

Salah satu elemen penting dalam pertanian presisi adalah pengelolaan kesuburan tanah, karena tanah yang subur memberikan nutrisi yang sedang untuk mendukung pertumbuhan tanaman. Dalam hal ini, teknologi sensor tanah digunakan untuk mengumpulkan data parameter tanah secara langsung, seperti kelembapan, pH, dan kandungan nutrisi, yang kemudian diolah untuk membantu pengambilan keputusan berbasis data.

Namun, penerapan pertanian presisi masih menghadapi tantangan, terutama dalam mengolah dan menginterpretasikan data dari sensor tanah yang kompleks. Penelitian oleh ini bertujuan menjawab tantangan tersebut dengan mengembangkan sistem klasifikasi kesuburan tanah berbasis *Support Vector Machine (SVM)* untuk menganalisis data sensor secara akurat dan efisien.

2.2.2. Kesuburan Tanah

Kesuburan tanah didefinisikan sebagai kemampuan tanah untuk mendukung pertumbuhan tanaman dengan menyediakan air, udara, dan nutrisi yang diperlukan. Kesuburan tanah dipengaruhi oleh tiga faktor utama: fisik, kimia, dan biologi.

Faktor-faktor yang Mempengaruhi Kesuburan Tanah :

1. Faktor Fisik:
 - a. Tekstur tanah: Tanah yang seimbang antara pasir, debu, dan liat memiliki kemampuan yang lebih baik untuk menahan air dan nutrisi.
 - b. Kelembapan Tanah: Kelembapan tanah merupakan faktor penting yang memengaruhi pelarutan dan pergerakan unsur hara seperti N, P, dan K. Dalam kondisi tanah yang lembap, ketersediaan nutrisi dapat meningkat karena aktivitas mikroorganisme dan proses kimia dalam tanah berlangsung lebih optimal, meskipun tingkat ideal kelembapan dapat bervariasi tergantung jenis tanah dan lingkungan. Tanah yang terlalu kering dapat menghambat pertumbuhan akar dan mikroorganisme (Borowik & Wyszowska, 2016).
 - c. Struktur Tanah: Struktur tanah yang baik memungkinkan akar tanaman tumbuh optimal, mempermudah sirkulasi udara, dan mempertahankan kelembapan tanah.

- d. Suhu Tanah: Suhu tanah memengaruhi proses biologis seperti aktivitas mikroorganisme dan penyerapan nutrisi oleh akar tanaman. Suhu optimal untuk sebagian besar tanaman pertanian adalah 20-30°C. Suhu di bawah 10°C atau di atas 35°C dapat memperlambat aktivitas mikroorganisme dan menurunkan ketersediaan nutrisi (Rahmatika Cahyaningprastiwi et al., 2021).
2. Faktor Kimia:
- a. pH tanah: Tanah yang terlalu asam atau basa dapat mengurangi ketersediaan nutrisi.
 - b. Nutrisi utama (NPK): Nitrogen, Fosfor, dan Kalium adalah nutrisi makro yang sangat penting untuk pertumbuhan tanaman (Gottemukkala et al., 2023).
 - c. Elektrokonduktivitas (EC): Mengindikasikan tingkat ketersediaan nutrisi larut dalam tanah.
3. Faktor Biologi:
- a. Mikroorganisme tanah: Mikroba membantu dalam dekomposisi bahan organik dan pelepasan nutrisi penting.
 - b. Bahan organik: Meningkatkan kemampuan tanah menahan air dan nutrisi.

2.2.3. Klasifikasi Kesuburan Tanah Berdasarkan Kalium Menurut Blesslin Sheeba et al. (2022)

Klasifikasi kesuburan tanah dapat dilakukan dengan menilai kadar unsur hara utama di dalam tanah, salah satunya adalah Kalium (K). Kalium merupakan nutrisi esensial yang berperan dalam berbagai proses fisiologis tanaman, seperti pembukaan stomata, pengangkutan air, serta sintesis protein dan pati. Oleh karena itu, kadar Kalium di dalam tanah dapat dijadikan sebagai indikator utama dalam menilai tingkat kesuburan tanah.

Dalam penelitian yang dilakukan oleh (Blesslin Sheeba et al., 2022), klasifikasi kesuburan berdasarkan kandungan Kalium dibagi menjadi lima kategori

berdasarkan rentang nilai Kalium yang terukur dari hasil sensor atau analisis tanah. Rincian klasifikasi tersebut disajikan dalam Tabel II. 12 berikut :

Tabel II. 12. Klasifikasi Menurut Blesslin Sheeba et al. (2022)

Rating	Kalium (K) (ppm)
Very High	> 300
High	200–300
Medium	151–200
Low	101–150
Very Low	< 100

Klasifikasi ini menjadi referensi utama dalam penelitian ini karena memberikan batasan kuantitatif yang jelas untuk mengevaluasi kesuburan tanah dari parameter Kalium secara langsung. Klasifikasi ini juga digunakan dalam berbagai penelitian modern yang mengintegrasikan teknologi *Internet of Things (IoT)* dan *machine learning* untuk pemantauan dan prediksi kesuburan tanah. Hal ini selaras dengan tujuan penelitian ini yang mengembangkan sistem klasifikasi kesuburan tanah berbasis sensor untuk pertanian presisi skala kecil.

2.2.4. *Internet of Things (IoT)*

Internet of Things (IoT) adalah jaringan perangkat fisik yang terhubung ke internet, yang dapat mengumpulkan, mengirim, dan menerima data tanpa intervensi manusia. *IoT* memungkinkan konektivitas antara berbagai perangkat melalui jaringan, memungkinkan komunikasi dan pertukaran data . *IoT* memanfaatkan teknologi sensor dan perangkat lunak untuk mengumpulkan informasi dari lingkungan dan memproses data tersebut untuk analisis lebih lanjut(Friha dkk., 2021).

