

PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK KLASIFIKASI IKAN HIAS CHANNA

Fariz Putra Rizki, Asril Adi Sunarto, Winda Apriandari
Program Studi Teknik Informatika, Fakultas Sains dan Teknologi
Universitas Muhammadiyah Sukabumi, Jawa Barat, Indonesia
farizputrarizki@gmail.com

ABSTRAK

Ikan Channa, atau yang dikenal sebagai ikan gabus, merupakan ikan air tawar yang mendiami berbagai habitat seperti sungai, danau, serta rawa. Ikan ini memiliki nilai ekonomi yang tinggi dan sangat digemari oleh para pecinta ikan hias karena warna tubuhnya yang menarik, yang terdiri dari perpaduan warna oranye kekuningan dan hitam dengan pola bunga. Budidaya ikan Channa telah menjadi sumber pendapatan yang signifikan bagi banyak petani ikan, terutama karena prospeknya yang menguntungkan bahkan pada lahan terbatas dengan sumber daya air yang minimal. Salah satu tantangan utama dalam budidaya ini adalah melakukan klasifikasi ikan secara akurat berdasarkan karakteristik fisiknya. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi ikan Channa dengan memanfaatkan algoritma Support Vector Machine (SVM), yang telah terbukti efektif dalam berbagai aplikasi lain, termasuk diagnosis medis. Sistem yang dikembangkan memanfaatkan data non-linier dengan enam variabel fisik: panjang, berat, warna, bentuk sirip, panjang sirip, dan bentuk kepala. Hasil penelitian ini menunjukkan bahwa penerapan algoritma Support Vector Machine (SVM) untuk klasifikasi ikan hias Channa memberikan hasil dengan akurasi sebesar 69%, presisi 66%, recall 69%, dan F1-score 67%. Algoritma SVM terbukti mampu mengklasifikasikan ikan Channa berdasarkan karakteristik fisik seperti panjang, berat, warna, bentuk sirip, panjang sirip, dan bentuk kepala secara akurat dan efisien.

Kata kunci : *Ikan channa, Support Vector Machine, Klasifikasi, Karakteristik Ikan*

1. PENDAHULUAN

Ikan Channa, yang juga dikenal sebagai ikan gabus, adalah salah satu spesies ikan air tawar yang sering ditemukan di berbagai jenis perairan seperti muara sungai, danau, dan rawa [1].

Ikan ini memiliki kemampuan adaptasi yang luar biasa, bahkan dalam kondisi lingkungan dengan kadar oksigen rendah, sehingga sering dipilih untuk budidaya ikan air tawar [1].

Sebagai predator karnivora, ikan Channa dewasa memangsa berbagai jenis makanan seperti serangga air, potongan hewan air, udang, dan detritus, yang menjadikannya komponen penting dalam rantai makanan di ekosistemnya [2].

Dalam dunia ikan hias, Channa sangat diminati oleh para penghobi ikan predator [2].

Ikan ini dikenal karena keindahan warnanya yang mencolok, dengan bagian bawah tubuh berwarna oranye kekuningan, bagian atas yang dominan hitam, serta corak bunga hitam dengan pinggiran putih di sisi tubuhnya [5].

Karena keindahan tersebut, ikan Channa sering dipelihara di akuarium atau kolam kecil untuk keperluan estetika dan rekreasi, menambah nilai bagi para penghobi [2].

Klasifikasi merupakan proses krusial dalam mengidentifikasi kelas atau kategori dari suatu objek berdasarkan karakteristik yang dimiliki [8].

Dalam komputasi, terdapat berbagai algoritma yang dapat digunakan untuk tujuan klasifikasi, termasuk Decision Tree (DT), Support Vector Machine (SVM), Neural Network (NN), K-Nearest Neighbors (K-NN), dan Naïve Bayes (NB) [8].

Di antara berbagai algoritma ini, SVM dikenal sebagai salah satu yang paling efektif untuk mencapai klasifikasi dengan tingkat akurasi yang tinggi [8].

Algoritma SVM bekerja dengan prinsip Structural Risk Minimization (SRM) yang bertujuan menemukan hyperplane optimal untuk memisahkan data antar kelas [7].

Penelitian sebelumnya telah menunjukkan bahwa algoritma SVM efektif dalam berbagai aplikasi, termasuk diagnosis medis, dengan tingkat akurasi yang bervariasi tergantung pada kompleksitas data dan model yang digunakan [7].

