML	<i>SGDRegressor, BayesianRidge, LassoLars, ARDRegression, PassiveAggressiveRegressor, TheilSenRegressor, LinearRegression, SVR, MLR, ANN</i>	Akurasi prediksi berdasarkan waktu	Normalisasi, pemisahan data training-testing	Intensitas cahaya, kelembapan, suhu udara, suhu air, EC, pH.	Memprediksi metrik kualitas tanaman selada
(Kaur <i>et al.</i> , 2024)	Pengumpulan data berbasis IoT; pemodelan ARIMA	Fokus pada prediksi pH; data historis jangka pendek	ARIMA dan auto-SARIMA	RMSE, AIC	Pengujian ADF, analisis ACF dan PACF	pH, TDS	Memprediksi nilai pH dalam sistem hidroponik

Tabel 2.1 Perbandingan penelitian terdahulu (lanjutan)

Penelitian	Compare	Contrast	Model Machine Learning	Metode Evaluasi	Data Preprocessing	Variabel yang Dinilai	Tujuan Penelitian
(Mehra <i>et al.</i> , 2018)	Fokus pada integrasi IoT dengan DNN untuk prediksi dan kontrol variabel hidroponik	Berbasis sistem IoT dengan pemrosesan data menggunakan Deep Neural Networks	Deep Neural Networks (DNN)	Akurasi prediksi (88,5%)	Pelabelan kontrol (0-7)	pH, suhu, kelembapan, intensitas cahaya	Mengembangkan sistem hidroponik berbasis IoT yang cerdas untuk memantau dan mengontrol parameter lingkungan secara otomatis melalui DNN untuk pertumbuhan tanaman tomat.
(Kularbphettong <i>et al.</i> , 2019)	Integrasi IoT dengan aplikasi mobile untuk pengendalian parameter hidroponik	Tidak menggunakan ML; fokus pada pemantauan dan pengendalian real-time melalui antarmuka mobile	Tidak menggunakan <i>machine learning</i>	Pengujian prototipe, uji fungsionalitas aplikasi	Tidak disebutkan	pH, suhu, kelembapan, tingkat air	Meningkatkan efisiensi pengelolaan hidroponik melalui aplikasi mobile yang memungkinkan kontrol dan pemantauan real-time terhadap parameter lingkungan.
(Lakshmana <i>n et al.</i> , 2020)	Sistem berbasis IoT menggunakan protokol MQTT	Fokus pada stabilitas dan efisiensi data dalam integrasi IoT	Tidak menggunakan <i>machine learning</i>	Pengujian stabilitas dan latensi data	Tidak disebutkan	Suhu, kelembapan, pH, tingkat air	Mengembangkan sistem IoT yang efisien untuk pengelolaan lingkungan hidroponik secara otomatis dan real-time.
(Adidrana & Surantha, 2019)	Menggunakan algoritma KNN untuk klasifikasi kondisi nutrisi hidroponik	Berfokus pada prediksi tindakan pengendalian nutrisi secara otomatis	K-Nearest Neighbors (KNN)	Akurasi klasifikasi (93,3%)	Tidak disebutkan	pH, TDS, suhu	Mengotomatisasi kontrol nutrisi hidroponik berbasis IoT untuk meningkatkan efisiensi dan produktivitas.
(Sihombing <i>et al.</i> , 2018)	Menggunakan Arduino Uno dengan kontrol berbasis aplikasi Android	Tidak menggunakan ML; fokus pada otomatisasi berbasis logika if-else	Tidak menggunakan <i>machine learning</i>	Uji prototipe dan keakuratan data	Tidak disebutkan	Ketinggian larutan, suhu	Mengembangkan sistem hidroponik otomatis untuk pengelolaan nutrisi berbasis microcontroller dengan kontrol melalui aplikasi Android.

Kajian literatur sistematis yang dilakukan oleh Rahmadi *et al.*, (2025), mengulas tren penelitian otomatisasi hidroponik dalam dua pendekatan utama, yaitu ‘*internet of things*’ dan ‘*machine Learning*’. Penelitian tersebut tidak hanya merangkum model-model dan teknologi yang digunakan dalam studi-studi terdahulu, tetapi juga melakukan pemetaan bibliometrik untuk mengidentifikasi arah perkembangan riset secara global. Salah satu hasil penting dari kajian ini adalah visualisasi jaringan kata kunci menggunakan perangkat lunak *VOSviewer* yang memetakan keterhubungan antar topik utama dalam bidang tersebut. Visualisasi *VOSviewer* dapat dilihat pada gambar 2.1 :



Gambar 2.1 Visualisasi VOSviewer Otomatisasi hidroponik

Pemetaan ini memberikan gambaran posisi strategis dari tema-tema riset seperti ‘*hydroponics*,’ ‘*IoT*,’ ‘*machine learning*,’ dan ‘*predictive model*’ yang membentuk kluster-kluster yang kuat dalam literatur ilmiah. Dengan mengutip hasil pemetaan tersebut, penelitian ini memperkuat justifikasi bahwa fokus pengembangan sistem simulasi otomatisasi berbasis *machine learning* yang ditawarkan masih merupakan celah yang belum banyak dieksplorasi secara mendalam dalam penelitian sebelumnya.

Hasil visualisasi yang ditampilkan pada gambar 2.1 (diambil dari Rahmadi et al., 2025) menunjukkan bahwa topik '*hydroponics*' berada di pusat jaringan yang dikelilingi oleh klaster kata kunci yang berkaitan erat seperti '*sensor*,' '*monitoring*,' '*automation*,' dan '*IoT*.' Topik-topik ini membentuk kelompok dominan yang menandakan bahwa sebagian besar riset saat ini masih terkonsentrasi pada aspek pengawasan dan pengendalian sistem berbasis IoT. Di sisi lain, klaster yang berisi kata kunci seperti '*machine learning*,' '*growth*,' dan '*prediction*' terlihat lebih terpisah, menunjukkan bahwa pendekatan prediktif masih belum terintegrasi secara menyeluruh ke dalam sistem otomatisasi hidroponik.

Pemetaan ini memperjelas bahwa terdapat peluang penelitian yang besar dalam menggabungkan antara sistem prediksi berbasis data dengan simulasi pertumbuhan yang dapat digunakan secara praktis. Penelitian disertasi ini hadir untuk menjawab gap tersebut dengan mengembangkan model prediktif berbasis Prophet dan mengintegrasikannya ke dalam sistem aplikasi simulasi berbasis web (HydroSim).

2.2 Dasar Teori

2.2.1 Hidroponik

Hidroponik merupakan metode bercocok tanam modern yang menggunakan larutan nutrisi tanpa memerlukan tanah sebagai media tanam. Istilah hidroponik berasal dari bahasa Yunani, yaitu *hydro* yang berarti air dan *ponos* yang berarti kerja, sehingga hidroponik berarti "pengelolaan dengan air" (Bambang et al., 2017). Sistem ini memungkinkan tanaman untuk tumbuh dengan menyalurkan nutrisi penting secara langsung melalui air, tanpa memerlukan media tanah, selama unsur hara, mineral, dan pH dijaga dalam kisaran optimal (Perry, 2012). Dalam perkembangannya, hidroponik telah menjadi solusi pertanian modern yang dapat diusahakan di berbagai lokasi, baik di lahan terbuka, perkotaan, bahkan di apartemen, dan dapat dilakukan sepanjang tahun tanpa bergantung pada musim (Wirawan, 2016).

Sistem hidroponik menawarkan berbagai keunggulan. Pertama, tanaman yang dihasilkan memiliki kualitas yang lebih baik karena lingkungan pertumbuhannya bersih dan aman dari serangan hama serta penyakit, sehingga hasil panen lebih sehat

dan berkualitas tinggi (Wirawan, 2016). Kedua, hidroponik hemat tempat karena tidak memerlukan lahan yang luas untuk menghasilkan kuantitas tanaman yang banyak. Ketiga, sistem ini hemat pupuk karena nutrisi dalam air langsung diserap oleh akar, sehingga lebih efektif dibandingkan metode konvensional. Keempat, tanaman tumbuh lebih cepat karena pasokan nutrisi diterima secara tepat dan konsisten. Kelima, hidroponik menghemat tenaga dan waktu karena kebutuhan lahan kecil memungkinkan efisiensi dalam pengelolaan (Harsono, 2020).

Namun, hidroponik juga memiliki beberapa kekurangan. Untuk skala usaha, biaya awal instalasi relatif tinggi karena membutuhkan peralatan khusus dan sistem yang tepat, seperti NFT yang dikembangkan oleh Allen Cooper pada tahun 1965 (Perry, 2012). Selain itu, diperlukan keterampilan dan pemahaman yang baik terkait pengelolaan nutrisi dalam air, mengingat keberhasilan sistem ini bergantung pada stabilitas nutrisi, mineral, dan pH dalam larutan (Harsono, 2020).

Seiring dengan kemajuan teknologi, hidroponik telah berkembang pesat dan digunakan secara luas, terutama untuk produksi komersial sayuran, buah, tanaman obat, dan tanaman hias. Sistem ini menjadi populer karena kepraktisannya, terutama di kawasan perkotaan, memungkinkan masyarakat untuk menghasilkan produk pertanian berkualitas tinggi dengan memanfaatkan ruang yang terbatas (Perry, 2012; Wirawan, 2016).

2.2.1.1 Variabel Hidroponik

Variabel kunci dalam sistem hidroponik meliputi suhu, kelembapan, pH, dan konsentrasi nutrisi dalam larutan. Rentang suhu ideal untuk pertumbuhan tanaman adalah antara 18°C hingga 25°C, sedangkan tingkat pH yang optimal berkisar antara 5,5 hingga 6,5. Konsentrasi nutrisi diukur menggunakan EC, dengan rentang ideal untuk tanaman selada antara 1,2 hingga 1,8 mS/cm. Selain itu, kelembapan ideal untuk tanaman hidroponik adalah antara 50% hingga 70% (López Mora *et al.*, 2024). Jika variabel-variabel ini tidak dikontrol dengan baik, tanaman dapat mengalami stres, yang berdampak negatif terhadap hasil panen (Rajendiran & Rethnaraj, 2024). Tabel 2.2, merangkum nilai-nilai ideal untuk variabel kunci dalam budidaya selada menggunakan sistem hidroponik.