Dalam bidang pertanian presisi, *IoT* menjadi salah satu teknologi kunci untuk memantau dan mengelola berbagai aspek pertanian secara otomatis dan *real-time*(G. K. Kumar et al., 2024). Dengan *IoT*, petani dapat mengumpulkan data dari berbagai perangkat yang tersebar di lahan pertanian, menganalisis data tersebut, dan mengambil keputusan berbasis informasi.

Keunggulan *IoT* dalam Pertanian:

1. Pemantauan *Real-time*: Sistem *IoT* dapat memberikan informasi terkini tentang kondisi lahan, seperti kelembapan tanah, suhu udara, dan kadar nutrisi, yang memengaruhi pertumbuhan tanaman.
2. Efisiensi Operasional: Dengan data yang dihasilkan oleh *IoT*, petani dapat mengoptimalkan penggunaan air, pupuk, dan pestisida, sehingga mengurangi pemborosan.
3. Automasi Sistem Pertanian: *IoT* memungkinkan pengendalian jarak jauh, seperti mengatur irigasi otomatis berdasarkan data kelembapan tanah.

Pengambilan Keputusan Berbasis Data: *IoT* memberikan data yang relevan dan terukur untuk mendukung pengelolaan lahan yang lebih baik.

2.2.5. Smote (Synthetic Minority Over-sampling Technique)

Ketidakseimbangan kelas (*class imbalance*) merupakan masalah umum dalam klasifikasi, terutama ketika salah satu kelas jauh lebih dominan dibanding kelas lainnya. Hal ini dapat menyebabkan model pembelajaran mesin menjadi lebih condong kelas mayoritas dan mengabaikan kelas minoritas, sehingga performa klasifikasi secara keseluruhan menjadi menurun. Salah satu metode yang populer digunakan untuk mengatasi hal ini adalah *SMOTE* (*Synthetic Minority Over-sampling Technique*).

SMOTE pertama kali diperkenalkan oleh (Chawla et al., 2002) sebagai metode oversampling berbasis interpolasi, di mana data baru dari kelas minoritas dihasilkan secara sintesis berdasarkan tetangga terdekat (*nearest neighbors*) dari instance kelas minoritas. Teknik ini tidak hanya menambahkan jumlah data minoritas, tetapi juga memperkaya variasi pola data dalam ruang fitur, sehingga dapat membantu model untuk mengenali karakteristik kelas minoritas secara lebih baik (Chawla et al., 2002).

Dalam berbagai penelitian, penggunaan *SMOTE* terbukti dapat meningkatkan performa model klasifikasi, terutama bila dikombinasikan dengan algoritma *Support Vector Machine* (*SVM*). *SVM* merupakan metode klasifikasi yang bekerja dengan mencari *hyperplane* optimal untuk memisahkan kelas-kelas dalam ruang berdimensi tinggi. Namun, *SVM* sensitif terhadap

ketidakseimbangan kelas, sehingga penerapan *SMOTE* dapat memberikan distribusi data yang lebih seimbang dan meningkatkan akurasi prediksi.

Sebagai contoh, penelitian oleh (Rahman et al., 2023) menunjukkan bahwa penerapan *SMOTE* sebelum pelatihan model *SVM* mampu meningkatkan akurasi klasifikasi penyakit menjadi 91,4%, dibandingkan dengan 84,7% pada data yang tidak diseimbangkan..

Dengan demikian, penerapan *SMOTE* dalam penelitian ini bertujuan untuk meningkatkan akurasi dan generalisasi model klasifikasi kesuburan tanah berbasis data sensor, terutama ketika jumlah sampel pada masing-masing kelas kesuburan tidak merata.

2.2.6. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma *machine learning* yang digunakan untuk tugas klasifikasi dan regresi. Algoritma ini bekerja dengan mencari *Hyperplane* terbaik yang memisahkan data ke dalam kelas-kelas yang berbeda di ruang dimensi tinggi (G. K. Kumar et al., 2024). Dalam Penelitian oleh ini, *SVM* digunakan untuk mengklasifikasikan kesuburan tanah berdasarkan parameter yang diukur oleh sensor tanah NPK RS485 seperti suhu, kelembapan, pH, Nitrogen (N), Fosfor (P), dan Kalium (K).

Keunggulan utama *SVM* adalah kemampuannya dalam menangani:

1. Data yang kompleks: *SVM* dapat bekerja dengan data yang kompleks, seperti data tanah dengan berbagai parameter.
2. Data *Non-linear*: Dengan memanfaatkan fungsi *kernel*, *SVM* dapat memetakan data *non-linear* ke ruang dimensi lebih tinggi agar dapat dipisahkan dengan lebih mudah.
3. *Overfitting* yang Minim: *SVM* memiliki mekanisme regularisasi yang menjaga generalisasi model pada data baru.

Dalam konteks Penelitian oleh ini, *SVM* bertujuan untuk memisahkan data tanah ke dalam kategori kesuburan, seperti sangat subur, subur, kurang subur, atau tidak subur, berdasarkan pola yang ditemukan dalam data sensor.

Berikut adalah Komponen Utama *SVM* :

1. *Hyperplane*

Hyperplane adalah garis atau bidang yang memisahkan data ke dalam kelas-kelas tertentu. *Hyperplane* terbaik adalah yang memiliki margin maksimum antara titik-titik data dari kedua kelas. Dalam Penelitian oleh ini, *Hyperplane* digunakan untuk memisahkan kategori kesuburan tanah berdasarkan parameter seperti pH dan kandungan NPK.

