

BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif yang memiliki fokus terhadap analisis data numerik dalam melakukan klasifikasi kualitas air kolam budidaya ikan nila dengan algoritma *Support Vector Machine (SVM)* (Abdurrahman, Rahmat, dan Sihananto 2023). Data yang digunakan pada penelitian ini merupakan data yang didapatkan dari Dinas Perikanan dan Peternakan Kabupaten Purwakarta yang mencakup sejumlah parameter kualitas air. Seluruh data tersebut dianalisis secara komputasional menggunakan algoritma *SVM* dengan *kernel Radial Basis Function (RBF)* guna mengidentifikasi pola serta hubungan antar parameter yang berkontribusi terhadap klasifikasi kelayakan kualitas air.

Tujuan utama dari analisis adalah untuk mengembangkan model yang mampu memetakan kondisi kualitas air menjadi dua kategori, yaitu “Layak” dan “Tidak Layak” untuk budidaya ikan nila. Pendekatan ini dipilih karena mampu menangkap kompleksitas hubungan non-linier antar variabel kualitas air, seperti suhu, pH, kekeruhan, oksigen terlarut (DO), dan amonia, yang tidak selalu mudah dijelaskan secara eksplisit melalui metode statistik konvensional. Melalui pemanfaatan algoritma *SVM* dengan kernel *RBF*, model dirancang untuk belajar dari data historis dan menemukan *hyperplane* optimal yang dapat memisahkan kedua kelas dengan margin terbesar, sekaligus meminimalkan kesalahan klasifikasi. Selain itu, data yang digunakan dalam penelitian ini telah melalui proses prapengolahan (*preprocessing*) terlebih dahulu, meliputi pembersihan data (*data cleaning*), normalisasi, serta pengkodean label target, untuk memastikan bahwa

Tazkiah Kamilah Aulia, 2025

**PENERAPAN SUPPORT VECTOR MACHINE (SVM) UNTUK ANALISIS KUALITAS AIR PADA
BUDIDAYA IKAN NILA DI KABUPATEN PURWAKARTA**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

data dalam kondisi siap analisis dan bebas dari *noise* atau *outlier* yang dapat memengaruhi kinerja model. Proses pelatihan dan pengujian model dilakukan dengan metode pembagian data (*data split*) menggunakan rasio 60:40, di mana 60% data digunakan untuk pelatihan model dan 40% sisanya digunakan untuk evaluasi kinerja model. Evaluasi terhadap performa model dilakukan dengan menggunakan metrik-metrik seperti akurasi, *precision*, *recall*, *f1-score*, serta visualisasi *confusion matrix* untuk memahami distribusi prediksi yang dihasilkan oleh model.

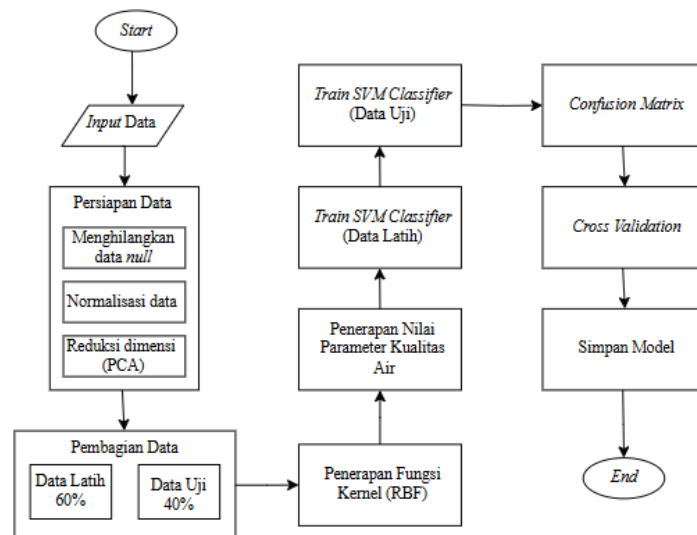
Berdasarkan desain penelitian ini, diharapkan hasil akhir yang diperoleh tidak hanya dapat memberikan informasi yang akurat mengenai status kelayakan kualitas air di kolam budidaya ikan nila, tetapi juga dapat berkontribusi dalam memberikan dasar ilmiah untuk pengambilan keputusan oleh pihak terkait, baik dalam konteks perbaikan manajemen kualitas air maupun dalam penentuan prioritas intervensi terhadap kolam yang berisiko.