Tabel 2.2 Variabel ideal untuk pertumbuhan selada pada sistem hidroponik

(sumber: Sambo & et al., 2019)

Variable	Ideal Range
Suhu udara	18°C - 25°C
pH	5.5 - 6.5
Kelembapan	50% - 70%
EC (mS/cm)	1.2 - 1.8
Kepekatan Nutrisi (ppm)	800 - 1000 ppm

Sistem hidroponik modern sering kali menggunakan sensor otomatis untuk memantau variabel-variabel ini secara *real-time*, memungkinkan penyesuaian segera jika terjadi perubahan drastis pada kondisi lingkungan. Sensor-sensor ini dapat mendeteksi suhu, pH, kelembapan, dan kadar nutrisi, serta mengirimkan data secara *real-time* ke *server* untuk analisis dan penyesuaian otomatis (Sambo & et al., 2019). Simulasi otomatis yang berbasis data lingkungan musiman sangat penting untuk menjaga stabilitas dan efisiensi pertumbuhan tanaman, terutama selama kondisi cuaca yang tidak terduga (Ullah *et al.*, 2019).

2.2.1.2 Kebutuhan Hidroponik

Kebutuhan hidroponik dapat dikelompokkan menjadi dua aspek utama, yaitu kebutuhan lingkungan dan kebutuhan air. Kebutuhan lingkungan meliputi faktor-faktor seperti cahaya, suhu, dan kelembapan. Cahaya diperlukan untuk proses fotosintesis, sementara suhu dan kelembapan harus berada pada tingkat optimal agar tanaman tidak mengalami stres akibat kondisi yang terlalu panas, terlalu dingin, atau kelembapan yang berlebihan (Tjendapati, 2017). Di sisi lain, kebutuhan air menjadi sangat vital karena dalam sistem hidroponik, air menggantikan tanah sebagai media untuk memberikan nutrisi kepada tanaman. Nutrisi yang terkandung dalam air harus dijaga pada tingkat yang sesuai karena sangat memengaruhi pertumbuhan dan kesehatan tanaman (Isnain, 2020).

Beberapa indikator utama kebutuhan air dalam hidroponik meliputi pH, EC, dan TDS. pH mengukur tingkat keasaman atau kebasaan larutan, dengan skala logaritmik dari 0 hingga 14. Air murni memiliki pH netral sebesar 7, sedangkan larutan nutrisi untuk tanaman hidroponik idealnya memiliki pH dalam rentang 5,5 hingga 6,5. Rentang ini memastikan unsur hara penting tersedia dan dapat diserap secara optimal oleh tanaman (Isnain, 2020).

EC mengukur daya hantar listrik larutan, yang mencerminkan konsentrasi unsur hara terlarut. EC dihitung dalam satuan mS/cm, dengan nilai ideal yang berbeda untuk setiap jenis tanaman tergantung pada kebutuhannya. EC memberikan gambaran tentang jumlah nutrisi yang tersedia dalam larutan dan penting untuk memastikan bahwa nutrisi yang diberikan tidak kurang atau berlebihan (Isnani, 2020).

TDS mengukur jumlah nutrisi terlarut dalam larutan dan dinyatakan dalam satuan ppm (*part per million*). 1 ppm setara dengan 1 mg/liter atau 1 gram/1.000 liter air. TDS menunjukkan total nutrisi yang tersedia untuk diserap oleh akar tanaman. Memantau dan menjaga kandungan TDS dalam kondisi ideal penting untuk memastikan nutrisi yang cukup dan seimbang bagi tanaman (Isnani, 2020).

Setiap tanaman sayuran mempunyai kebutuhan pH, EC dan TDS yang berbeda sesuai dengan karakteristik masing-masing sayuran. Selain faktor lingkungan seperti suhu dan kelembaban ataupun lokasi penanaman juga dapat mempengaruhi kebutuhan sayuran terhadap kandungan air (Tjendapati, 2017). Rentang kebutuhan nilai TDS, EC dan pH untuk berbagai sayuran dapat dilihat pada tabel 2.3.

Tabel 2.3 Kebutuhan TDS, EC, dan pH sayuran daun

(Sumber: Tjendapati, 2017)

Sayuran	TDS (ppm)	EC (mS/cm)	pH Ideal
Asparagus	980-1.260	1,4-1,8	6,0-6,8
Basil	700-1.120	1,0-1,6	5,5-6,5
Bayam	1.260-1.610	1,8-2,3	6,0-7,0
Brokoli	1.960-2.450	2,8-3,5	6,0-6,8
Bunga kol	1.050-1.400	1,5-2,0	6,5-7,0
Daun bawang	1.260-1.540	1,8-2,2	6,0-6,5
Daun Mint	1.400-1.680	2,0-2,4	5,5-6,0
Kailan	1.050-1.400	1,5-2,0	5,5-6,5
Kangkung	1.050-1.400	1,5-2,0	5,5-6,5
Kemangi	700-1.120	1,0-1,6	5,5-6,5
Kubis/kol	750-2.100	2,5-3,1	6,5-7,0
Pakcoi	1.050-1.400	1,5-2,0	7,0
Peterseli	560-1.260	0,8-1,8	5,5-6,0
Romein	560-840	0,8-1,2	6,0-7,0

Tabel 2.3 Kebutuhan TDS, EC, dan pH sayuran daun (lanjutan)

Sayuran	TDS (ppm)	EC(mS/cm)	pH Ideal
Sawi/caisim/sosin	1.050-1.400	1,5-2,0	5,5-6,5
Sawi pahit	840-1.680	1,6-2,4	6,0-6,5
Selada	560-840	0,8-1,2	6,0-7,0
Selada Endive	1.400-1.680	2,0-2,4	5,5
Selada Lororosa	560-840	0,8-1,2	6,0-7,0
Selada Air	560-840	0,8-1,2	6,0-7,0
Selada butterhead	560-840	0,8-1,2	6,0-7,0
Seledri	1.260-1.680	1,8-2,4	6,5

Selain indikator air, faktor lingkungan seperti suhu, kelembapan, intensitas cahaya, dan konsentrasi CO₂ di udara juga memengaruhi efektivitas penyerapan nutrisi oleh tanaman hidroponik. Hubungan antara indikator lingkungan dan kandungan air menunjukkan bahwa rasio yang tepat antara faktor-faktor ini sangat penting untuk menjaga keberlanjutan pertumbuhan tanaman dalam sistem hidroponik (Singh & Singh, 2017). Dengan demikian, pengelolaan indikator pH, EC, dan TDS, serta kondisi lingkungan yang optimal, menjadi kunci keberhasilan dalam sistem hidroponik.

Kebutuhan hidroponik pada tanaman selada dalam sistem NFT dipengaruhi oleh faktor lingkungan dan kandungan air, yang saling berhubungan. Indikator utama kebutuhan air meliputi pH, TDS, dan EC. Nilai ideal dari indikator-indikator yang terlihat pada tabel 2.3 tidak bersifat baku, karena dapat dipengaruhi oleh faktor lingkungan seperti suhu, kelembapan udara, dan letak geografis kebun. Misalnya, kebun di daerah panas dengan suhu tinggi cenderung meningkatkan penguapan, sehingga nilai TDS dan EC perlu dinaikkan untuk menyesuaikan dengan kondisi lingkungan (Murata *et al.*, 2013).

Secara umum, nilai ideal untuk tanaman selada adalah sebagai berikut: pH larutan nutrisi dalam kisaran 5,5–6,5 untuk memastikan penyerapan nutrisi yang optimal, EC berkisar antara 1,2–1,6 mS/cm, dan TDS antara 400–800 ppm untuk fase remaja serta 800–1200 ppm untuk fase pembesaran. Pada fase remaja, kebutuhan nutrisi air lebih rendah (400–500 ppm) karena akar tanaman belum berkembang sempurna dan nutrisi yang berlebih dapat menyebabkan akar busuk.

Sedangkan pada fase pembesaran, kebutuhan nutrisi meningkat menjadi 800–1000 ppm karena tanaman membutuhkan lebih banyak nutrisi untuk mendukung pertumbuhan yang lebih besar (Tjendapati, 2017).

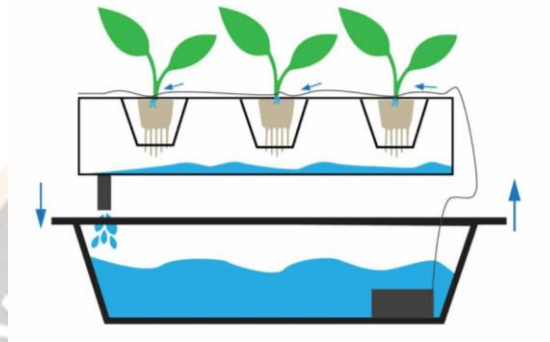
Faktor lingkungan seperti suhu, kelembapan relatif, dan intensitas cahaya juga memengaruhi efektivitas kebutuhan air. Suhu ideal untuk selada adalah 18–24°C pada siang hari dan 10–16°C pada malam hari, dengan kelembapan relatif 50–70% serta intensitas cahaya 1000–1500 *lux*. Lingkungan yang sesuai ini membantu mengoptimalkan pertumbuhan tanaman tanpa menyebabkan stres atau gangguan fisiologis (Murata *et al.*, 2013)

Pada fase semai, kebutuhan lingkungan menjadi lebih dominan dibandingkan kebutuhan air. Bibit selada hanya memerlukan penyemprotan air untuk menjaga kelembapan media tanam tanpa larutan nutrisi, karena akar belum berkembang sempurna. Suhu lingkungan tidak boleh terlalu panas, namun tetap memerlukan cahaya untuk mempercepat pertumbuhan tunas. Dengan demikian, pengelolaan pH, TDS, EC, dan faktor lingkungan secara holistik sangat penting untuk menjaga keberlangsungan dan optimalisasi pertumbuhan tanaman hidroponik.