2. *Kernel Trick*

- a. Untuk menangani data *non-linear*, *SVM* menggunakan fungsi *kernel* yang memetakan data ke ruang dimensi yang lebih tinggi. Fungsi *kernel* yang umum digunakan meliputi *Linear Kernel* Untuk data yang dapat dipisahkan secara *linear*. *Polynomial Kernel*: Untuk hubungan *non-linear* tingkat rendah.
- b. *Radial Basis Function (RBF) Kernel*: Untuk data yang sangat kompleks dan *non-linear*.

3. *Support Vectors*

- a. *Support vectors* adalah data titik yang paling dekat dengan *Hyperplane*. Titik-titik ini menentukan posisi *Hyperplane* dan memengaruhi hasil klasifikasi.
- b. Dalam konteks Penelitian oleh ini, *support vectors* membantu menentukan batasan antara kategori tanah yang subur dan kurang subur.

Pemilihan *hyperparameter* yang tepat sangat penting untuk meningkatkan performa model *SVM*. *Hyperparameter* utama yang perlu diatur meliputi :

1. *C (Regularization Parameter)*

- a. Mengontrol seberapa besar model ingin menghindari mis-klasifikasi. Nilai *C* yang tinggi mencoba untuk memisahkan semua data dengan benar, tetapi berisiko *Overfitting* (model terlalu "menghafal" data latih sehingga performanya buruk pada data baru).

- b. Dalam Penelitian oleh ini, nilai C akan dioptimalkan untuk mencapai keseimbangan antara akurasi model dan generalisasi.

2. Gamma (Parameter *Kernel*)

Gamma menentukan pengaruh setiap titik data dalam fungsi *kernel*. Nilai gamma yang terlalu tinggi dapat menyebabkan *Overfitting* (model terlalu "menghafal" data latih sehingga performanya buruk pada data baru), Nilai gamma yang rendah berarti pengaruh setiap titik data lebih luas, menghasilkan model yang lebih sederhana. Namun, ini dapat menyebabkan *underfitting* (model terlalu sederhana sehingga gagal menangkap pola dalam data).

3. *Kernel*

Memilih *kernel* yang sesuai (*Linear*, *Polynomial*, *RBF*) berdasarkan pola data. Uji coba akan dilakukan untuk menentukan *kernel* yang menghasilkan akurasi terbaik.

Setelah model *SVM* dilatih menggunakan data sensor, performanya akan dievaluasi menggunakan beberapa metrik:

1. *Precision*

- a. *Precision* mengukur proporsi prediksi positif yang benar terhadap seluruh prediksi positif.
- b. Rumus pada (II. 1 :

(II. 1)

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

- c. Dalam konteks Penelitian oleh ini, precision menunjukkan akurasi model dalam mengklasifikasikan tanah ke kategori subur tanpa salah memprediksi tanah yang kurang subur sebagai subur.

2. *Recall (Sensitivity)*

a. *Recall* mengukur proporsi prediksi positif yang benar terhadap seluruh data positif aktual.

b. Rumus pada (II. 2 :

(II. 2)

$$\text{Recall} = \frac{\text{True Positives (TP)}}{(\text{TP})\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

c. *Recall* dalam Penelitian oleh ini mengindikasikan sejauh mana model dapat mendeteksi semua kategori tanah yang subur.

3. *F1-Score*

a. *F1-Score* adalah rata-rata harmonis dari precision dan recall. Metrik ini memberikan gambaran keseluruhan tentang performa model.

b. Rumus pada (II. 3):

(II. 3)

$$\text{F1 - Score} = 2 \cdot \frac{\text{Precssion} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

c. *F1-Score* sangat berguna jika terdapat ketidakseimbangan jumlah data antar kategori kesuburan tanah.

4. *Support*

Support adalah jumlah data aktual yang termasuk dalam setiap kategori. Metrik ini membantu memahami distribusi data dalam evaluasi model.

5. Akurasi

a. Akurasi mengukur persentase prediksi yang benar terhadap keseluruhan data.

b. Rumus pada (II. 4) :

(II. 4)

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data}}$$

Akurasi akan menjadi metrik utama dalam mengevaluasi performa model klasifikasi kesuburan tanah

2.3. Alasan Pemilihan Teori

Pemilihan teori dan konsep dalam penelitian ini didasarkan pada relevansi dan kontribusinya terhadap tujuan utama penelitian, yaitu mengembangkan sistem klasifikasi kesuburan tanah menggunakan algoritma *Support Vector Machine (SVM)* dan data sensor berbasis *IoT*. Berikut adalah alasan pemilihan teori utama:

1. Jurnal (Blesslin Sheeba et al., 2022)
 - a. Aturan klasifikasi dari jurnal (Blesslin Sheeba et al., 2022) dipilih sebagai dasar pelabelan data karena memberikan batasan kuantitatif yang jelas dalam mengklasifikasikan kesuburan tanah, khususnya berdasarkan nilai Kalium (K). Rentang nilai Kalium dibagi menjadi lima tingkat kesuburan, yang kemudian disesuaikan menjadi tiga kelas (rendah, sedang, tinggi) agar lebih sesuai dengan kebutuhan penelitian.
 - b. Validitas pendekatan ini didukung oleh penerapannya dalam penelitian berbasis sensor dan algoritma klasifikasi tanah secara modern, serta adanya acuan numerik yang dapat diadaptasi ke dalam sistem berbasis data sensor. Penelitian tersebut menjadi salah satu referensi ilmiah yang relevan dan aplikatif untuk konteks klasifikasi kesuburan tanah berbasis teknologi.
2. *Support Vector Machine (SVM)*

- a. *SVM* dipilih sebagai algoritma utama karena kemampuannya untuk menangani data berdimensi tinggi dan kompleksitas hubungan *non-linear* antar parameter tanah.
- b. *SVM* mampu menghasilkan *Hyperplane* terbaik yang memisahkan kelas-kelas data, sehingga cocok untuk tugas klasifikasi seperti mengkategorikan tanah ke dalam kelas Sangat Subur, Subur, Kurang Subur, dan Tidak Subur.
- c. Dengan optimasi *hyperparameter* (*C* dan *Gamma*), *SVM* dapat mencapai keseimbangan antara akurasi model dan kemampuan generalisasi.