Berdasarkan keunggulan SVM, penelitian ini bertujuan mengembangkan sistem klasifikasi untuk ikan hias Channa menggunakan algoritma SVM. Penelitian ini memanfaatkan data non-linier dengan enam variabel fisik utama, yaitu panjang, berat, warna, bentuk sirip, panjang sirip, dan bentuk kepala, dengan tujuan mencapai klasifikasi yang akurat dan efisien [7].

2. TINJAUAN PUSTAKA

2.1. Ikan Channa

Ikan Channa, yang merupakan spesies gabus, sering ditemukan di sungai-sungai di Indonesia. Ikan ini tergolong karnivora dengan beragam jenis dan warna yang ada di Indonesia, banyak di antaranya digunakan sebagai ikan hias karena ciri khas dan warnanya yang menarik. Dalam komunitas pecinta ikan hias, Channa dikenal sebagai genus ikan predator dari keluarga Channidae, yang berasal dari habitat air tawar di Asia dan Afrika, namun juga umum ditemukan di Indonesia. Dijuluki "*Snakehead*

Emperor" karena bentuk kepalanya yang menyerupai kepala ular, ikan *Channa* memiliki pola tubuh yang unik dan berbagai jenis, sehingga sangat diminati oleh pecinta ikan predator. Popularitas ini juga mendorong harga jual yang tinggi, menciptakan peluang usaha baru bagi para penggemar ikan.[1]

2.2. Taksonomi ikan *Channa*

Ikan *Channa*, yang termasuk dalam kelas Actinopterygii dan berasal dari keluarga Channidae, adalah ikan karnivora dengan tubuh berbentuk hampir bulat dan memanjang, yang menjadi lebih pipih ke arah belakang. Ikan ini memiliki punggung cembung dan perut rata, dengan kepala yang pipih dan tanpa jari-jari sirip keras. Panjang tubuh ikan gabus bervariasi, bisa mencapai 1 hingga 50 cm. Warna tubuhnya biasanya hitam di bagian punggung dan putih di bagian perut.[2]

Taksonomi ikan *Channa* adalah sebagai berikut:

Kingdom: Animalia
Filum: Chordata
Kelas: Actinopterygii
Ordo: Perciformes
Famili: Channidae
Genus: Channa
Spesies: Channa striata

2.3. Machine Learning

Machine Learning atau Pembelajaran Mesin adalah teknik dalam Kecerdasan Buatan (AI) yang bertujuan untuk meniru dan menggantikan peran manusia dalam berbagai kegiatan pemecahan masalah. Secara singkat, *Machine Learning* memungkinkan mesin untuk belajar dan bekerja secara mandiri tanpa perlu arahan dari pengguna. Algoritma *machine learning* umumnya dibagi menjadi tiga kategori utama: supervised learning, unsupervised learning, dan reinforcement learning. Penting untuk memahami ketiga jenis algoritma ini sebelum menerapkan solusi berbasis *machine learning*, karena masing-masing memiliki fungsi dan tujuan yang berbeda.[8]

2.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma pembelajaran terawasi yang bertujuan menganalisis data dan mengenali pola, digunakan dalam analisis klasifikasi dan regresi. Tujuan utama SVM adalah membangun model dari data pelatihan yang mampu memprediksi nilai target dari data uji. SVM menggunakan ruang hipotesis berupa fungsi linier dalam ruang fitur berdimensi tinggi dan dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi serta bias pembelajaran dari teori pembelajaran statistik. SVM bertujuan untuk menemukan hyperplane terbaik yang dapat memisahkan dua kelas pada ruang input. Hyperplane ini membagi ruang vektor menjadi dua bagian, masing-masing merepresentasikan kelas yang berbeda. Hyperplane terbaik adalah yang memisahkan kelas -1 dan kelas +1 secara optimal.[8]

2.5. Knowledge discovery in databases (KDD)

Knowledge discovery in databases adalah proses yang mencakup hasil dari data mining, yaitu proses mengekstraksi pola dan kecenderungan dari data, kemudian mengubah hasil tersebut menjadi informasi yang akurat dan mudah dipahami. Istilah data mining dan *knowledge discovery in databases* sering digunakan secara bergantian untuk menggambarkan proses menemukan informasi tersembunyi dalam basis data yang besar. Meskipun kedua istilah ini memiliki konsep yang berbeda, mereka saling berhubungan. Data mining adalah salah satu tahap dalam keseluruhan proses *knowledge discovery in databases*. [7]