2.2.1.3 *Nutrient Film Technique* (NFT)

Nutrient Film Technique (NFT) merupakan salah satu jenis sistem hidroponik yang dikembangkan pertama kali oleh Dr. A.J Cooper di Glasshouse Crops Research Institute, Littlehampton, Inggris pada akhir tahun 1960-an dan berkembang pada awal 1970-an secara komersial. NFT mengacu pada penggunaan saluran atau tabung kecil untuk sirkulasi berkelanjutan dari larutan nutrisi melalui susunan tanaman. Ini adalah sistem berbasis listrik dengan saluran ditempatkan miring sehingga nutrisi dapat dipindahkan ke segmen yang lebih tinggi juga untuk memfasilitasi pergerakan nutrisi. Konsep dasar NFT ini adalah suatu metode budidaya tanaman dengan akar tanaman tumbuh pada lapisan nutrisi yang dangkal dan tersirkulasi sehingga tanaman dapat memperoleh cukup air, nutrisi dan oksigen. Tanaman tumbuh dalam lapisan *polyethylene* dengan akar tanaman terendam dalam air yang berisi larutan nutrisi yang disirkulasikan secara terus menerus dengan pompa (Affan, 2004).

Pada sistem NFT, bidang tanam diletakkan pada kemiringan rendah (kira-kira 5 derajat) sehingga air akan mengalir mengikuti gaya gravitasi. Air bernutrisi yang dialirkan akan menangkap oksigen dari udara sehingga akan semakin baik bagi pertumbuhan tanaman. Bagian akar diletakkan pada bidang aliran air ini sehingga tumbuhan dapat menyerap nutrisi yang dibutuhkan (Harsono, 2020).



Gambar 2.2 Gambaran sistem NFT

Melalui gambar 2.2 kita memperoleh gambaran sederhana dari sistem NFT, kita dapat melihat bahwa air nutrisi mengalir dari nutrisi air melalui pompa, diserap oleh akar tanaman dan kembali lagi ke tandon nutrisi air.

2.2.1.3.1 Kelebihan Hidroponik NFT

Hidroponik, khususnya teknik NFT, telah menjadi pilihan utama dalam skala besar atau komersial karena keunggulannya dalam meningkatkan produktivitas tanaman. NFT pertama kali dikenalkan pada tahun 1960-an, dengan prinsip dasar mengalirkan air kaya nutrisi secara terus menerus ke akar tanaman. Aliran air ini, yang diletakkan pada bidang miring sekitar 5 derajat, memungkinkan penyerapan oksigen dari udara, memberikan manfaat besar bagi pertumbuhan tanaman (Harsono, 2020). Sistem ini mendukung kontrol yang baik terhadap kebutuhan air dan nutrisi, termasuk menyesuaikan konsentrasi larutan nutrisi sesuai umur dan jenis tanaman, serta memungkinkan penelitian dengan variabel terkontrol (Affan, 2004).

Keunggulan utama hidroponik dibandingkan dengan kultur tanah terletak pada efisiensi penggunaan air dan pupuk, serta pengaturan nutrisi yang lebih presisi. Hidroponik juga memungkinkan ketersediaan pangan di daerah dengan tanah yang sulit ditanami atau miskin unsur hara, mengurangi dampak dari penyakit tanah, dan

meningkatkan kepadatan tanam sehingga menghasilkan produktivitas per hektar yang lebih tinggi (Resh, 2022). Sistem ini memungkinkan siklus tanam yang lebih pendek dengan hasil tanaman yang lebih seragam dan berkualitas tinggi, karena media yang digunakan steril dan terhindar dari serangan hama dan penyakit yang umumnya ditemukan di tanah (Harsono, 2020)

Namun, sistem hidroponik, termasuk NFT, memiliki beberapa keterbatasan. Biaya modal awal yang tinggi, ketergantungan pada listrik, serta risiko penyebaran penyakit seperti *Fusarium* dan *Verticillium* melalui sistem air menjadi tantangan utama. Meskipun demikian, banyak dari kekurangan ini dapat diatasi dengan pemilihan varietas tanaman yang tahan penyakit, penggunaan batang bawah, dan alat pengujian nutrisi yang lebih baik (Resh, 2022). Secara keseluruhan, keunggulan hidroponik dalam efisiensi, produktivitas, dan adaptabilitasnya menjadikannya solusi unggul dalam memenuhi kebutuhan pertanian modern.

Hidroponik menawarkan berbagai kelebihan dan keunggulan jika dibandingkan dengan pertanian konvensional. Pada tabel 2.4 disajikan perbandingan hidroponik dan pertanian konvensional.

Tabel 2.4 Perbandingan hidroponik dan pertanian konvensional

Aspek	Hidroponik	Pertanian Konvensional
Lahan	Tidak membutuhkan lahan luas bisa mengoptimalkan area sempit	Perlu lahan luas kuantitas tanaman tergantung lahan
Hama	Mengurangi resiko hama dan penyakit dari tanah	Kesuburan alami tanah mempengaruhi pertumbuhan perlu pestisida kimia
Biaya	Memerlukan biaya investasi awal yang tinggi untuk perawatan, biaya operasional tinggi untuk beberapa sistem yang perlu aliran air terus menerus	Tidak perlu investasi awal memanfaatkan lahan tanah, biaya operasional lebih rendah dengan pemupukan berkala
Kontrol	Mempunyai kendali penuh atas kondisi nutrisi dan lingkungan	Tergantung pada kondisi lingkungan dan cuaca yang alami, semakin besar kuantitas tanaman dan lahan kontrol lebih sulit
Penggunaan Air	Air cenderung lebih hemat karena air berputar terus dan lebih efektif terserap tanaman	Membutuhkan banyak air terutama pada musim kemarau yang bisa meningkatkan biaya
Manfaat Lingkungan	Mengurangi polusi tanah karena tidak menggunakan pestisida atau herbisida yang terserap tanah	Menyerap nutrisi alami dari tanah, namun berpotensi mencemari tanah jika pemupukan dan pestisida tidak terkendali
Efisiensi Nutrisi	Nutrisi langsung diserap oleh akar, sehingga lebih efisien dan sedikit yang terbuang	Efisiensi lebih rendah karena nutrisi bisa terserap, melebar atau hilang di dalam tanah.
Kebutuhan ruang	Dapat memaksimalkan ruang, bisa ditumpuk secara vertikal, cocok untuk daerah perkotaan	Membutuhkan ruang yang luas, sulit dilakukan implementasi secara vertikal
Produktivitas	Tanaman tumbuh lebih cepat dan stabil sepanjang tahun terlepas dari musim	Tergantung musim dan cuaca, yang mempengaruhi waktu dan hasil panen

2.2.1.3.2 Kekurangan Hidroponik NFT

Sistem hidroponik semakin populer dalam beberapa tahun terakhir sebagai metode untuk menghasilkan tanaman berkualitas tinggi secara lebih berkelanjutan dan efisien dalam penggunaan sumber daya. Namun, metode budidaya yang inovatif ini tidak terlepas dari kelemahan yang harus diatasi untuk memastikan keberlanjutan jangka panjangnya (Thomas & Bose, 2023). Salah satu teknik hidroponik yang paling banyak digunakan adalah NFT, yang memiliki tantangan tersendiri dalam penerapannya.

Nutrient Film Technique adalah sistem hidroponik di mana akar tanaman direndam dalam aliran dangkal air yang kaya nutrisi. Metode ini menawarkan sejumlah keunggulan, seperti efisiensi tinggi dalam pengelolaan air dan nutrisi. Namun, metode ini juga memiliki beberapa kelemahan. Salah satu kelemahan utama NFT adalah tingginya biaya modal dan energi yang diperlukan untuk implementasi. Infrastruktur dan peralatan canggih yang diperlukan untuk menjaga kondisi lingkungan optimal, seperti suhu, kelembapan, dan konsentrasi nutrisi, dapat menjadi sangat mahal. Hal ini membatasi adopsi teknologi ini, terutama di kalangan petani dengan sumber daya terbatas (Jensen & Collins, 1985).

Selain itu, sistem NFT sangat sensitif terhadap perubahan kondisi lingkungan. Gangguan pada aliran air atau ketidakseimbangan nutrisi dapat berdampak serius pada kesehatan dan produktivitas tanaman. Sensitivitas ini menjadikan sistem NFT sulit dikelola, terutama bagi petani yang kurang berpengalaman, dan dapat menyebabkan kerugian tanaman yang signifikan jika tidak segera ditangani.

Kelemahan lainnya dari sistem NFT adalah meningkatnya risiko wabah penyakit dan serangan hama. Sifat sistem hidroponik yang terisolasi dapat menciptakan lingkungan yang lebih rentan terhadap penyebaran patogen dan hama, karena tanaman tidak mendapatkan perlindungan alami seperti yang ditemukan pada sistem berbasis tanah (Thomas & Bose, 2023).

Meskipun menghadapi tantangan-tantangan ini, teknik NFT dan sistem hidroponik lainnya terus dieksplorasi dan disempurnakan. Peneliti dan petani bekerja untuk mengatasi kelemahan-kelemahan tersebut dan meningkatkan efisiensi serta keberlanjutan metode budidaya ini. Misalnya, penelitian terkini

mengeksplorasi penggunaan varietas tanaman yang lebih tahan terhadap penyakit serta sistem *monitoring* otomatis untuk memantau nutrisi dan kondisi lingkungan secara *real-time* (Resh, 2022). Upaya ini menunjukkan potensi besar hidroponik sebagai solusi pertanian modern yang berkelanjutan.