3. Sensor *IoT* NPK RS485

- a. Penggunaan teknologi *IoT*, khususnya sensor NPK RS485, memungkinkan pengukuran *real-time* parameter tanah seperti pH, Nitrogen, Fosfor, Kalium, suhu, dan kelembapan. Teknologi ini mendukung kebutuhan pertanian presisi dengan menyediakan data yang akurat dan efisien.
- b. Sensor ini dipilih karena kompatibilitasnya dengan sistem *IoT*, kemudahan implementasi, dan jangkauan pengukuran parameter yang relevan dengan kesuburan tanah.

4. Pertanian Presisi

Konsep pertanian presisi dipilih sebagai kerangka penelitian karena sejalan dengan kebutuhan modern dalam mengoptimalkan hasil pertanian melalui teknologi. Dengan pendekatan ini, pengelolaan tanah dapat dilakukan secara lebih efektif berdasarkan data yang terukur.

5. Validasi dengan Sampel Tanah Nyata

Validasi menggunakan sampel tanah nyata dipilih untuk memastikan hasil klasifikasi model *SVM* sesuai dengan kondisi lapangan. Pendekatan ini memberikan landasan empiris yang kuat bagi hasil penelitian.

Dengan kombinasi teori ini, penelitian diharapkan dapat menghasilkan model klasifikasi yang akurat, efisien, dan relevan dengan kebutuhan pertanian modern di Indonesia.

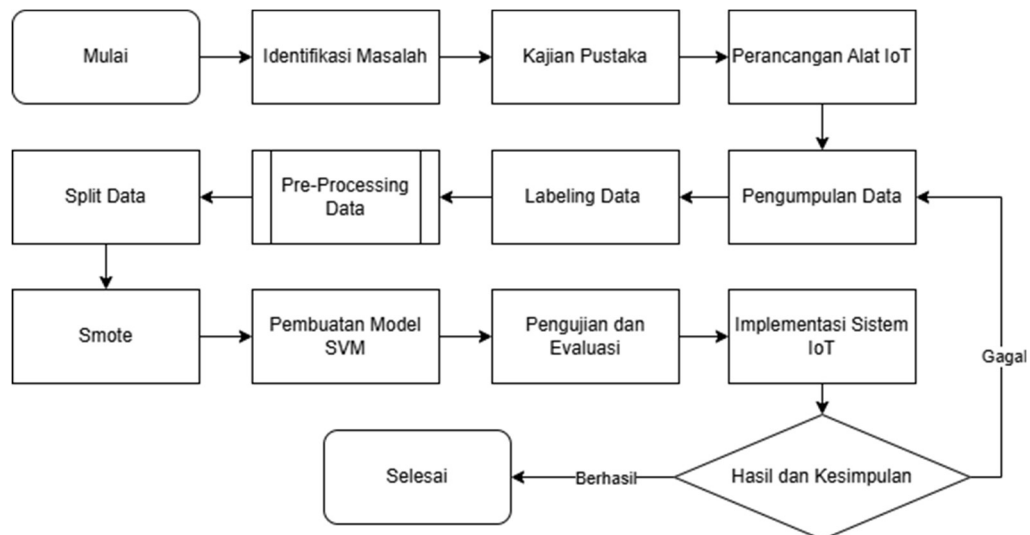
BAB III

METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan dan pendekatan yang digunakan dalam pelaksanaan penelitian, mulai dari perencanaan hingga evaluasi hasil. Metodologi dirancang untuk memastikan bahwa proses penelitian berjalan secara sistematis dan dapat dipertanggungjawabkan. Adapun uraian dalam bab ini meliputi:

3.1. Alur Penelitian

Berikut ini adalah *Flowchart* Penelitian pada Gambar III. 1 *Flowchart* Penelitian sebagai berikut:

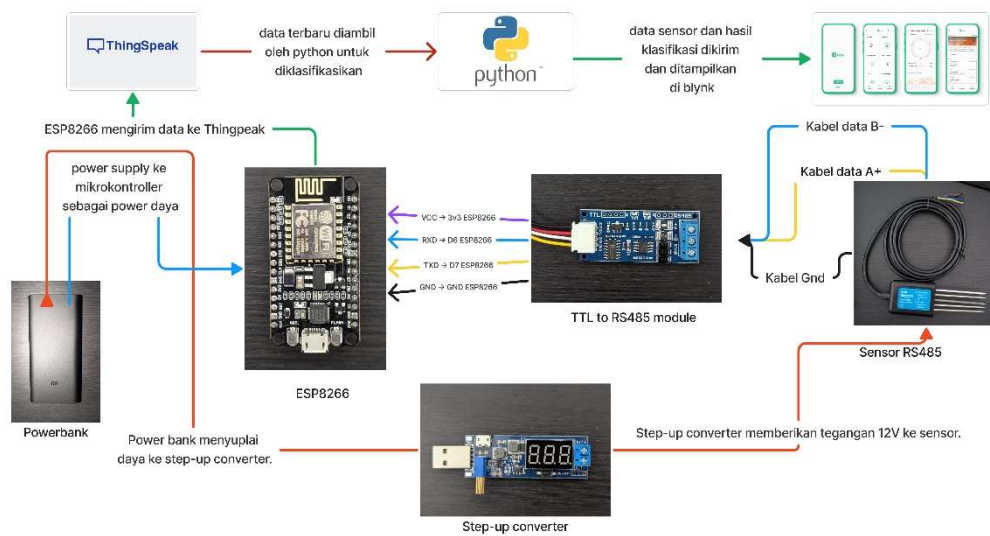


Gambar III. 1 *Flowchart* Penelitian

3.2. Desain Sistem

Pada sub-bab ini, dijelaskan komponen-komponen alat yang digunakan dalam desain sistem *IoT* untuk klasifikasi kesuburan tanah berbasis *machine learning*.