2.6. Python

Python adalah bahasa pemrograman yang sangat populer dan digunakan oleh perusahaan besar serta pengembang untuk membuat berbagai jenis aplikasi, termasuk aplikasi desktop, web, dan mobile. Diciptakan oleh Guido van Rossum di Belanda pada tahun 1990, nama *Python* berasal dari acara televisi favoritnya, Monty Python's Flying Circus. Meskipun awalnya dikembangkan sebagai proyek hobi, *Python* telah berkembang menjadi bahasa pemrograman yang sangat umum digunakan di industri dan pendidikan, berkat kemudahan penggunaannya, kesederhanaannya, sintaks yang intuitif, dan berbagai pustaka yang tersedia[4]

2.7. Google Colab

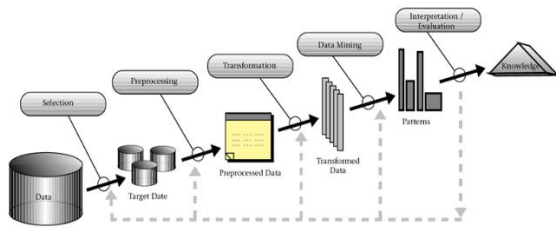
Google Colab adalah layanan komputasi awan yang disediakan oleh Google untuk mendukung pengembangan dan penelitian ilmiah. Dikenal juga sebagai "Colab," produk dari *Google Research* ini memungkinkan pengguna menulis dan mengeksekusi kode *Python* langsung melalui *browser*, menjadikannya sangat cocok untuk keperluan *machine learning*, analisis data, dan pendidikan. Selain itu, *Google Colab* mendukung kolaborasi antar pengembang aplikasi, memfasilitasi kerja sama tim secara efisien. Meskipun *Google Colab* belum begitu dikenal oleh masyarakat umum, layanan ini banyak digunakan oleh developer dan programmer yang membutuhkan lingkungan pemrograman yang fleksibel.[3]

2.8. Flask

Flask merupakan kerangka kerja aplikasi web WSGI yang ringan, dirancang untuk memudahkan pengembangan aplikasi dengan proses yang cepat dan sederhana, sehingga mengurangi waktu pemuatan. Meskipun *Flask* merupakan pembungkus minimalis yang menggabungkan Werkzeug dan Jinja, ia telah berkembang menjadi salah satu kerangka kerja aplikasi web *python* yang paling populer, terutama karena kemampuannya untuk menangani peningkatan kompleksitas aplikasi secara efektif.[3]

3. METODE PENELITIAN

Penelitian ini bertujuan untuk menyajikan gambaran komprehensif mengenai rencana pelaksanaan penelitian, khususnya dalam penerapan klasifikasi penyakit pada ikan *Channa* hias menggunakan algoritma *Support Vector Machine* (SVM). Selain menguraikan secara rinci proses pengumpulan data yang didasarkan pada informasi dari instansi terkait, penelitian ini juga mencakup penjelasan lebih lanjut mengenai berbagai kebutuhan yang diperlukan. Rencana kegiatan akan dipaparkan secara terperinci, mencakup langkah-langkah yang akan diambil selama proses penelitian.



Gambar 1. Alur penelitian SVM menggunakan kdd

3.1. Data Selection

Pada tahap ini, data yang diperoleh melalui proses seleksi data, disesuaikan dengan kebutuhan penelitian. Data yang dikumpulkan dari berbagai sumber, termasuk komunitas pecinta ikan hias dan sumber-sumber terpercaya lainnya, terdiri dari metadata terkait ikan hias *Channa* yang akan diklasifikasikan. Metadata tersebut meliputi variabel seperti Panjang, Berat, Bentuk Sirip, Panjang Sirip, Bentuk Kepala, serta Kelayakan.

3.2. Preprocessing Data

Tahap ini melibatkan pengolahan data guna meningkatkan kualitas dan konsistensi data. Proses yang dilakukan meliputi mencari missing value, information gain dan mutual information. Langkah-langkah ini bertujuan memastikan data siap digunakan dalam pelatihan model SVM guna klasifikasi ikan hias *Channa* dengan performa optimal. Setelah diproses, data dibagi menjadi set pelatihan dan set pengujian untuk mengevaluasi kinerja model.