2.2.1.4 Hidroponik Otomatis

Sistem hidroponik otomatis telah menjadi subjek penelitian yang luas karena keunggulannya dalam memberikan *throughput* yang tinggi, efisiensi operasional, dan mengurangi kebutuhan akan tenaga kerja manual. Otomatisasi memungkinkan berbagai tugas dalam sistem hidroponik, seperti pengukuran, perbandingan, perhitungan, dan koreksi, untuk dilakukan oleh instrumen secara berulang. Proses ini mendukung kelancaran operasi, pengendalian keamanan, dan mutu produk secara konsisten (Affan, 2004). Selain itu, otomatisasi dinilai lebih efektif dibandingkan pengelolaan manual karena komputer dapat menjalankan tugas-tugas tersebut dengan lebih cepat dan konsisten (Modu *et al.*, 2020).

Dalam konteks hidroponik, otomatisasi tidak hanya mendukung efisiensi tetapi juga optimalisasi pertumbuhan tanaman. Untuk mencapai otomatisasi yang efektif, diperlukan model yang dapat memahami kondisi variabel lingkungan dan tanaman, seperti suhu, kelembapan, pH, TDS, dan EC. Model ini akan memproses data *input* menjadi informasi yang relevan untuk mendukung pengelolaan hidroponik melalui simulasi otomatis yang disajikan. Dengan pendekatan ini, otomatisasi hidroponik dapat meningkatkan produktivitas dan kualitas hasil tanaman secara signifikan.

2.2.2 Machine Learning

Machine learning, sebagai cabang kecerdasan buatan, memungkinkan komputer untuk belajar dari data, mengenali pola, dan membuat keputusan atau prediksi tanpa harus diprogram secara eksplisit. Dalam berbagai bidang, termasuk pertanian, *machine learning* telah menjadi alat yang sangat efektif untuk mengoptimalkan proses dan meningkatkan hasil melalui algoritma prediktif. Di sektor pertanian, *machine learning* secara luas diterapkan untuk memprediksi hasil panen, mengelola sumber daya, dan memantau kondisi lingkungan yang memengaruhi pertumbuhan tanaman. Penggunaan *machine learning* dalam sistem hidroponik membantu petani

membuat keputusan yang lebih cepat dan akurat berdasarkan data *real-time* yang dikumpulkan dari sensor di lapangan (Folorunso & et al., 2023).

Salah satu pendekatan umum dalam *machine learning* adalah *supervised learning*, di mana model dilatih menggunakan data berlabel untuk memprediksi hasil atau keputusan tertentu. Dalam hidroponik, *supervised learning* sering digunakan untuk memprediksi variabel yang memengaruhi pertumbuhan tanaman, seperti suhu, kelembapan, tingkat pH, dan konsentrasi nutrisi. Algoritma yang sering digunakan dalam konteks ini meliputi *Random Forest*, *Decision Trees*, *Support Vector Machines* (SVM), dan *Neural Networks*. Algoritma-algoritma ini membantu petani mengelola sumber daya secara efisien dan meningkatkan produktivitas tanaman (Hamiane *et al.*, 2024). Sebagai contoh, *Random Forest* sangat berguna untuk memprediksi hubungan non-linear antara variabel lingkungan yang memengaruhi hasil panen dalam sistem hidroponik (Cedric & et al., 2022).

Aplikasi *machine learning* dalam pertanian, khususnya hidroponik, juga melibatkan model seperti *Long Short-Term Memory* (LSTM) dan ARIMA yang dirancang untuk menangani data deret waktu (*time-series*). LSTM, sebagai jenis *neural network*, digunakan untuk memproses data berurutan dan sangat cocok untuk memprediksi variabel terkait waktu seperti suhu atau kelembapan dalam sistem hidroponik (Cho *et al.*, 2019). Di sisi lain, ARIMA digunakan untuk menganalisis data musiman dan membuat prediksi jangka panjang berdasarkan tren historis. Kedua model ini sangat penting dalam mengoptimalkan sistem hidroponik dengan menjaga kondisi ideal untuk pertumbuhan tanaman (van Klompenburg *et al.*, 2020).

Dalam sistem hidroponik berbasis IoT, sensor mengumpulkan data *real-time* yang kemudian diproses oleh algoritma *machine learning*. Teknologi ini memungkinkan simulasi otomatis yang menyesuaikan parameter utama pertumbuhan tanaman, seperti pasokan nutrisi dan pengendalian suhu. Model ini memberikan rekomendasi prediktif kepada petani, memungkinkan mereka untuk mengelola sistem hidroponik secara efisien tanpa perlu intervensi manual secara terus-menerus (Rahman *et al.*, 2024). Secara keseluruhan, integrasi *machine*

learning telah merevolusi pengelolaan hidroponik, mulai dari pemantauan variabel lingkungan hingga otomatisasi proses produksi (Folorunso & *et al.*, 2023).

Selain itu, *machine learning* meningkatkan akurasi prediksi hasil panen dengan menganalisis variabel lingkungan yang memengaruhi pertumbuhan tanaman. Sebagai contoh, model regresi dapat memprediksi perilaku pertumbuhan tanaman dalam sistem hidroponik berdasarkan data historis, memungkinkan petani untuk membuat keputusan yang lebih terinformasi terkait manajemen tanaman. Dengan adopsi teknologi *machine learning* yang semakin luas dalam pertanian, model prediktif ini memiliki potensi signifikan untuk meningkatkan efisiensi dan produktivitas di lingkungan pertanian modern (Chopra & Khurana, 2023; Folorunso & *et al.*, 2023).

2.2.2.1 Model ARIMA

Model *Autoregressive Integrated Moving Average* (ARIMA) adalah metode statistik yang banyak digunakan untuk menganalisis dan memprediksi data deret waktu (*time-series*). Model ini sangat populer karena kemampuannya untuk memproses data dengan pola musiman, tren, dan komponen stokastik, yang sering ditemukan dalam data set terkait peramalan harga, cuaca, serta, dalam konteks ini, pertanian dan hidroponik. ARIMA sangat efektif dalam menangani data non-stasioner, di mana distribusi data berubah seiring waktu—pola yang sering terjadi pada variabel lingkungan seperti suhu, kelembapan, dan kadar nutrisi dalam sistem hidroponik (dos Santos de Jesus & da Silva Gomes, 2023).

Model ARIMA terdiri dari tiga komponen utama: *Autoregressive* (AR), *Integrated* (I), dan *Moving Average* (MA). Komponen AR memprediksi nilai masa depan berdasarkan nilai masa lalu dari variabel tersebut. Komponen I (Integrated) digunakan untuk mengubah data non-stasioner menjadi data stasioner dengan menghitung perbedaan antara data secara berurutan. Komponen MA (*Moving Average*) memprediksi nilai masa depan berdasarkan kesalahan (*residual*) dari prediksi sebelumnya (Angelo *et al.*, 2023). Model ARIMA biasanya ditulis sebagai $ARIMA(p, d, q)$, di mana p adalah urutan dari komponen *autoregressive*, d adalah tingkat *differencing* untuk membuat data menjadi stasioner, q adalah urutan dari

komponen *moving average* (Elseidi, 2024). Rumus umum ARIMA dapat dinyatakan sebagai berikut:

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (2.1)$$

Dalam persamaan ini, Y_t mewakili nilai aktual pada waktu t , c adalah konstanta, ϕ_i adalah koefisien AR, θ_j adalah koefisien MA, dan ϵ_t adalah *residual error* pada waktu t . Model ARIMA berfokus pada dua aspek utama dari data deret waktu: tren dan *residual error*. Dalam konteks ini, tren yang diamati dari data masa lalu digunakan untuk memprediksi nilai masa depan, sementara pengaruh fluktuasi acak diminimalkan (Faruk, 2010).

Model ARIMA dirancang untuk menganalisis dan memprediksi data deret waktu dengan memperhitungkan hubungan masa lalu dan pola residual. Model ini memiliki tiga komponen utama: *Autoregressive* (AR), *Integrated* (I), dan *Moving Average* (MA).

1. Komponen *Autoregressive* (AR): Komponen AR menggambarkan hubungan antara nilai variabel pada waktu tertentu (Y_t) dengan nilai masa lalunya (Y_{t-1} , Y_{t-2} , ..., Y_{t-p}). Nilai masa depan diprediksi sebagai kombinasi linear dari nilai masa lalu.
2. Komponen *Integrated* (I): Komponen *Integrated* digunakan untuk menangani data non-stasioner, yaitu data di mana rata-rata, variansi, atau pola musiman berubah seiring waktu.
3. Komponen *Moving Average* (MA): Komponen MA memperhitungkan *residual error* (ϵ_t) dari prediksi sebelumnya untuk memperbaiki prediksi masa depan.

Proses ARIMA dimulai dengan memeriksa apakah data bersifat stasioner. Jika tidak, model menerapkan *differencing* untuk menstabilkan data dengan menghitung perbedaan antara nilai-nilai berurutan hingga data menjadi stasioner. Setelah data stasioner, komponen *autoregressive* (AR) memprediksi nilai masa depan berdasarkan data masa lalu, sedangkan komponen *moving average* (MA) memperhitungkan *error* prediksi dari model AR (Jin *et al.*, 2022). Pendekatan ini

memungkinkan ARIMA menangani data deret waktu dengan pola musiman dan tren yang kompleks, sehingga cocok untuk memprediksi variabel lingkungan seperti suhu, kelembapan, dan kadar nutrisi dalam sistem hidroponik (Elseidi, 2024).

Salah satu keunggulan utama model ARIMA adalah kemampuannya menangani data non-stasioner, yang sering menjadi tantangan dalam analisis deret waktu. Namun, ARIMA juga memiliki keterbatasan, seperti fleksibilitas yang lebih rendah dalam mengelola data yang sangat kompleks atau non-linear, di mana model yang lebih canggih seperti *Long Short-Term Memory* (LSTM) atau Neural Networks cenderung memberikan hasil yang lebih baik. Oleh karena itu, dalam beberapa penelitian, ARIMA dikombinasikan dengan model lain untuk meningkatkan akurasi prediksi (dos Santos de Jesus & da Silva Gomes, 2023; Omar *et al.*, 2016).