Berikut adalah diagram sistem *IoT* Gambar III. 2 berikut :



Gambar III. 2. Diagram Sistem IoT

Gambar III. 2 tersebut merupakan desain sistem klasifikasi kesuburan tanah berbasis IoT. Sistem ini menggunakan sensor RS485 untuk membaca parameter tanah (Nitrogen, Fosfor, Kalium, pH, suhu, dan kelembapan), yang kemudian dikirim ke *mikrokontroler ESP8266* melalui modul *TTL to RS485*.

Sensor mendapatkan suplai daya 12V dari *step-up converter*, yang mengambil sumber tenaga dari *powerbank*. Data sensor dikirim oleh *ESP8266* ke platform *Thingspeak*, lalu diambil oleh program *Python* untuk diklasifikasikan menggunakan model *machine learning*. Hasil klasifikasi dan data sensor ditampilkan melalui aplikasi *Blynk*.

Dengan sistem ini, data tanah dapat diukur, dikirimkan, dan dianalisis untuk membantu petani menentukan langkah-langkah dalam meningkatkan kesuburan tanah.

3.3. Alat dan Bahan

Subbab ini menjelaskan secara rinci mengenai perangkat keras, perangkat lunak, dan data yang digunakan dalam proses penelitian ini. Pemilihan alat dan bahan didasarkan pada kebutuhan sistem klasifikasi kesuburan tanah berbasis data sensor dengan pendekatan *machine learning*.

3.3.1. Perangkat Keras (Hardware)

1. *ESP8266 NodeMCU*

Merupakan mikrokontroler berbasis *WiFi* yang digunakan sebagai pusat kendali dan akuisisi data dari sensor serta pengiriman data ke platform *cloud*.

2. Sensor Tanah RS485 NPK + pH + Suhu + Kelembapan

Sensor ini digunakan untuk mengukur parameter kimia dan fisik tanah, yaitu kadar nitrogen (N), fosfor (P), kalium (K), pH tanah, suhu, dan kelembapan.

3. Modul TTL to RS485

Berfungsi sebagai konverter komunikasi antara mikrokontroler dan sensor RS485 agar data dapat terbaca dengan benar.

4. *Step-Up Converter (DC-DC Boost Converter)*

Digunakan untuk meningkatkan tegangan dari *powerbank* 5V ke tegangan yang dibutuhkan sensor (12V).

5. *Powerbank 5V*

Sebagai sumber daya portabel untuk memberikan catu daya ke sistem mikrokontroler dan sensor.

6. *Breadboard* dan Kabel *Jumper*

Digunakan untuk merakit dan menghubungkan komponen elektronik selama proses prototipe dan pengujian alat.

3.3.2. Perangkat Lunak (Software)

Perangkat lunak yang digunakan dalam penelitian ini meliputi:

1. *Arduino IDE*

Digunakan untuk menulis, mengunggah, dan menjalankan program pada mikrokontroler *ESP8266*.

2. *Python (Jupyter Notebook/Google Colab)*

Digunakan untuk *preprocessing* data, normalisasi, pelatihan model *machine learning*, evaluasi model, serta klasifikasi data.

3. *Library Python*

Beberapa pustaka yang digunakan antara lain:

pandas, *numpy*: untuk pengolahan dan analisis data

scikit-learn: untuk proses *machine learning* seperti *SVM*, *GridSearchCV*, *LabelEncoder*

imblearn: untuk teknik *oversampling* (*SMOTE*)

matplotlib, seaborn: untuk visualisasi data dan evaluasi model

4. *Thingspeak*

Digunakan sebagai platform *cloud IoT* untuk menyimpan dan memonitor data sensor secara *online*.

5. *Blynk*

Digunakan untuk memonitoring data dari *cloud IoT* sekaligus menampilkan hasil klasifikasi kesuburan tanah dari data sensor yang didapat di *Thingspeak*

6. *Excell*

Digunakan untuk melakukan *labeling* sampel tanah

3.3.3. Data Penelitian

Data yang digunakan merupakan data hasil pembacaan sensor tanah yang dikumpulkan melalui *mikrokontroler ESP8266* dan dikirimkan ke platform *Thingspeak* setiap 30 detik. *Dataset* terdiri dari parameter:

1. Nitrogen (N)
2. Phosphorus (P)
3. Kalium (K)
4. pH tanah
5. Suhu tanah
6. Kelembapan tanah

Setiap data di-labeli berdasarkan kondisi nyata tanah yang diamati secara manual. Sampel tanah dibagi menjadi tiga kategori:

a. Rendah

Tanah yang terlihat gersang, kering, dan cenderung tidak memiliki vegetasi.

b. Sedang

Tanah yang berada di pot atau lahan dengan perawatan biasa, masih memiliki vegetasi namun tidak terlalu subur.

c. Tinggi

Tanah media tanam yang kaya akan nutrisi atau telah diberi pupuk secara berkala dan menunjukkan vegetasi yang tumbuh baik.