3.3. Transformation

Pada tahap *Transformation* data dalam pengolahan ikan hias *Channa*, data dibagi menjadi dua kategori utama, yaitu Data Training dan Data Testing. Pembagian ini sangat penting untuk memastikan bahwa model yang dikembangkan mampu melakukan prediksi dan validasi secara efektif dengan menggunakan data baru yang belum pernah diakses sebelumnya. Selain pembagian data, tahap transformasi juga mencakup konversi beberapa variabel dari format asli ke format yang lebih mudah dianalisis. Misalnya, variabel seperti warna, bentuk sirip, dan bentuk kepala yang awalnya berbentuk string diubah menjadi format numerik. Transformasi ini

memungkinkan data untuk diolah dalam analisis statistik dan model prediktif yang akan dikembangkan lebih lanjut dalam klasifikasi ikan hias *Channa*.

3.4. Data Mining

Tahap *data mining* bertujuan untuk mengidentifikasi pola dan informasi penting dalam data yang telah diproses. Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk melakukan klasifikasi ikan hias *Channa*. Proses dimulai dengan melatih model SVM menggunakan *Data Training*, di mana model belajar mengenali pola berdasarkan fitur-fitur yang telah diekstraksi.

3.5. Evaluation

Pada tahap *evaluation*, kinerja model SVM diuji untuk menilai tingkat akurasi dan efektivitasnya dalam mengklasifikasikan ikan hias *Channa*. Proses ini dilakukan dengan menggunakan data *testing* yang belum pernah digunakan selama pelatihan model. Evaluasi utama melibatkan pengukuran *akurasion*, *precision*, *recall*, dan *F1-score*. Akurasi mengindikasikan seberapa sering model menghasilkan prediksi yang tepat.

3.6. Deployment

Pada tahap akhir proyek ini, Pengembangan model dan mengintegrasikannya ke dalam sebuah situs web sederhana. Proyek penelitian ini melibatkan penggunaan *Flask* sebagai kerangka kerja utama untuk mengimplementasikan model tersebut.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Data yang digunakan meliputi berbagai atribut ikan hias *Channa* yang diperoleh dari berbagai sumber diantaranya :

Tabel 1. Variabel *Dataset*

1	Berat	Variabel ini digunakan untuk mengukur massa tubuh ikan, yang dapat memberikan indikasi terkait kesehatan, usia, dan spesiesnya.
2	Panjang	Variabel ini berfungsi untuk menentukan ukuran keseluruhan ikan, yang dapat berkaitan dengan spesies dan tahap perkembangannya.
3	Warna	Variabel ini digunakan untuk mengidentifikasi spesies serta variasi morfologi di dalam spesies yang sama.
4	Bentuk Sirip	Variabel ini digunakan untuk menunjukkan bagaimana ikan beradaptasi dengan lingkungan hidupnya serta meningkatkan daya tarik estetika dalam konteks ikan hias.
5	Panjang Sirip	Variabel dapat memberikan informasi tentang karakteristik fisik yang berbeda antara spesies.
6	Bentuk Kepala	Variabel ini membantu dalam mengidentifikasi spesies dan memahami adaptasi morfologi.

7	Kelayakan	Variabel ini mencatat kategori apakah ikan hias Channa dianggap layak atau tidak layak berdasarkan atribut-atribut lainnya. Variabel ini merupakan target prediksi yang ingin ditentukan menggunakan model klasifikasi berdasarkan variabel-variabel lainnya
---	-----------	--

Selanjutnya, data akan diproses menggunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan ikan hias Channa berdasarkan berbagai variabel yang telah dikumpulkan. Algoritma SVM dipilih karena kemampuannya dalam memisahkan data ke dalam kategori yang berbeda dengan margin optimal. Hasil klasifikasi ini akan dianalisis untuk menilai akurasi dan keandalan model dalam mengidentifikasi spesies Channa yang sesuai berdasarkan atribut seperti berat, panjang, warna, bentuk sirip, panjang sirip, bentuk kepala, dan kelayakan.