3.2 Metodologi Penelitian

Metodologi penelitian pada dasarnya memuat urutan kerja yang disusun secara sistematis dan dijadikan pedoman utama dalam upaya mencapai tujuan penelitian. Setiap tahapan dalam metodologi tidak hanya menggambarkan langkah-langkah teknis yang dilakukan, melainkan juga mencerminkan kerangka berpikir ilmiah. Dalam konteks penelitian ini, metodologi dirancang untuk memastikan bahwa data yang digunakan dapat diproses secara tepat, dianalisis dengan metode yang relevan (Kilani & Kobziev 2016).

Tahapan penelitian disusun secara berkesinambungan mulai dari pengumpulan data, *preprocessing* data untuk memastikan kelayakan data, penerapan algoritma klasifikasi, hingga tahap evaluasi model yang bertujuan menilai kinerja metode yang digunakan. Hubungan antar tahapan ini bersifat integratif, sehingga hasil pada tahap sebelumnya menjadi dasar bagi tahap

berikutnya, dan kesalahan pada salah satu tahap dapat berdampak pada kualitas keseluruhan hasil penelitian. Penelitian ini dilaksanakan berdasarkan tahapan-tahapan sistematis yang telah ditetapkan. Seluruh tahapan tersebut diposisikan sebagai siklus penelitian yang terstruktur, mulai dari input data kualitas air kolam budidaya ikan nila hingga diperolehnya output berupa informasi klasifikasi kelayakan kualitas air. Untuk memberikan pemahaman yang lebih jelas mengenai alur kegiatan penelitian yang dilakukan dari awal hingga akhir, keseluruhan desain metodologi disajikan dalam bentuk bagan alir (*flowchart*) sebagaimana ditampilkan pada Gambar 3.1



Gambar 3.1 Alur Penelitian

Secara umum, penelitian ini dilakukan melalui serangkaian tahapan yang saling berkaitan, dimulai dari persiapan dan pra-pemrosesan data, hingga pembentukan serta evaluasi model. Melalui rangkaian proses tersebut, dihasilkan sebuah model klasifikasi yang dapat digunakan untuk menilai kelayakan kualitas air berdasarkan parameter-parameter yang tersedia. Meskipun demikian, performa model tetap bergantung pada kualitas data masukan serta batasan parameter yang digunakan dalam penelitian ini. Oleh karena itu, metode penelitian yang disusun

diharapkan mampu memberikan gambaran awal mengenai potensi penerapan algoritma SVM dalam klasifikasi kualitas air, sekaligus menjadi dasar bagi pengembangan penelitian lebih lanjut dengan data yang lebih beragam.

3.3 Alat dan Bahan Penelitian

Pelaksanaan penelitian memerlukan penggunaan alat sebagai aspek penting untuk memastikan proses pengambilan dan pengukuran data berjalan sesuai prosedur. Keberadaan alat yang tepat dan sesuai standar sangat berpengaruh terhadap kualitas data yang dihasilkan. Penelitian ini menggunakan beberapa alat utama yang diperlukan untuk mendapatkan data parameter kualitas air. Alat yang digunakan meliputi reagen uji dengan merek dan spesifikasi tertentu, serta alat ukur seperti pH meter, DO meter, dan TDS meter untuk memperoleh data parameter kualitas air secara akurat. Berikut daftar lengkap alat beserta merek yang digunakan disajikan pada Tabel 3.1.

Tabel 3. 1 Alat Pengambilan Data

Alat Penelitian	Merek
pH Meter	pH TDS EC Temp (4 in 1) Meter EZ9908
TDS Meter	HM Digital TDS-3
DO Meter	JPB-70A Pen-Shapped Dissolved Oxygen Tester
Kit Uji Ammonia	SERA Ammonium/Ammonia-Test (NH ₄ /NH ₃) 3x15 mL
Kit Uji Nitrat	SERA Nitrat Test (NO ₃) 3x15 ml
Kit Uji Nitrit	SERA Nitrit Test (NO ₂) 2x15 ml
Kit Uji Fosfat	SERA Fosfat Test (PO ₄) 2x15 ml

Seluruh alat yang tercantum pada Tabel 3.1 digunakan sesuai dengan prosedur selama proses pengambilan sampel air. Penggunaan berbagai alat tersebut bertujuan untuk memperoleh data parameter kualitas air yang dibutuhkan, sehingga dapat memberikan gambaran yang menyeluruh mengenai kondisi lingkungan perairan yang diteliti. Pemilihan alat dilakukan berdasarkan kesesuaian dengan jenis parameter yang diukur serta ketersediaan sumber daya yang mendukung pelaksanaan penelitian ini. Agar dapat menganalisis data diperlukan bahan penelitian yang terangkum di Tabel 3.2