Dalam bidang pertanian dan sistem hidroponik, ARIMA telah banyak diterapkan untuk memprediksi variabel lingkungan yang memengaruhi pertumbuhan tanaman. Sebagai contoh, Menculini & et al., (2021) membandingkan kinerja ARIMA dengan model *deep learning* dalam peramalan harga pangan dan menemukan bahwa meskipun ARIMA andal untuk prediksi jangka pendek, model *deep learning* lebih cocok untuk menangani data deret waktu yang lebih kompleks. Dalam sistem hidroponik, ARIMA digunakan untuk memprediksi tingkat kelembapan dan suhu, yang sangat penting untuk menjaga lingkungan tumbuh yang stabil. Kombinasi ARIMA dengan data sensor *real-time* memungkinkan petani hidroponik membuat keputusan lebih cepat dan akurat dalam mengelola kondisi lingkungan (Menculini & et al., 2021).

Selain itu, ARIMA sering digunakan untuk meramalkan suhu dan pola cuaca, yang merupakan faktor kunci dalam pertanian dan hidroponik. Sebagai contoh, Elseidi (2024) menggunakan ARIMA untuk memprediksi data suhu frekuensi tinggi dan menggabungkannya dengan model Prophet untuk meningkatkan akurasi. Temuan mereka menunjukkan bahwa model *hybrid* ARIMA-Prophet memberikan prediksi yang lebih akurat dibandingkan dengan ARIMA saja, terutama ketika data menunjukkan pola musiman atau tren yang kuat (Elseidi, 2024).

Dalam studi lain, Pindiga (2022) membandingkan ARIMA dengan model Facebook Prophet dalam memprediksi indeks saham, yang juga menunjukkan pola musiman dan tren serupa dengan data pertanian. Hasilnya menunjukkan bahwa Prophet cenderung lebih unggul dibandingkan ARIMA dalam menangani data musiman yang kompleks, tetapi ARIMA tetap menjadi pilihan yang kuat untuk prediksi jangka pendek atau data set yang lebih sederhana (Pindiga, 2022).

Aplikasi ARIMA dalam bidang pertanian tidak terbatas pada prediksi variabel lingkungan. Model ini juga digunakan untuk meramalkan hasil panen berdasarkan data pertumbuhan historis dan kondisi cuaca. (Kasthuri & Selvakumar, 2021) menggunakan ARIMA yang dikombinasikan dengan *Neural Networks* untuk memprediksi hasil produksi pangan, memberikan prediksi hasil panen yang lebih akurat dalam skenario di mana variabel lingkungan sangat berfluktuasi. Pendekatan *hybrid* ini semakin populer dalam prediksi deret waktu, terutama di sektor pertanian yang sangat bergantung pada data musiman dan kondisi cuaca yang berubah (Kasthuri & Selvakumar, 2021).

Secara keseluruhan, ARIMA adalah model yang kuat dan serbaguna untuk analisis dan peramalan data deret waktu. Dengan kemampuannya memproses data non-stasioner dan menangani tren serta pola musiman, ARIMA sangat cocok digunakan dalam pertanian hidroponik. Namun, untuk dataset yang lebih kompleks atau non-linear, model hybrid atau teknik *deep learning* mungkin memberikan hasil yang lebih baik. Meskipun demikian, ARIMA tetap menjadi salah satu model yang paling banyak digunakan untuk peramalan deret waktu karena kesederhanaan dan keandalannya dalam memproses data yang relatif sederhana (dos Santos de Jesus & da Silva Gomes, 2023; Faruk, 2010).

2.2.2.2 Model Prophet

Model Prophet adalah alat peramalan deret waktu yang dikembangkan oleh Facebook (sekarang Meta) yang dirancang untuk menangani data dengan pola musiman, tren, dan *outlier*. Prophet sering digunakan untuk memprediksi data yang menunjukkan ketidakstabilan dalam tren, seperti perubahan mendadak atau pola musiman yang tidak teratur. Salah satu keunggulan utamanya adalah kemampuannya untuk menangani celah (*missing data*) dan *outlier* secara efektif,

sambil tetap memberikan prediksi yang akurat bahkan dalam kondisi yang berubah dengan cepat (Jin *et al.*, 2022).

Berbeda dengan model deret waktu tradisional seperti ARIMA, Prophet menggunakan pendekatan aditif, di mana komponen tren, musiman, dan efek *outlier* diperlakukan secara terpisah dan kemudian digabungkan untuk menghasilkan peramalan akhir. Model ini mengasumsikan bahwa data deret waktu dapat diuraikan menjadi tiga komponen utama: tren $g(t)$, musiman $s(t)$, dan efek *holiday* atau *event* $h(t)$, dengan tambahan komponen residual atau *noise* ϵt .

Secara matematis, model *Prophet* dapat dijelaskan dengan persamaan berikut (Taylor & Letham, 2018):

$$y(t) = g(t) + s(t) + h(t) + \epsilon t \quad (2.2)$$

- $y(t)$ prediksi nilai pada waktu t ,
- $g(t)$ komponen tren,
- $s(t)$ komponen musiman,
- $h(t)$ merepresentasikan *holidays* atau *special events*,
- dan ϵt adalah *noise*.

Komponen tren ($g(t)$) berfungsi menangkap perubahan jangka panjang dalam data dan dapat direpresentasikan dengan model linear *piecewise* atau *logistic growth*. Linear *piecewise* menggunakan serangkaian segmen linear dengan kemiringan (k) yang dapat berubah pada titik-titik tertentu, sedangkan *logistic growth* digunakan pada data dengan batas kapasitas, seperti populasi maksimum, yang mengikuti pola pertumbuhan eksponensial sebelum akhirnya melambat. Komponen musiman ($s(t)$) memanfaatkan representasi *Fourier* untuk menggambarkan pola berulang, seperti perubahan harian, mingguan, atau tahunan. Komponen ini menjelaskan bagaimana data mengikuti pola yang teratur dalam jangka waktu tertentu. Sementara itu, komponen hari libur ($h(t)$) memungkinkan pengguna menambahkan dampak dari kejadian tertentu, seperti libur nasional, yang signifikan terhadap data.

Prophet menggunakan regresi linear berbasis segmen (*piecewise linear regression*) atau pertumbuhan logistik (*logistic growth*) untuk menangkap tren

dalam data. Salah satu keunggulannya adalah kemampuannya untuk secara otomatis menyesuaikan jumlah titik perubahan (*change points*) dalam tren, memungkinkan model untuk mengakomodasi perubahan arah yang tiba-tiba. Komponen musiman didefinisikan berdasarkan periode tertentu, seperti tahunan, mingguan, atau bulanan, sehingga memungkinkan *Prophet* untuk menangkap pola musiman yang berulang dengan lebih fleksibel. Komponen liburan atau acara memungkinkan inklusi faktor eksternal seperti hari libur atau acara musiman yang dapat memengaruhi data (Liu *et al.*, 2023).

Tahapan dalam penerapan *Prophet* dimulai dengan identifikasi komponen data, di mana data dipecah menjadi tren, musiman, dan dampak hari libur. Selanjutnya, algoritma mendeteksi titik perubahan tren secara otomatis, meskipun pengguna dapat menetapkan titik perubahan secara manual. Kemudian, pola musiman dianalisis menggunakan model *Fourier*, yang memungkinkan identifikasi perubahan musiman dalam berbagai skala waktu. Dampak hari libur atau kejadian khusus juga ditambahkan ke model untuk meningkatkan akurasi prediksi. Dampak hari libur atau kejadian khusus juga ditambahkan ke model untuk meningkatkan akurasi prediksi. Setelah semua komponen ini ditentukan, parameter model seperti kemiringan tren (k) dan koefisien *Fourier* dioptimalkan. Langkah terakhir adalah menghasilkan prediksi berdasarkan kombinasi aditif dari semua komponen.

Sebagai contoh, jika digunakan untuk memprediksi jumlah pengunjung sebuah taman bermain, model *Prophet* akan memanfaatkan komponen tren untuk menunjukkan peningkatan jumlah pengunjung selama musim panas hingga mencapai kapasitas maksimum, komponen musiman untuk mengidentifikasi puncak kunjungan selama akhir pekan atau liburan sekolah, dan komponen hari libur untuk menangkap dampak dari libur nasional seperti Natal. Dengan memisahkan komponen-komponen ini, *Prophet* menghasilkan prediksi yang akurat sekaligus transparan dalam menunjukkan bagaimana tiap faktor memengaruhi hasil akhir.

Prophet sangat intuitif untuk digunakan karena secara otomatis mendeteksi dan menangani data yang hilang dalam deret waktu. Model ini juga dapat menyesuaikan *interval* prediksi dengan memberikan kepercayaan lebih besar pada komponen tren

dan musiman dibandingkan model deret waktu tradisional (Satrio *et al.*, 2021). Selain itu, Prophet memungkinkan pengguna untuk menyesuaikan interval prediksi, memberikan fleksibilitas dalam hal akurasi dan margin kesalahan sesuai dengan kebutuhan pengguna (Mo *et al.*, 2023).

Salah satu aplikasi umum *Prophet* adalah dalam peramalan harga saham dan data ekonomi, yang memerlukan prediksi akurat untuk menangani tren musiman yang berfluktuasi dan sering tidak stabil. Sebagai contoh, Jin *et al.*, (2022) menggunakan Prophet untuk memprediksi harga saham Google, sementara Angelo *et al.*, (2023) membandingkan kinerja Prophet dan ARIMA dalam meramalkan harga Bitcoin. Temuan mereka menunjukkan bahwa Prophet unggul dalam menangkap tren musiman yang kompleks dan menangani data dengan fluktuasi signifikan (Angelo *et al.*, 2023; Jin *et al.*, 2022).