4.2. Preprocessing Data

Preprocessing data merupakan tahap krusial dalam penelitian ini untuk memastikan bahwa data yang digunakan dalam model *Support Vector Machine* (SVM) berada dalam kondisi terbaik. Tahap ini melibatkan berbagai langkah yang dirancang untuk membersihkan, mengubah, dan menyiapkan data mentah agar sesuai untuk analisis lebih lanjut.

4.3. Mencari Missing Value

Pencarian missing values dilakukan dengan memanfaatkan library Pandas dan NumPy. Berikut adalah kode yang digunakan untuk mendeteksi dan mengatasi missing values tersebut:

```
[ ] df.isnull().sum()
Panjang      0
Berat         0
Warna         0
Bentuk Sirip  0
Panjang Sirip 0
Bentuk Kepala 0
Kelayakan     0
dtype: int64
```

Gambar 2. Mencari missing value

Fungsi `df.isnull().sum()` digunakan untuk menghitung jumlah missing values di setiap kolom. Hasilnya kemudian ditampilkan untuk memberikan gambaran mengenai kolom mana yang memiliki missing values dan seberapa banyak.

4.4. Information Gain

Information Gain digunakan untuk mengevaluasi fitur-fitur dalam dataset ikan hias Channa. Perhitungan Information Gain menghasilkan fitur-fitur yang memiliki pengaruh terbesar terhadap hasil klasifikasi. Setelah menghitung Information Gain, diperoleh sebuah DataFrame yang terdiri dari dua kolom: 'Feature' dan 'Information Gain'. Kolom 'Feature' mencantumkan nama-nama fitur yang telah diubah

melalui One-Hot Encoding, sementara kolom 'Information Gain' menunjukkan nilai Information Gain untuk masing-masing fitur tersebut.

	Feature	Information Gain
0	onehot_Warna_Biru	0.826425
1	onehot_Warna_Coklat	0.820726
2	onehot_Warna_Hitam	0.852799
3	onehot_Warna_Kuning	0.821786
4	onehot_Bentuk_Sirip_Anal	0.819142
5	onehot_Bentuk_Sirip_Caudal	0.818927
6	onehot_Bentuk_Sirip_Dorsal	0.818863
7	onehot_Bentuk_Sirip_Pectoral	0.819417
8	onehot_Bentuk_Kepala_Bulat	0.820288
9	onehot_Bentuk_Kepala_Lebar datar	0.819267
10	onehot_Bentuk_Kepala_Moncong runcing	0.820086
11	onehot_Bentuk_Kepala_Panjang pipih	0.819422
12	remainder_Panjang	0.942221
13	remainder_Berat	0.844051
14	remainder_Panjang Sirip	0.826866

Gambar 3. Information gain

Atribut 'Panjang' memiliki Information Gain tertinggi (0.0285), menunjukkan bahwa 'Panjang' adalah fitur yang paling informatif dalam menentukan kelas 'Kelayakan' dibandingkan atribut lain dalam dataset ini. Atribut 'Berat' memiliki Information Gain berikutnya sebesar (0.0071), tetapi nilainya masih jauh lebih rendah dibandingkan dengan 'Panjang'. Atribut lainnya, seperti 'Warna', 'Bentuk Sirip', 'Panjang Sirip', dan 'Bentuk Kepala', memiliki Information Gain yang sangat rendah, yang menunjukkan bahwa mereka memberikan sedikit informasi dalam menentukan kelas 'Kelayakan'.

4.5. Mutual Information

	Feature	Mutual Information Score
0	onehot_Warna_Biru	0.005389
1	onehot_Warna_Coklat	0.001415
2	onehot_Warna_Hitam	0.027371
3	onehot_Warna_Kuning	0.029722
4	onehot_Bentuk_Sirip_Anal	0.001128
5	onehot_Bentuk_Sirip_Caudal	0.005414
6	onehot_Bentuk_Sirip_Dorsal	0.000000
7	onehot_Bentuk_Sirip_Pectoral	0.000000
8	onehot_Bentuk_Kepala_Bulat	0.000000
9	onehot_Bentuk_Kepala_Lebar datar	0.000842
10	onehot_Bentuk_Kepala_Moncong runcing	0.000000
11	onehot_Bentuk_Kepala_Panjang pipih	0.004672
12	remainder_Panjang	0.036626
13	remainder_Berat	0.000927
14	remainder_Panjang Sirip	0.010571