Tabel 3. 2 Bahan Penelitian

Bahan Penelitian	Fungsi
Python v3.9.7	Bahasa pemrograman untuk menganalisis data dan melakukan visualisasi data.
Visual Studio Code v1.60.1.0	<i>Integrated development environment (IDE).</i>
Google Colab	<i>Platform</i> untuk menjalankan <i>notebook Python</i> .
Microsoft Excel 2019	<i>Software spreadsheet</i> untuk analisis data.

Bahan penelitian yang ditampilkan pada Tabel 3.2 merupakan perangkat lunak yang digunakan untuk mendukung proses pengolahan, analisis, serta visualisasi data penelitian. Pemilihan perangkat lunak tersebut didasarkan pada kemampuan masing-masing aplikasi dalam menyediakan fitur yang relevan dengan kebutuhan analisis kualitas air. Dengan adanya perangkat lunak tersebut, data hasil pengukuran dapat diolah secara sistematis, divisualisasikan secara informatif, serta dianalisis dengan metode statistik maupun machine learning sehingga mendukung validitas hasil penelitian.

Tersedianya alat dan bahan penelitian sebagaimana tercantum pada Tabel 3.1 dan Tabel 3.2 memungkinkan proses penelitian dilaksanakan secara terstruktur. Alat berperan dalam mendukung pengambilan sampel di lapangan, sedangkan bahan berupa perangkat lunak digunakan untuk mengolah serta menganalisis data hasil pengukuran. Seluruh komponen tersebut menjadi dasar dalam penerapan metode penelitian yang dijelaskan pada bagian berikutnya, sehingga tahapan penelitian dapat berlangsung secara sistematis dan menghasilkan data yang valid.

3.4 Populasi dan Sampel

3.4.1 Populasi

Populasi dalam suatu penelitian memegang peranan yang sangat penting, karena dari populasi seluruh informasi yang relevan dapat diperoleh. Secara umum, populasi dapat diartikan sebagai keseluruhan elemen yang menjadi perhatian dalam penelitian, baik berupa objek maupun subjek, yang memiliki ciri-ciri atau karakteristik tertentu sesuai dengan fokus kajian yang sedang diteliti. Dengan kata lain, populasi merupakan himpunan besar dari individu, kelompok, atau unit analisis yang memiliki kesamaan sifat, sehingga dapat digunakan sebagai dasar dalam pengambilan kesimpulan penelitian. Keberadaan populasi menjadi penting karena penelitian tidak akan dapat berjalan tanpa adanya sumber data yang jelas, dan dari populasi inilah nantinya peneliti akan menentukan bagian yang lebih kecil, yaitu sampel, untuk dianalisis secara lebih mendalam (Subhaktiyasa 2024).

Populasi pada penelitian ini adalah seluruh kolam budidaya ikan nila yang tercatat pada Dinas Perikanan dan Peternakan Kabupaten Purwakarta. Berdasarkan data yang diperoleh, jumlah populasi tersebut mencapai 224 kolam ikan nila.

3.4.2 Sampel

Sampel dapat diartikan sebagai sebagian kecil dari populasi yang dipilih untuk dijadikan sumber data utama dalam suatu penelitian. Dengan kata lain, sampel merupakan representasi dari populasi yang digunakan untuk menggambarkan kondisi keseluruhan tanpa harus meneliti setiap elemen yang ada di dalam populasi tersebut (Pitaloka, 2022).

Adapun sampel pada penelitian ini terdiri atas 166 kolam budidaya ikan nila yang dipilih dari total populasi sebanyak 224 kolam. Sampel tersebut berasal dari beberapa kecamatan utama yang di Kabupaten Purwakarta antara lain Bungursari, Babakancikao, Jatiluhur, Purwakarta, Pasawahan, Sukatani, Plered, Darangdan, Wanayasa, Kiarapedes, hingga Maniis.