Prophet juga banyak digunakan dalam sektor energi dan cuaca. Sebagai contoh, Elseidi, (2024) mengombinasikan Prophet dengan ARIMA untuk meramalkan data suhu, menghasilkan akurasi yang lebih baik dibandingkan dengan penggunaan ARIMA saja. Pendekatan *hybrid* ini memungkinkan peramalan yang lebih komprehensif dengan menangkap karakteristik berbeda dari data deret waktu (Elseidi, 2024).

Manfaat Prophet terletak pada fleksibilitasnya untuk digunakan pada data yang tidak sempurna, kemampuannya menangkap pola yang kompleks, serta transparansinya dalam memisahkan komponen utama. Model ini sering digunakan untuk peramalan deret waktu di berbagai bidang, seperti logistik, keuangan, hingga pertanian. Penelitian sebelumnya menunjukkan keunggulan Prophet dalam menangani data musiman yang kompleks, seperti yang diungkapkan oleh Taylor dan Letham (2018) serta Nguyen *et al.* (2022), menjadikannya salah satu algoritma paling andal untuk kebutuhan peramalan modern.

Secara keseluruhan, Prophet adalah model yang kuat dan fleksibel untuk meramalkan data deret waktu yang kompleks, terutama ketika data memiliki pola musiman yang tidak teratur. Dengan pendekatan adaptifnya dan fleksibilitas dalam menangani *outlier*, Prophet menawarkan keunggulan signifikan di berbagai sektor, termasuk ekonomi, pertanian, energi, dan peramalan cuaca. Aplikasinya yang luas

meliputi prediksi harga pangan dan hasil panen hingga peramalan permintaan energi dan pola cuaca (Menculini & *et al.*, 2021).

2.2.3 Evaluation Matrix

Dalam mengevaluasi kinerja model prediktif, dua metrik yang paling umum digunakan adalah *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE). Metrik-metrik ini sangat penting dalam menilai seberapa baik model memprediksi data serta memberikan wawasan mengenai tingkat kesalahan antara nilai prediksi dan nilai aktual.

Pemilihan evaluasi kinerja model prediktif merupakan aspek penting dalam penelitian berbasis data deret waktu. Dalam penelitian ini, RMSE dan MAE dipilih sebagai matriks evaluasi utama karena sifatnya yang intuitif, sederhana, dan efektif untuk mengukur kesalahan prediksi antara hasil model dan data aktual. RMSE mengukur akar dari rata-rata kuadrat selisih antara data aktual dan prediksi, memberikan bobot lebih pada kesalahan yang besar, sehingga ideal untuk mendeteksi deviasi signifikan. Hal ini menjadikan RMSE relevan untuk analisis yang membutuhkan sensitivitas terhadap error besar, seperti pengendalian variabel lingkungan dalam sistem hidroponik (Chai & Draxler, 2014). Sementara itu, MAE memberikan rata-rata nilai absolut dari kesalahan prediksi, tanpa memberikan bobot berlebih pada *outlier*, sehingga lebih stabil dalam mengevaluasi performa model secara keseluruhan (Willmott & Matsuura, 2005).

Pendekatan berbasis integral, seperti *Integral of Time-weighted Squared Error* (ITSE), *Integral of Time-weighted Absolute Error* (ITAE), *Integral of Squared Error* (ISE), dan *Integral of Absolute Error* (IAE), meskipun bermanfaat dalam sistem kontrol dinamis, dianggap kurang relevan dalam konteks penelitian ini. ITSE dan ITAE memperhitungkan bobot waktu dalam evaluasi kesalahan, sehingga lebih cocok untuk sistem kendali yang memprioritaskan respons waktu, seperti pengendalian suhu boiler atau pengoptimalan robotika (Åström & Murray, 2008). Sementara itu, ISE dan IAE, yang masing-masing mengukur integral dari kuadrat kesalahan dan integral dari nilai absolut kesalahan, lebih tepat digunakan untuk memastikan minimisasi kesalahan kumulatif dalam jangka panjang, seperti pada

pengendalian motor listrik (Franklin et al., 2015). Karena penelitian ini berfokus pada prediksi deret waktu yang tidak memerlukan bobot waktu, penggunaan matriks ini kurang sesuai.

Sebagai perbandingan, ITAE dan ITSE sering diterapkan pada sistem yang memerlukan optimalisasi waktu dan respons cepat. Kedua indikator ini banyak digunakan dalam sistem kendali otomatis karena memberikan bobot tambahan pada kesalahan yang terjadi dalam waktu lebih lama. ITAE mengutamakan meminimalkan kesalahan yang berlangsung dalam durasi panjang, sementara ITSE lebih sensitif terhadap kesalahan besar yang terjadi di awal proses. Sebagai contoh, ITAE cocok untuk mengevaluasi performa sistem kendali suhu dalam mengurangi waktu pemulihan dari gangguan, sedangkan ITSE sering digunakan dalam sistem robotik atau otomasi yang memerlukan akurasi jangka panjang.

Di sisi lain, ISE dan IAE juga merupakan indikator kinerja yang umum dalam sistem kendali dinamis. ISE menghitung total kuadrat kesalahan, yang sangat sensitif terhadap kesalahan besar, sehingga memberikan penalti yang lebih tinggi pada deviasi signifikan. Sebaliknya, IAE menghitung jumlah absolut dari kesalahan, yang memberikan gambaran lebih stabil terhadap rata-rata kesalahan tanpa memperhitungkan besar kecilnya deviasi secara eksponensial. ISE sering digunakan ketika tujuan utama adalah memastikan bahwa kesalahan besar harus diminimalkan, sedangkan IAE digunakan jika kestabilan keseluruhan sistem menjadi prioritas.

Namun, dalam penelitian ini, pendekatan berbasis integral seperti ITSE, ITAE, ISE, dan IAE tidak diterapkan karena fokus utama tidak berada pada sistem kendali dinamis yang memerlukan respons waktu atau optimasi *real-time*. Penelitian ini berkaitan dengan prediksi pertumbuhan hidroponik berbasis model *machine learning* yang lebih membutuhkan akurasi prediksi daripada respons waktu atau kestabilan sistem secara langsung. Oleh karena itu, metrik yang digunakan harus mampu mengevaluasi performa model prediktif secara akurat tanpa mempertimbangkan faktor waktu.

Keputusan untuk menggunakan RMSE dan MAE didukung oleh berbagai literatur yang menunjukkan keunggulan kedua metrik ini dalam mengevaluasi

model prediktif, terutama dalam konteks data deret waktu. Menurut Chai dan Draxler (2014), RMSE dan MAE memberikan perspektif evaluasi yang saling melengkapi: RMSE sangat sensitif terhadap error besar karena menghitung akar kuadrat dari rata-rata kuadrat kesalahan, sedangkan MAE memberikan gambaran rata-rata yang lebih representatif terhadap kesalahan keseluruhan karena menghitung rata-rata absolut dari selisih antara nilai prediksi dan nilai aktual.

Dalam sistem hidroponik, fluktuasi nutrisi atau suhu dapat menyebabkan kesalahan besar yang memengaruhi pertumbuhan tanaman. RMSE menjadi sangat berguna dalam mendeteksi dan mengevaluasi kesalahan besar tersebut, memberikan informasi penting terkait potensi penyimpangan ekstrem yang dapat merugikan. Sementara itu, MAE memberikan gambaran yang lebih stabil mengenai performa keseluruhan model dengan menghitung rata-rata error tanpa memperbesar pengaruh kesalahan ekstrem. Kombinasi kedua metrik ini memungkinkan analisis yang lebih komprehensif dan akurat terhadap performa model prediksi hidroponik.

Dengan demikian, penggunaan RMSE dan MAE lebih sesuai dalam penelitian ini karena keduanya menawarkan pendekatan yang lebih relevan untuk mengevaluasi performa model prediktif. Pendekatan berbasis integral seperti ITAE, ITSE, ISE, dan IAE cenderung lebih cocok untuk sistem kendali dinamis yang memerlukan optimasi waktu dan respons real-time, yang bukan menjadi fokus utama dalam penelitian ini. Oleh karena itu, RMSE dan MAE dipilih sebagai metrik evaluasi utama untuk memastikan akurasi dan keandalan prediksi dalam sistem hidroponik berbasis *machine learning*.

2.2.3.1 Root Mean Square Error (RMSE)

RMSE adalah metrik umum lainnya untuk mengevaluasi akurasi model prediktif. RMSE menghitung akar kuadrat dari rata-rata perbedaan kuadrat antara nilai prediksi dan nilai aktual. Berbeda dengan MAE yang berfokus pada perbedaan absolut, RMSE memberikan bobot lebih besar pada kesalahan yang lebih besar karena kesalahan tersebut dikuadratkan sebelum dirata-ratakan.

Rumus umum untuk RMSE adalah sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.3)$$

- y_i merupakan nilai asli,
- $\hat{y}_i$ merupakan nilai yang diprediksi,
- n merupakan jumlah data.

RMSE sangat berguna ketika kesalahan yang lebih besar dianggap lebih tidak diinginkan dibandingkan dengan kesalahan yang lebih kecil, karena RMSE memberikan penalti lebih besar pada kesalahan besar. Dalam sebuah studi oleh Hamiane *et al.* (2024), RMSE digunakan untuk mengukur akurasi model *hybrid* LSTM, ARIMA, dan Prophet dalam memprediksi GDP masa depan, menunjukkan bahwa RMSE sangat efektif dalam mengidentifikasi perbedaan signifikan antara prediksi dan data aktual pada periode tertentu. Demikian pula, Mokhtar *et al.*, (2022) menerapkan RMSE untuk memprediksi hasil panen hidroponik, mengungkapkan bahwa RMSE lebih sensitif terhadap *outlier* dibandingkan dengan MAE, sehingga menjadikannya metrik evaluasi yang penting ketika berhadapan dengan data yang sangat bervariasi (Hamiane *et al.*, 2024; Mokhtar *et al.*, 2022).