Klasifikasi label tersebut dilakukan berdasarkan pengamatan lapangan sebagai representasi dari *ground truth* yang menjadi acuan dalam pelabelan *dataset*

3.4. Pengumpulan Data

Sub-bab ini menjelaskan proses pengumpulan data yang dilakukan untuk mendukung pengembangan model klasifikasi kesuburan tanah berbasis Support Vector Machine (*SVM*). Data dikumpulkan menggunakan sistem IoT yang dirancang khusus dengan sensor tanah NPK RS485, yang terintegrasi dengan mikrokontroler untuk pengiriman data. Proses ini fokus pada pengambilan data mentah yang mencakup parameter-parameter tanah yaitu suhu, kelembapan, pH, dan kandungan nutrisi utama

3.4.1. Proses Pengumpulan Data

Proses pengumpulan data dilakukan melalui langkah-langkah berikut:

1. Penentuan Lokasi Pengambilan Sampel
 - a. Lokasi percobaan dipilih berdasarkan variasi jenis tanah, yakni tanah pertanian, tanah kosong, atau tanah dengan kelembapan yang bervariasi.
 - b. Sampel tanah diambil dari beberapa titik berbeda dalam area penelitian, sehingga data yang diperoleh mencerminkan kondisi tanah yang beragam.
2. Pengaturan Perangkat *IoT*
 - a. Sensor tanah NPK RS485 digunakan untuk mengukur parameter utama: suhu, kelembapan, pH, Nitrogen (N), Fosfor (P), dan Kalium (K).
 - b. Sensor dihubungkan ke *mikrokontroler ESP8266* melalui RS485-to-TTL converter, yang berfungsi sebagai antarmuka penghubung.

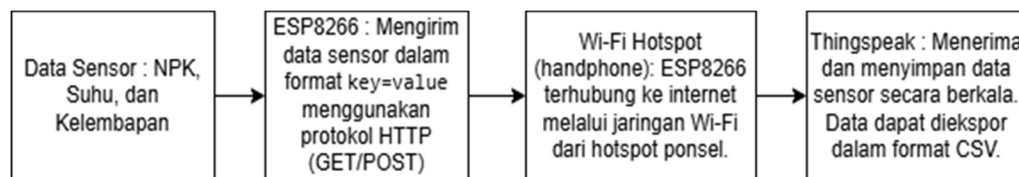
- c. Data yang diperoleh dari sensor dikirimkan ke server *Thingspeak* untuk disimpan dan dianalisis.

3. Pengambilan Data di Lapangan

- a. Sensor ditempatkan langsung di tanah pada kedalaman 5-10 cm untuk memastikan hasil pengukuran akurat.
- b. *Mikrokontroler* membaca data dari sensor, mengonversinya menjadi format digital, dan mengirimkan data tersebut melalui *Wi-Fi* dari *hotspot handphone* ke server *Thingspeak*

4. Arsitektur Pengiriman Data

Berikut adalah arsitektur pengiriman data dari data sensor, kemudian dikirimkan oleh *ESP8266*, lalu menggunakan *Wi-Fi* untuk dikirimkan ke *Thingspeak* pada Gambar III. 3



Gambar III. 3. Arsitektur Pengiriman data

3.4.2. Interval Waktu Pengambilan Data

Pengumpulan data dilakukan dengan interval waktu tertentu, yang dirancang untuk menangkap variasi parameter tanah secara optimal:

1. Pengujian pengambilan data

Data diambil setiap 30 detik selama sesi pengujian awal. Interval ini dirancang untuk memahami perubahan cepat dalam parameter seperti suhu dan kelembapan.

2. Pengamatan Jangka Panjang

Untuk pengamatan rutin, data diambil setiap 30 menit atau 1 jam. Interval ini memungkinkan analisis pola parameter tanah dalam periode harian atau mingguan.

Pengaturan interval pengambilan data dapat disesuaikan berdasarkan kebutuhan penelitian dan volume data yang dihasilkan.

3.5. Labeling Data

Pelabelan data merupakan tahap penting dalam proses klasifikasi karena kualitas label akan sangat memengaruhi performa model. Pada penelitian ini, label kesuburan tanah diberikan berdasarkan nilai kalium dan juga validasi observasi langsung terhadap kondisi fisik dan lingkungan tanah yang diamati secara nyata (*ground truth*).

3.5.1. Interval Waktu Pengambilan Data

Penelitian ini melakukan klasifikasi kesuburan tanah berdasarkan kadar unsur hara Kalium (K) yang diperoleh dari hasil pembacaan sensor tanah berbasis IoT. Penentuan kategori mengacu pada Blesslin Sheeba et al. (2022), yang mengklasifikasikan Kalium ke dalam lima tingkat kesuburan pada Tabel III. 1 ini :

Tabel III. 1. Tabel Klasifikasi Menurut Blesslin Sheeba et al. (2022)

Rating	Kalium (K) (ppm)
<i>Very High</i>	> 300
<i>High</i>	200–300
<i>Medium</i>	151–200
<i>Low</i>	101–150
<i>Very Low</i>	< 100

Untuk menyederhanakan sistem klasifikasi dan menyesuaikan dengan distribusi data lapangan, kategori ini digabung menjadi tiga kelas utama, seperti ditunjukkan pada Tabel III. 2

Tabel III. 2. Contoh Pelabelan Data

Rating	Label Penelitian Ini
<i>Very High, High</i>	Tinggi
<i>Medium</i>	Sedang
<i>Low, Very Low</i>	Rendah

Proses pelabelan dilakukan secara manual berdasarkan nilai Kalium yang diperoleh dari sensor. Dengan demikian, setiap sampel data memiliki label kesuburan yang mengacu pada klasifikasi ilmiah berbasis kadar unsur hara Kalium

3.5.2. Validasi Ground Truth Berdasarkan Observasi Lapangan

Untuk meningkatkan keandalan pelabelan data, dilakukan validasi berdasarkan *ground truth* observasional, yaitu pengamatan langsung terhadap kondisi fisik tanah dan vegetasi di lokasi pengambilan data. Validasi dilakukan dengan mencatat indikator visual berikut:

- Kategori Tinggi: Vegetasi tumbuh subur dan lebat, warna tanah cenderung gelap dan lembap.
- Kategori Sedang: Vegetasi ada namun pertumbuhannya sedang, tanah tidak terlalu gembur atau kering.
- Kategori Rendah: Tidak terdapat vegetasi atau hanya sedikit, tanah tampak gersang dan kering.