3.5 Teknik Penelitian

3.5.1 Teknik Pembagian Data

Pada tahap ini, dataset yang berjumlah 166 sampel hasil pengukuran kualitas air terlebih dahulu dipersiapkan dengan proses pembersihan data dan normalisasi. Data yang digunakan hanya mencakup variabel numerik, yaitu parameter-parameter fisik dan kimia air seperti pH, suhu air, *total dissolved solid* (TDS), *dissolved oxygen* (DO), amonia, nitrit, nitrat, serta fosfat, data kualitas air yang berjumlah 166 sampel dibagi menjadi dua bagian menggunakan metode *train-test split*. *Train-test split* merupakan salah satu teknik validasi sederhana yang bertujuan untuk memisahkan dataset menjadi data latih (*training set*) dan data uji (*testing set*) dengan perbandingan tertentu (Oktafiani dkk., 2023)

Penelitian ini menggunakan rasio 60:40, yaitu sebanyak 60% data (99 sampel) digunakan sebagai data latih, sedangkan 40% sisanya (67

sampel) digunakan sebagai data uji. Data latih berfungsi untuk membangun dan melatih model *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF), sedangkan data uji digunakan untuk mengukur performa model terhadap data yang belum pernah dilihat sebelumnya. Pemilihan rasio 60:40 didasarkan pada pertimbangan agar model memperoleh cukup banyak data untuk proses pembelajaran sekaligus menyediakan data uji yang memadai untuk mengevaluasi kemampuan generalisasi model. Dengan pembagian ini, diharapkan hasil evaluasi model tidak hanya menunjukkan akurasi pada data latih, tetapi juga dapat menggambarkan kinerja model pada data baru secara lebih objektif.

Gambar 3.2 akan menunjukkan lima sampel pertama dari data latih dan data uji yang menyertakan informasi lokasi, parameter kualitas air, serta label klasifikasi (Layak atau Tidak Layak). Gambar 3.2 tersebut disajikan untuk memberikan gambaran awal mengenai struktur data yang digunakan dalam penelitian ini. Dengan menampilkan sebagian sampel data, diharapkan pembaca dapat memahami karakteristik utama dari variabel yang dianalisis serta distribusi label klasifikasi yang menjadi dasar dalam proses pelatihan dan pengujian model.

Data Latih:				
	lokasi (kelompok)	Derajat Keasaman (pH)	Temperatur Air (°C)	\
10	BBI 1	8.42	32.4	
98	cihuni 2	7.75	26.8	
125	pasawahan anyar 6	8.24	29.0	
0	warungkadu	7.90	30.0	
123	pasawahan anyar 4	7.70	29.2	
	Oksigen Terlarut (DO; mg/L)	Ammonia (NH3 NH4; mg/L)		Label
10	2.9	0.0		Tidak Layak
98	5.5	0.0		Layak
125	3.3	0.0		Layak
0	6.1	0.5		Tidak Layak
123	4.8	0.5		Tidak Layak
Data Uji:				
	lokasi (kelompok)	Derajat Keasaman (pH)	Temperatur Air (°C)	\
100	cihuni 4	8.00	27.0	
136	pasawahan 5	8.32	29.0	
78	cilegog 2	8.27	30.4	
56	cilangkap 6	7.21	28.2	
97	cihuni 1	7.90	26.3	
	Oksigen Terlarut (DO; mg/L)	Ammonia (NH3 NH4; mg/L)		Label
100	6.2	0.0		Layak
136	5.9	0.0		Layak
78	6.1	0.0		Layak
56	6.0	0.0		Layak
97	5.3	0.0		Layak

Gambar 3. 2 Data Latih & Data Uji

Gambar 3.2 tersebut terlihat bahwa 5 data awal memuat nilai parameter numerik kualitas air beserta label hasil kategorisasi. Data latih digunakan untuk melatih model SVM RBF, sedangkan data uji dimanfaatkan untuk mengukur kinerja model dalam memprediksi label pada data yang belum pernah digunakan dalam proses pelatihan.

3.5.2 Teknik Analisis Data

Teknik analisis data dimulai dengan tahap pengumpulan data kualitas air dari kolam budidaya ikan nila, yang mencakup parameter-parameter seperti suhu, pH, oksigen terlarut, kadar amonia, dan parameter lainnya yang relevan. Data tersebut diperoleh dari Dinas Perikanan dan Peternakan Kabupaten Purwakarta. Setelah data terkumpul, dilakukan proses *preprocessing* data untuk memastikan mutu dan konsistensi data yang akan digunakan. Tahap ini melibatkan pembersihan data untuk mengatasi nilai-nilai yang hilang, atau data yang tidak valid. Selain itu, data dinormalisasi menggunakan metode seperti *min-max scaling* atau *standardization* agar setiap parameter memiliki skala yang setara, sehingga tidak ada parameter yang mendominasi analisis.