Gambar 4. Mutual Information

Mutual Information adalah metode lain yang digunakan untuk mengukur hubungan antara fitur dan target dalam dataset. Dalam proses klasifikasi ikan hias Channa menggunakan algoritma *Support Vector Machine* (SVM), Mutual Information untuk setiap fitur akan dihitung. Kemudian, dua DataFrame akan dibuat: satu untuk hasil *Information Gain* dan satu lagi untuk hasil *Mutual Information*. Berdasarkan

perhitungan *Mutual Information*, fitur "Panjang" memiliki nilai tertinggi sebesar 0.019723, yang menunjukkan bahwa panjang ikan hias Channa adalah fitur yang paling berpengaruh dalam menentukan kelayakan ikan tersebut. Fitur ini secara signifikan mengurangi ketidakpastian dan efektif dalam membedakan antara kelas "layak" dan "tidak layak".

4.6. Transformation

Pada tahap transformation, akan dilakukan beberapa proses, antara lain:

4.7. Memisahkan Fitur Dan Label

Pada tahap ini, fitur (X) dan label (y) dalam dataset akan dipisahkan. Kolom 'Kelayakan' berfungsi sebagai label atau target yang akan diprediksi, sehingga kolom ini dipisahkan dari dataset fitur. X mencakup semua kolom selain 'Kelayakan', sementara y berisi nilai-nilai dari kolom 'Kelayakan'.

4.8. One-Hot Encoding

One-Hot Encoding mengonversi setiap nilai kategori menjadi vektor biner. Setiap kategori unik dalam fitur diubah menjadi vektor di mana hanya satu elemen yang bernilai 1 (menunjukkan adanya kategori tersebut), sementara elemen lainnya bernilai 0. Setelah penerapan One-Hot Encoding, setiap fitur kategorikal dalam dataset diubah menjadi beberapa kolom biner, di mana masing-masing kolom mewakili satu kategori dari fitur tersebut. Hasil transformasi ini adalah sebuah matriks yang berisi kombinasi nilai 0 dan 1.

```
Fitur Kategorikal Setelah One-Hot Encoding:
[[ 0.  0.  1. ...  1.  7.  3.]
 [ 1.  0.  0. ... 24. 12.  3.]
 [ 0.  0.  0. ... 19.  6.  2.]
 ...
 [ 0.  0.  0. ... 20. 10.  5.]
 [ 0.  1.  0. ... 12.  8.  4.]
 [ 1.  0.  0. ...  5.  4.  2.]]
```

Gambar 5. One-Hot Encoding

4.9. Membagi Data Training Dan Data Testing

```
print("Jumlah Baris Dan Kolom")
print(" Data Set Awal :", df.shape)
print(" x_train :", x_train.shape)
print(" y_train :", y_train.shape)
print(" x_test :", x_test.shape)
print(" y_test :", y_test.shape)

Jumlah Baris Dan Kolom
Data Set Awal : (700, 7)
x_train : (560, 6)
y_train : (560,)
x_test : (140, 6)
y_test : (140,)
```

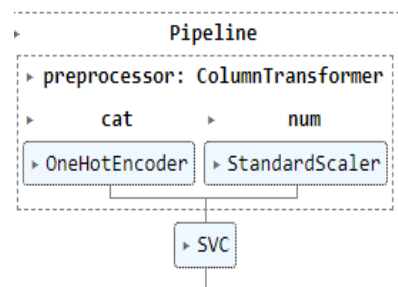
Gambar 6. Pembagian dataset

Dataset akan dibagi menggunakan fungsi `train_test_split` menjadi 2 bagian yaitu menjadi data *training* dan data *testing*.

Data dibagi menjadi 6 bagian untuk pelatihan, dengan total 560 data, dan 2 bagian untuk pengujian, berjumlah 140 data. Proses pengacakan dilakukan sebanyak 42 kali untuk memastikan validitas data dan meningkatkan akurasi model.