Penggunaan RMSE dan MAE secara bersamaan memberikan gambaran yang lebih komprehensif tentang kinerja model. Sementara MAE memberikan gambaran umum tentang rata-rata kesalahan, RMSE menekankan kesalahan yang lebih besar, yang dapat menjadi penting dalam aplikasi di mana meminimalkan kesalahan ekstrem sangat diperlukan. Oleh karena itu, dalam banyak studi, kedua metrik ini sering digunakan bersamaan untuk mengevaluasi model prediktif deret waktu, termasuk dalam aplikasi pertanian hidroponik dan prediksi energi (Folorunso & *et al.*, 2023).

2.2.3.2 Mean Absolute Error (MAE)

MAE mengukur rata-rata perbedaan absolut antara nilai prediksi dan nilai aktual. MAE memberikan gambaran umum tentang seberapa besar kesalahan yang dibuat model secara rata-rata, tanpa memperhatikan apakah kesalahan tersebut bernilai positif atau negatif, karena semua kesalahan diperlakukan sama. Rumus umum untuk menghitung MAE adalah sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.4)$$

- y_i merupakan nilai asli,
- $\hat{y}_i$ merupakan nilai yang diprediksi,
- n merupakan jumlah data.

MAE memberikan hasil yang intuitif dan mudah dipahami karena menunjukkan besarnya kesalahan dalam satuan yang sama dengan data asli. Sebagai contoh, dalam sebuah studi oleh Kenyi & Yamamoto (2024), MAE digunakan untuk mengevaluasi kinerja model SARIMA-Prophet dalam memprediksi aliran air, dan hasilnya menunjukkan bahwa MAE membantu mengidentifikasi tingkat rata-rata kesalahan prediksi pada setiap titik waktu (Kenyi & Yamamoto, 2024). Studi lain oleh Angelo *et al.*, (2023) menggunakan MAE untuk membandingkan kinerja model ARIMA dan Prophet dalam memprediksi harga Bitcoin, menunjukkan bagaimana MAE dapat mengukur akurasi prediksi deret waktu pada dataset yang kompleks (Angelo *et al.*, 2023).

2.2.3.3 Teknik Validasi Model

Validasi model adalah langkah esensial dalam pembelajaran mesin untuk memastikan bahwa model dapat melakukan generalisasi pada data baru yang belum pernah dilihat sebelumnya. Proses ini bertujuan untuk menghindari masalah seperti *overfitting* atau *underfitting*. Teknik validasi model yang sering digunakan mencakup validasi sederhana dengan data uji, *K-Fold Cross-Validation* (KCV), dan variasinya, seperti *Virtual K-Fold Cross Validation* (VKCV) dan *Multiple Predicting K-Fold Cross Validation* (MPCV). Teknik-teknik ini tidak hanya

membantu dalam mengevaluasi kinerja model tetapi juga mendukung proses pemilihan model terbaik (Bergmeir et al., 2018). Pendekatan validasi sederhana melibatkan pembagian dataset menjadi dua subset utama: data latih (*training data*) dan data uji (*testing data*). Model dilatih menggunakan data latih, kemudian kinerjanya dievaluasi berdasarkan performa pada data uji. Hasil validasi biasanya dinyatakan dalam bentuk metrik seperti *Mean Squared Error* (MSE), RMSE, atau *Accuracy*. Namun, teknik ini dapat menghasilkan bias yang tinggi terutama pada dataset kecil karena performa model sangat bergantung pada cara data dibagi (Rodríguez et al., 2010). *K-Fold Cross Validation* (KCV) adalah metode validasi yang lebih komprehensif dibandingkan dengan validasi sederhana. Teknik ini membagi dataset menjadi k subset (*folds*), di mana setiap lipatan secara bergantian digunakan sebagai data uji sementara sisanya sebagai data latih. Hasil evaluasi dari setiap lipatan dirata-rata untuk menghasilkan estimasi kinerja model secara keseluruhan.

Rumus untuk menghitung *Mean Squared Error* (MSE) dalam KCV adalah:

$$MSE_{KCV} = \frac{1}{K} \sum_{i=1}^K Error_i \quad 2.5$$

Di mana $Error_i$ adalah kesalahan pada lipatan validasi ke- i . Teknik ini mengurangi bias estimasi dan memberikan hasil yang lebih stabil dibandingkan pembagian data tunggal.

Pemilihan nilai k penting untuk mengontrol keseimbangan antara bias dan varians. Nilai k yang kecil, seperti 5 atau 10, biasanya digunakan untuk memberikan estimasi yang stabil. Namun, pada data deret waktu, penggunaan KCV standar dapat melanggar asumsi independensi antar data, sehingga diperlukan adaptasi seperti *time-series split* atau *block cross-validation* (Bergmeir et al., 2018).

Virtual K-Fold Cross Validation (VKCV) adalah variasi dari KCV yang dirancang untuk mengurangi kebutuhan komputasi. Teknik ini memanfaatkan proyeksi matriks untuk mengestimasi kesalahan prediksi tanpa perlu melatih ulang

model pada setiap lipatan. VKCV sangat efisien dalam konteks model linier. Rumus estimasi kesalahan pada VKCV adalah:

$$\epsilon_{VKCV} = \frac{1}{N} \sum_{i=1}^N \left(\frac{r_i}{1 - h_{ii}} \right)^2 \quad 2.6$$

Di mana r_i adalah residu untuk data ke- i dan h_{ii} adalah elemen diagonal dari matriks proyeksi (*hat matrix*). Metode ini memberikan estimasi yang cepat, meskipun lebih cocok untuk model linier dibandingkan model non-linier (Alippi & Roveri, 2010).

MPCV (*Multiple Predicting K-Fold Cross Validation*) adalah metode yang memodifikasi KCV dengan mengalokasikan lebih banyak lipatan untuk validasi daripada pelatihan. Dalam metode ini, $k-1$ lipatan digunakan untuk validasi dan hanya satu lipatan untuk pelatihan. Setiap pengamatan menghasilkan beberapa nilai prediksi yang dirata-rata untuk menghasilkan prediksi akhir. Teknik ini secara signifikan mengurangi varians validasi dan meningkatkan stabilitas hasil (Jung, 2018). Rumus untuk menghitung *Mean Squared Error* (MSE) dalam MPCV adalah:

$$MSE_{MPCV} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad 2.7$$

Di mana y_i adalah nilai aktual, dan $\hat{y}_i$ adalah rata-rata dari prediksi $k-1$ lipatan untuk setiap pengamatan. Validasi model pada data deret waktu memerlukan perhatian khusus karena adanya ketergantungan serial antar data. *Block Cross-Validation* atau *h-block cross-validation* sering digunakan untuk menghindari kebocoran informasi antar subset pelatihan dan validasi. Teknik ini membagi dataset dengan memperhatikan korelasi antar data untuk memastikan hasil validasi yang sah (Bergmeir et al., 2018). Teknik validasi model seperti KCV, VKCV, dan MPCV memberikan kerangka kerja yang solid untuk mengevaluasi kinerja model secara adil dan akurat. Pemilihan teknik yang sesuai harus mempertimbangkan karakteristik data dan kebutuhan komputasi. Dalam konteks data deret waktu,

modifikasi teknik validasi menjadi sangat penting untuk memastikan hasil yang dapat digeneralisasi.

2.2.4 Correlation Matrix

Correlation matrix adalah tabel berbentuk matriks yang menunjukkan nilai hubungan korelasi antar variabel dalam sebuah dataset. Setiap elemen dalam matriks ini merepresentasikan koefisien korelasi antara dua variabel, yang nilainya berada pada rentang -1 hingga 1. Nilai +1 menunjukkan hubungan positif sempurna, -1 menunjukkan hubungan negatif sempurna, sedangkan nilai 0 menunjukkan tidak adanya hubungan linear antara variabel tersebut. Menurut Cohen *et al.* (2003), *correlation matrix* merupakan alat analisis yang membantu menggambarkan pola hubungan antar variabel, terutama dalam analisis eksplorasi data. Diagonal utama matriks ini selalu memiliki nilai 1, karena setiap variabel memiliki korelasi sempurna dengan dirinya sendiri. Dalam analisis data, matriks ini dapat dihitung menggunakan metode korelasi seperti *Pearson*, *Spearman*, atau *Kendall*. Korelasi *Pearson* digunakan untuk mengukur kekuatan hubungan linear antar variabel, sedangkan metode *Spearman* dan *Kendall* cocok untuk hubungan monotonik atau yang tidak sepenuhnya linear (Rumsey, 2016).

Cara kerja dari *correlation matrix* adalah menghitung nilai koefisien korelasi antar pasangan variabel menggunakan metode statistik, seperti korelasi *Pearson*, *Spearman*, atau *Kendall*. Korelasi *Pearson* mengukur kekuatan hubungan linear antar variabel, sedangkan *Spearman* dan *Kendall* digunakan untuk mengukur hubungan monotonik atau hubungan nonlinear yang bersifat *rank-based* (Rumsey, 2016). Dalam matriks, diagonal utama (diagonal dari kiri atas ke kanan bawah) selalu memiliki nilai 1, karena setiap variabel memiliki korelasi sempurna dengan dirinya sendiri.

Koefisien korelasi *Pearson* r antara dua variabel X dan Y dihitung dengan rumus berikut:

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 * \sum (Y_i - \bar{Y})^2}} \quad 2.8$$

Di mana:

- X_i dan Y_i adalah nilai observasi dari variabel X dan Y,
- $\bar{X}$ dan $\bar{Y}$ adalah rata-rata nilai dari X dan Y,
- r adalah nilai koefisien korelasi Pearson.

Manfaat utama dari *correlation matrix* adalah untuk mempermudah analisis hubungan antar variabel. Matriks ini digunakan untuk mengidentifikasi hubungan signifikan antar variabel, mengevaluasi multikolinearitas dalam model *regresi*, serta mengurangi dimensi data melalui teknik seperti analisis komponen utama (*Principal Component Analysis*). McCarthy (2018) menekankan bahwa *correlation matrix* dapat membantu mendeteksi variabel mana yang memiliki hubungan kuat sehingga menjadi dasar dalam pemodelan prediktif. Gambar 2.3 adalah contoh visualisasi lain dari *correlation matrix* menggunakan *dataset* sederhana. Matriks ini menggunakan gaya warna viridis untuk memberikan kontras yang lebih tinggi. Setiap nilai korelasi ditampilkan secara langsung pada sel, dan gradasi warna menggambarkan hubungan positif atau negatif antar variabel.