Setelah data diproses, dilakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA). *Principal Component Analysis* (PCA) merupakan salah satu metode reduksi dimensi yang bertujuan menyederhanakan variabel-variabel dalam data dengan cara mentransformasikannya ke dalam himpunan variabel baru yang disebut komponen utama (*principal component*). *Principal Component Analysis* (PCA) adalah metode yang digunakan untuk mempermudah analisis data dengan mentransformasikan data secara linier menjadi sistem koordinat baru yang memiliki varian terbesar. Dengan menggunakan metode ini, dimensi data dapat dikurangi tanpa mengubah sifat-sifat penting data secara

signifikan (Abdillah & Prianto., 2019). Komponen utama ini dibentuk dari kombinasi linier variabel asal, sehingga tetap mempertahankan sebagian besar variasi dari data. Secara matematis, komponen utama dapat dinyatakan sebagai berikut, $PC_i = Z \cdot w_i$.

Pada persamaan PCA, Z merepresentasikan data hasil standardisasi, yaitu matriks data yang telah ditransformasikan sehingga memiliki rata-rata nol dan standar deviasi satu. Vektor w_i merupakan *eigenvector* ke- i yang menunjukkan arah utama dari komponen ke- i . Sementara itu, PC_i adalah nilai komponen utama ke- i yang dihasilkan dari perkalian antara matriks data terstandarisasi dengan *eigenvector* tersebut.

Penerapan PCA, data yang semula memiliki banyak variabel dapat dipadatkan menjadi sejumlah komponen utama yang lebih ringkas, tetapi tetap membawa informasi utama yang diperlukan untuk pelatihan model. Hal ini bermanfaat untuk mengurangi kompleksitas perhitungan, mengatasi potensi multikolinearitas antar variabel, serta meningkatkan kinerja model klasifikasi. Setelah melalui proses PCA, langkah selanjutnya adalah membagi dataset menjadi dua bagian, yaitu data latih (*train*) dan data uji (*test*), dengan rasio 60:40. Data latih digunakan untuk melatih model *Support Vector Machine* (SVM), sedangkan data uji dipergunakan untuk mengevaluasi performa model. Pada tahap pelatihan model, kernel SVM yang digunakan dipilih berdasarkan karakteristik data. Setelah pemilihan kernel yang sesuai, model SVM dilatih dengan parameter-parameter tertentu, seperti nilai C (*regularization parameter*) dan gamma. Proses ini dapat melibatkan penyetelan parameter menggunakan metode *grid search* untuk menemukan kombinasi parameter terbaik. Model yang telah dilatih kemudian dievaluasi menggunakan data uji untuk mengukur kinerjanya. Evaluasi dilakukan dengan menggunakan berbagai metrik, untuk mengukur

tingkat kebenaran prediksi. Selain itu, *confusion matrix* digunakan untuk memberikan analisis rinci mengenai prediksi yang benar dan salah. Untuk memastikan performa model stabil dan dapat diandalkan, diterapkan teknik validasi silang (*cross-validation*), data dibagi menjadi beberapa subset untuk melatih dan menguji model secara bergantian. Teknik analisis data ini diharapkan mampu menghasilkan model SVM yang andal dalam mengklasifikasikan kualitas air kolam budidaya ikan nila, sekaligus memberikan rekomendasi yang berguna bagi pengelolaan kolam yang lebih optimal.