Hasil observasi visual tersebut dibandingkan dengan label hasil klasifikasi berbasis sensor Kalium. Jika terdapat ketidaksesuaian antara kondisi visual dan hasil sensor, maka dilakukan peninjauan ulang terhadap data dan kemungkinan *noise* dari pembacaan sensor.

Validasi ini mengikuti prinsip yang digunakan dalam penelitian berbasis sensor dan observasi lapangan seperti pada Sinha et al. (2021), di mana *ground truth* visual digunakan untuk menilai akurasi pelabelan sensor terhadap kondisi nyata di lapangan.

Dengan pendekatan ini, pelabelan berdasarkan *ground truth* lapangan dianggap paling realistis dan mendekati kondisi nyata. Selain itu, metode ini lebih sesuai untuk sistem berbasis sensor yang digunakan langsung di lapangan tanpa perlu uji laboratorium yang memakan biaya dan waktu lebih banyak.

3.5.3. Contoh Pelabelan Data

Format untuk sampel pelabelan data seperti pada Tabel III. 3

Tabel III. 3 Format Pelabelan Data

Waktu	Suhu (°C)	Kelembapan (%)	pH	N (ppm)	P (ppm)	K (ppm)	Kategori Kesuburan
10.01	27	55	6,8	220	470	460	Tinggi
11.00	30	50	6,2	80	187	175	Sedang

12.00	32	75	7.8	26	106	99	Rendah
13.01	25	65	6.1	218	551	548	Tinggi
14.01	35	40	8	26	108	100	Rendah
16.01	31	64	9.4	0	208	186	Sedang (Tinjau Ulang)

3.6. Pre-Processing Data

Sub-bab ini menjelaskan tahap *pre-processing* data yang dilakukan untuk mempersiapkan data hasil pengumpulan menjadi dataset yang bersih, terstruktur, dan siap digunakan dalam pelatihan dan pengujian model *Support Vector Machine (SVM)*. *Pre-processing* data mencakup pembersihan data, normalisasi data, dan pengorganisasian dataset.

3.6.1. Tujuan Pre-Processing Data

Pre-processing dilakukan untuk:

1. Meningkatkan Kualitas Data: Menghapus data yang tidak valid, mengurangi *noise*, dan menangani data yang hilang.
2. Menyamakan Skala Parameter: Melakukan normalisasi agar setiap parameter berada pada skala yang sama, sehingga algoritma *SVM* dapat bekerja optimal.
3. Menyusun Dataset Terstruktur: Membentuk dataset yang siap digunakan untuk pelatihan dan pengujian model.

3.6.2. Tahapan Pre-Processing Data

Tahapan pre-processing yang dilakukan adalah sebagai berikut:

1. Pembersihan Data (*Data Cleaning*)

Data mentah dari sensor dapat memiliki berbagai masalah, seperti data yang hilang. Langkah-langkah pembersihan data meliputi:

Penanganan Data Hilang:

- a. Metode Imputasi: Menggantikan data yang hilang dengan rata-rata (mean) atau nilai median dari parameter yang sama.
- b. Penghapusan Data Tidak Valid: Data yang hilang secara signifikan (>50% parameter hilang) dihapus untuk menjaga integritas dataset.

2. Normalisasi Data

Parameter tanah memiliki skala nilai yang berbeda-beda, seperti pH dengan rentang 0-14, sementara Nitrogen dalam persen (<1%). Untuk menyamakan skala, normalisasi dilakukan menggunakan Min-Max Scaling dengan rumus pada (III. 1) :

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (III. 1)$$

Contoh:

Jika suhu berkisar antara 15°C hingga 35°C, nilai 25°C akan dinormalisasi menjadi (III. 2):

$$X_{norm} = \frac{25 - 15}{35 - 15} = 0.5 \quad (III. 2)$$

3. Pengorganisasian Dataset

Setelah data dibersihkan dan dinormalisasi, dataset diorganisasikan ke dalam format tabel dengan kolom fitur (input) dan label (Output). Fitur mencakup parameter tanah yakni suhu, kelembapan, pH, Nitrogen, Fosfor, dan Kalium, sementara label merupakan kategori kesuburan tanah.

3.7. Split Data

Sub-bab ini menjelaskan proses pembagian dataset yang telah melalui tahap pre-processing menjadi data latih (training data) dan data uji (testing data). Pembagian ini bertujuan untuk memastikan bahwa model *Support Vector Machine (SVM)* dapat dilatih dan diuji secara terpisah, sehingga performa model dapat dievaluasi secara objektif dan menghindari masalah *Overfitting*.

3.7.1. Tujuan Split Data

Pembagian dataset bertujuan untuk:

1. Melatih Model: Data latih digunakan untuk membangun model *SVM* yang mampu mempelajari pola hubungan antara parameter tanah dan kategori kesuburan tanah.
2. Menguji Model: Data uji digunakan untuk mengukur performa model pada data baru yang belum pernah dilihat sebelumnya.

Memastikan Generalisasi Model: Dengan menggunakan data uji yang terpisah, model dapat dievaluasi apakah mampu bekerja dengan baik pada data baru di luar dataset pelatihan.

3.7.2. Rasio Split Data

Dataset dibagi menjadi dua bagian:

1. Data Latih (Training Data):
 - a. 80% dari total dataset.
 - b. Digunakan untuk melatih model *SVM* agar dapat mengenali pola dalam data.
2. Data Uji (Testing Data):
 - a. 20% dari total dataset.
 - b. Digunakan untuk menguji performa model pada data yang belum pernah digunakan selama proses pelatihan.