Gambar 2.3 *Correlation Matrix*

Membaca correlation matrix melibatkan interpretasi nilai koefisien korelasi:

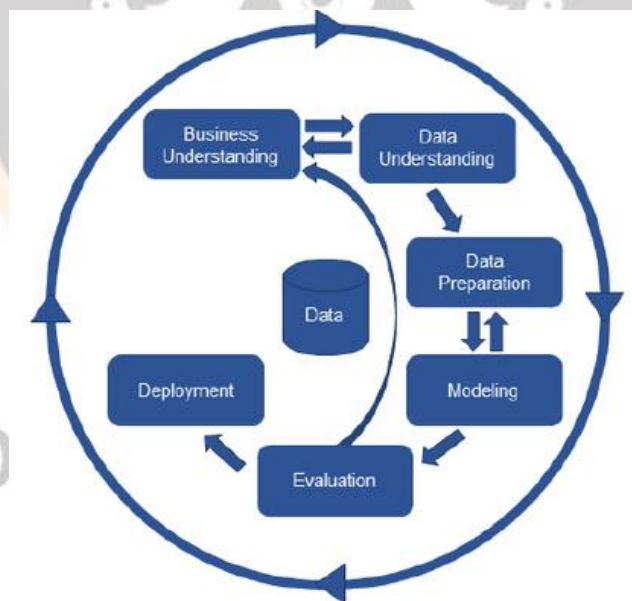
- $r > 0.7$: Hubungan positif yang kuat.
- $0.3 \leq r \leq 0.7$: Hubungan positif sedang.
- $0 \leq r < 0.3$: Hubungan positif lemah.

- $r = 0$: Tidak ada hubungan.
- $r < 0$: Hubungan negatif (arah hubungan berlawanan).

Dengan demikian, *correlation matrix* tidak hanya membantu memahami hubungan antar variabel, tetapi juga menjadi alat penting dalam analisis statistik yang lebih kompleks. Cohen et al. (2003), Rumsey (2016), dan McCarthy (2018) sepakat bahwa penggunaan *correlation matrix* memberikan wawasan mendalam untuk analisis data, terutama dalam mengidentifikasi pola hubungan yang signifikan.

2.2.5 Metodologi CRISP-DM

Metodologi CRISP-DM adalah kerangka kerja yang banyak diadopsi untuk menyusun dan melaksanakan proyek-proyek *data science*. Dikembangkan pada akhir 1990-an, CRISP-DM (yang merupakan singkatan dari *Cross-Industry Standard Process for Data Mining*) menyediakan proses komprehensif dan iteratif untuk pemecahan masalah berbasis data (Naparin & Saad, 2017). Gambar 2.4 memperlihatkan siklus metode CRISP-DM, inti dari CRISP-DM terdiri dari enam fase kunci: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan (Naparin & Saad, 2017; Nisbet *et al.*, 2018b).



Gambar 2.4 Siklus metode CRISP-DM

(Sumber: Khaqim *et al.*, (2022))

Fase pertama, pemahaman bisnis, berfokus pada penentuan tujuan dan persyaratan proyek dari perspektif bisnis (Naparin & Saad, 2017). Fase ini melibatkan perumusan masalah atau peluang yang ada, menetapkan tujuan yang spesifik dan terukur untuk proyek tersebut, serta mengidentifikasi sumber daya yang tersedia, kendala, dan pemangku kepentingan utama yang akan terlibat (Naparin & Saad, 2017; Nisbet *et al.*, 2018a).

Fase pemahaman data kemudian melibatkan pengumpulan, eksplorasi, dan *profiling* data yang relevan untuk memperoleh wawasan mengenai karakteristik dan kualitasnya (Nisbet *et al.*, 2018b). Ini dapat mencakup tugas seperti mengidentifikasi sumber data, memahami format dan struktur data, serta melakukan pemeriksaan kualitas data awal (Nisbet *et al.*, 2018b, 2018a).

Dengan pemahaman yang baik tentang konteks bisnis dan data yang tersedia, fase berikutnya adalah persiapan data. Fase ini sangat penting karena melibatkan pembersihan, transformasi, dan integrasi data ke dalam format yang sesuai untuk pemodelan (Nisbet *et al.*, 2018b). Tugas umum dalam persiapan data mencakup penanganan nilai yang hilang, mengatasi inkonsistensi data, dan rekayasa fitur untuk membuat variabel baru yang dapat meningkatkan kinerja model (Nisbet *et al.*, 2018b, 2018a).

Fase pemodelan melibatkan pemilihan dan penerapan teknik *data mining* yang sesuai pada data yang telah dipersiapkan, seperti algoritma regresi, klasifikasi, atau *clustering* (Naparin & Saad, 2017). Kinerja model kemudian dievaluasi pada fase berikutnya, di mana hasilnya dinilai berdasarkan tujuan proyek asli dan persyaratan pemangku kepentingan (Naparin & Saad, 2017; Nisbet *et al.*, 2018b).

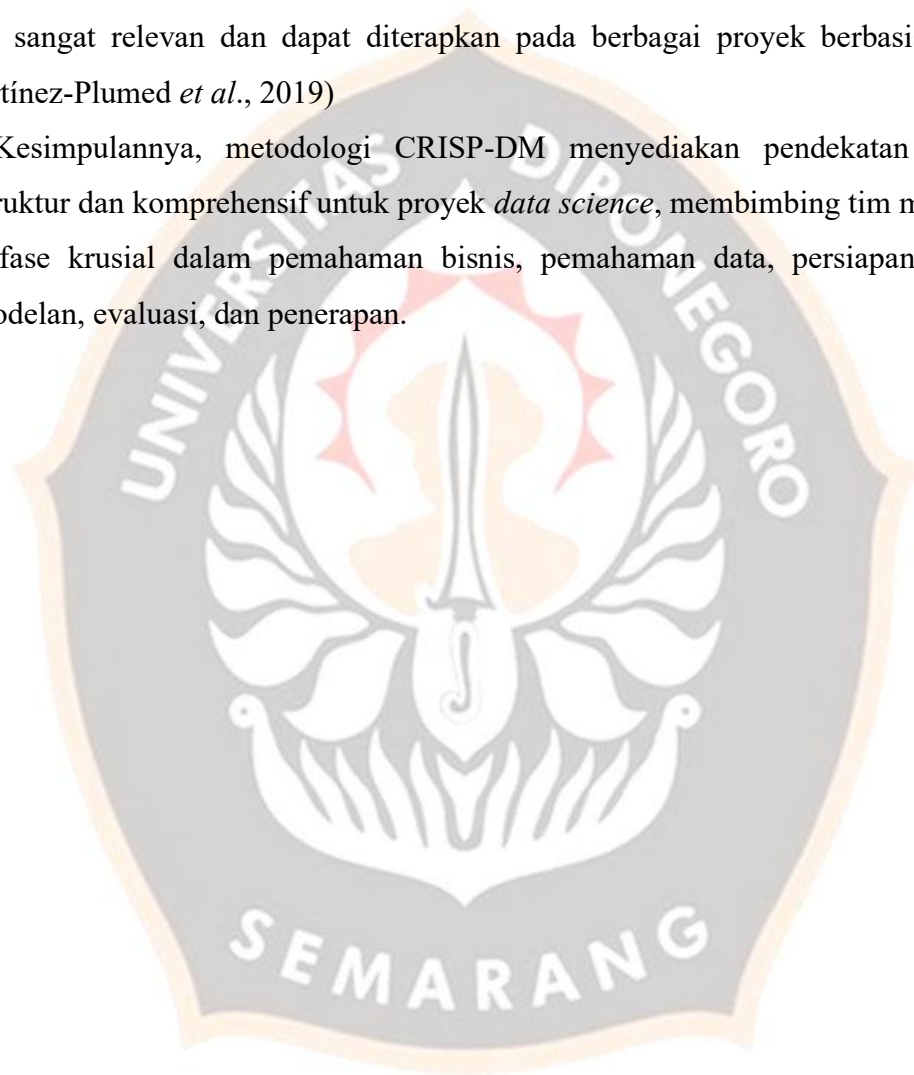
Terakhir, fase penerapan melibatkan integrasi model ke dalam proses bisnis dan pemantauan kinerjanya secara berkelanjutan untuk memastikan bahwa model tersebut terus menghasilkan hasil yang diinginkan (Naparin & Saad, 2017).

Metodologi CRISP-DM bersifat iteratif, yang berarti proses ini dapat kembali ke fase-fase sebelumnya seiring dengan munculnya wawasan, tantangan, atau persyaratan baru (Naparin & Saad, 2017; Nisbet *et al.*, 2018a). Fleksibilitas kerangka CRISP-DM telah berkontribusi pada adopsinya yang luas di komunitas *data science*, menjadikannya alat yang berharga bagi organisasi yang berusaha

untuk memperoleh wawasan yang dapat ditindaklanjuti dari data mereka (Martínez-Plumed *et al.*, 2019).

CRISP-DM telah menjadi metodologi dominan dalam komunitas data mining *and analytics* selama lebih dari dua dekade. Meskipun bidang ini telah berkembang dengan munculnya *data science*, prinsip dan struktur inti dari kerangka CRISP-DM tetap sangat relevan dan dapat diterapkan pada berbagai proyek berbasis data (Martínez-Plumed *et al.*, 2019)

Kesimpulannya, metodologi CRISP-DM menyediakan pendekatan yang terstruktur dan komprehensif untuk proyek *data science*, membimbing tim melalui fase-fase krusial dalam pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan.



SEKOLAH PASCASARJANA