3.6 Latar/Setting Penelitian

3.6.1 Waktu Penelitian

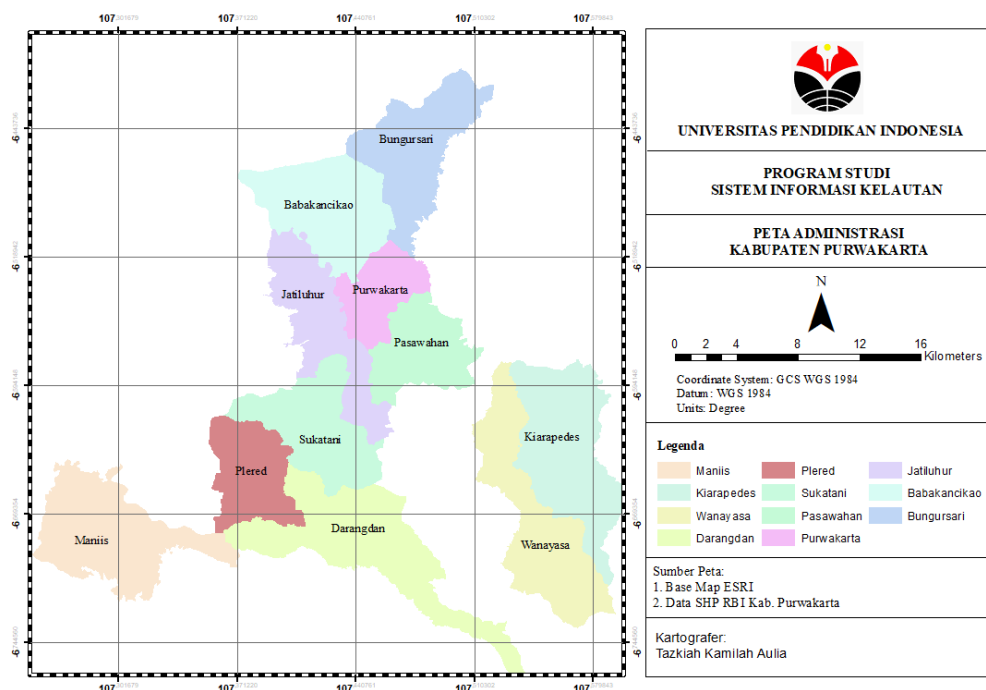
Penelitian ini dilaksanakan selama kurun waktu 5 bulan, dimulai pada tanggal Maret 2025 dan berakhir pada tanggal Agustus 2025, dengan pertimbangan bahwa rentang waktu tersebut dinilai cukup untuk mengakomodasi proses pengumpulan data yang lebih menyeluruh sekaligus memberikan ruang yang memadai bagi peneliti untuk melakukan analisis yang mendalam terhadap berbagai parameter kualitas air pada kolam budidaya ikan nila. Data kualitas air yang dianalisis mencakup pengumpulan data dari tahun 2024 hingga Maret 2025.

3.6.2 Tempat Penelitian

Penelitian ini dilakukan pada kelompok pembudidaya ikan nila yang tersebar di berbagai kecamatan di Kabupaten Purwakarta. Keberagaman lokasi tersebut dipilih agar data yang diperoleh lebih representatif, mencakup variasi kondisi lingkungan dan teknik budidaya yang ada di lapangan. Sebaran titik pengamatan tersebut mencakup beberapa kecamatan utama, antara lain Bungursari, Babakancikao, Jatiluhur, Purwakarta,

Pasawahan, Sukatani, Plered, Darangdan, Wanayasa, Kiarapedes, hingga Maniis.

Pemilihan lokasi mempertimbangkan intensitas kegiatan budidaya ikan nila di setiap kecamatan, sehingga data yang terkumpul dapat menggambarkan kondisi kualitas air pada skala wilayah yang lebih luas. Untuk memperjelas lokasi penelitian, peta wilayah Kabupaten Purwakarta disajikan pada Gambar 3.3 guna menunjukkan sebaran titik pengambilan data pada kolam budidaya ikan nila dengan total 166 sampel.



Gambar 3. 3 Peta Penelitian

Gambar 3.3 memperlihatkan “Peta Administrasi Kabupaten Purwakarta” yang membagi wilayah penelitian ke dalam beberapa kecamatan. Peta ini bertujuan untuk memberikan gambaran lokasi penelitian secara spasial, sekaligus menunjukkan batas-batas administrasi

Tazkiah Kamilah Aulia, 2025

PENERAPAN SUPPORT VECTOR MACHINE (SVM) UNTUK ANALISIS KUALITAS AIR PADA BUDIDAYA IKAN NILA DI KABUPATEN PURWAKARTA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

yang ada. Kabupaten Purwakarta terdiri atas sejumlah kecamatan. Setiap kecamatan ditampilkan dengan warna berbeda agar memudahkan dalam identifikasi wilayah. Informasi spasial yang digunakan mengacu pada sistem koordinat GCS WGS 1984 dengan satuan derajat, serta bersumber dari Base Map ESRI dan data SHP RBI Kabupaten Purwakarta. Peta ini menjadi acuan dasar dalam menentukan sebaran lokasi penelitian, serta penting untuk memahami kondisi geografis wilayah Purwakarta secara keseluruhan.