Dataset memiliki 1400 sampel:

- a. 1120 sampel digunakan sebagai data latih.
- b. 280 sampel digunakan sebagai data uji.

3.8. SMOTE

Smote dalam konteks ini bertujuan untuk mengatasi masalah ketidakseimbangan data antara kelas Tinggi, Sedang, dan Rendah. Pada penelitian ini data yang di smoting adalah data latih bukan data uji karena alasan fundamental terkait integritas evaluasi model dan mencegah kebocoran data

Berikut adalah tabel data setelah smote pada Tabel III. 4

Tabel III. 4 Tabel Distribusi Data setelah Smote

Kategori	Total	Data Latih (80%)	Data Uji (20%)
Tinggi	898	718	180
Sedang	317	718	63
Rendah	195	718	39

3.9. Rancangan Pemodelan dengan *SVM*

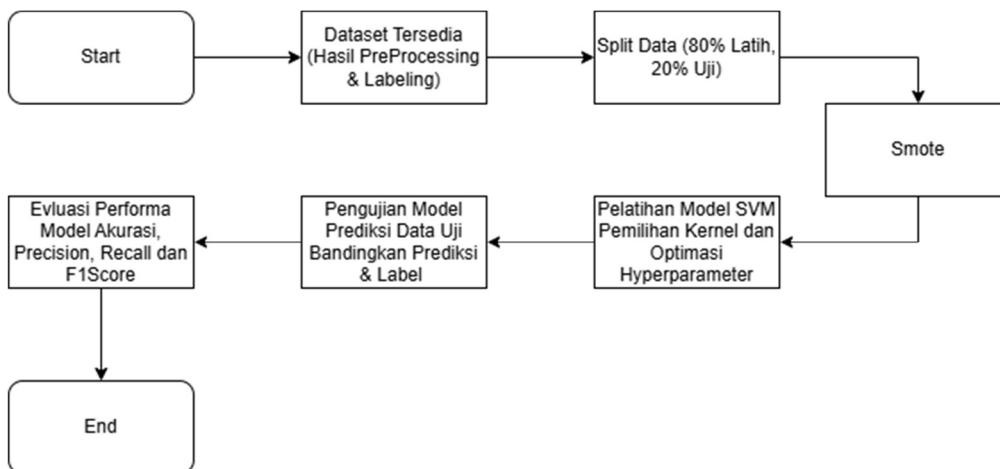
Pemodelan dengan Support Vector Machine (*SVM*) bertujuan untuk membangun model klasifikasi kesuburan tanah berdasarkan parameter yang diukur oleh sensor tanah. *SVM* dipilih karena kemampuannya dalam menangani data kompleks antar parameter tanah, yakni pH, Nitrogen, Fosfor, Kalium, suhu, dan kelembapan.

3.9.1. Tujuan Pemodelan

1. Klasifikasi Kesuburan Tanah: Membagi tanah ke dalam kategori Sangat Subur, Subur, Sedang Subur, Kurang Subur, atau Tidak Subur.
2. Optimasi Model: Mencari parameter terbaik (*hyperparameter*) untuk meningkatkan akurasi prediksi.
3. Evaluasi Performa: Mengukur performa model berdasarkan metrik yakni akurasi, precision, recall, dan *F1-Score*

3.9.2. Alur Pemodelan

Berikut adalah alur proses pemodelan menggunakan *SVM* Gambar III. 4



Gambar III. 4. Alur Pemodelan

Pada pelatihan model *SVM* kernel yang digunakan adalah linear, *RBF*, *Polynomial*, dan *sigmoid* dengan *hyperparameter* c yakni 0.1, 1, 10, dan 100 dan γ yakni Scale dan Auto. Menggunakan *library GridSearchCV* yang merupakan sebuah *class* di *library scikit-learn* (sering disingkat *sklearn*) pada *Python* yang digunakan untuk melakukan *hyperparameter tuning* atau mencari parameter terbaik untuk sebuah model *machine learning*. Dengan begitu pencarian kernel dan *hyperparameter* diambil yang memiliki tingkat akurasi yang terbaik.

3.10. Skenario Pengujian Sistem IoT, Model SVM, dan Evaluasi

Sub-bab ini menjelaskan skenario pengujian yang dilakukan untuk mengukur kinerja sistem *IoT*, model *Support Vector Machine (SVM)*, serta evaluasi keseluruhan sistem. Pengujian bertujuan untuk memastikan bahwa baik perangkat keras *IoT* maupun model klasifikasi bekerja dengan baik dalam mendukung tujuan penelitian.

3.10.1. Pengujian Sistem IoT

Pengujian sistem *IoT* dilakukan untuk memastikan bahwa perangkat keras dapat beroperasi dengan baik, stabil, dan akurat dalam mengukur serta mengirimkan data parameter tanah. Pengujian meliputi beberapa aspek berikut:

1. Pengujian Stabilitas Koneksi IoT

- a. Tujuan: Mengukur keandalan koneksi antara perangkat *IoT* (*ESP8266*) dan server *Blynk Cloud* selama periode waktu tertentu.
- b. Metode:
 - i. Memantau apakah data dikirimkan secara konsisten setiap interval waktu (setiap 30 detik).
 - ii. Mencatat kehilangan data
- c. Hasil yang Diharapkan: Data terkirim secara konsisten ke *Thingspeak* tanpa adanya jeda atau kehilangan data.

2. Pengujian Respons Sistem

- a. Tujuan: Mengukur waktu respons sistem dari pengukuran data oleh sensor hingga data ditampilkan pada aplikasi *Thingsview*.