4.10. Data Mining

Pada tahap ini, algoritma *Support Vector Machine* diterapkan untuk mengklasifikasikan dataset yang sudah diproses, dengan menggunakan SVC dan kernel linear. Pipeline digunakan untuk mengintegrasikan langkah-langkah *preprocessing* dan pembentukan model ke dalam satu alur kerja yang terstruktur, sehingga proses pelatihan model menjadi lebih efisien dan terkoordinasi. Untuk mendapatkan evaluasi yang lebih akurat dan mengurangi risiko *overfitting*, *cross-validation* dilakukan dengan membagi data latih menjadi beberapa subset (fold). Model kemudian dilatih dan dievaluasi pada subset yang berbeda, memastikan bahwa kinerja model tidak hanya baik pada subset data tertentu, tetapi juga pada data yang belum pernah dilihat sebelumnya.



Gambar 7. Pipeline

Setelah *cross-validation* selesai, model SVM dilatih menggunakan seluruh data yang telah diproses. Pipeline yang telah dibuat menyatukan proses preprocessing dan pelatihan model menjadi satu alur kerja yang terintegrasi.

4.11. Evaluation

Tahapan selanjutnya adalah menguji keakuratan hasil pemodelan yang telah dilakukan.

4.12. Accuracy

Accuracy adalah rasio prediksi benar terhadap total.

```
[14] from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_test, y_pred)
print(f"Akurasi: {accuracy:.2f}")

Akurasi: 0.69
```

Gambar 8. Perhitungan accuracy

4.13. Precision

Precision adalah proporsi prediksi positif yang benar

```
[15] from sklearn.metrics import precision_score
precision = precision_score(y_test, y_pred, average='weighted')
print(f"Precision: {precision:.2f}")
```

➤ Precision: 0.66

Gambar 9. Perhitungan Precision

4.14. Recall

Recall adalah proporsi dari total positif yang terdeteksi dengan benar

```
[16] from sklearn.metrics import recall_score
recall = recall_score(y_test, y_pred, average='weighted')
print(f"Recall: {recall:.2f}")
```

➤ Recall: 0.69

Gambar 10. Perhitungan Recall

4.15. F-1 Score

F1 Score adalah rata-rata dari precision dan recall

```
[17] from sklearn.metrics import f1_score
f1 = f1_score(y_test, y_pred, average='weighted')
print(f"F1-Score: {f1:.2f}")
```

➤ F1-Score: 0.67

Gambar 11. Perhitungan F-1 Score

Perhitungan di atas merupakan hasil dari perhitungan manual yang telah dilakukan. Berikut ini adalah implementasi dari program yang telah dibuat:

```
➤ Akurasi: 0.69
Laporan Klasifikasi:
```

	precision	recall	f1-score	support
Tidak Layak	0.33	0.26	0.29	35
Layak	0.77	0.83	0.80	105
accuracy			0.69	140
macro avg	0.55	0.54	0.54	140
weighted avg	0.66	0.69	0.67	140

Gambar 12. Hasil Evaluasi

4.16. Deployment

Klasifikasi ikan hias Channa diterapkan ke dalam sebuah situs web menggunakan framework Flask. Hal ini memungkinkan pengguna untuk dengan mudah mengevaluasi dan mengklasifikasikan ikan hias melalui antarmuka web. Pengguna dapat mengunggah data ikan hias dan memperoleh hasil klasifikasi dengan cepat dan efisien, sehingga mempermudah proses penilaian dan pengambilan keputusan terkait jenis ikan hias tersebut.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian terkait klasifikasi ikan hias Channa menggunakan algoritma *Support Vector Machine* (SVM), ditemukan bahwa dataset yang terdiri dari 700 data memberikan akurasi yang cukup baik sebesar 69%. Nilai-nilai metrik kinerja model lainnya adalah sebagai berikut : Precision 66%, Recall 69% dan F1-Score 67%.

Meskipun demikian, terdapat ketidakseimbangan dalam kinerja model antara kelas "Layak" dan "Tidak Layak". Kelas "Layak" menunjukkan kinerja yang jauh lebih baik, ditunjukkan oleh nilai precision, recall, dan f1-score yang lebih tinggi. Sebaliknya, rendahnya nilai precision dan recall pada kelas "Tidak Layak" menunjukkan bahwa model mengalami kesulitan dalam mengklasifikasikan data dengan tepat dalam kategori tersebut. Saran untuk penelitian berikutnya mencakup mencoba algoritma klasifikasi lain seperti *Random Forest*, *Gradient Boosting*, atau *Neural Networks* untuk mengevaluasi apakah ada peningkatan kinerja dalam klasifikasi. Selain itu, disarankan untuk menambah variabel baru guna menciptakan data yang lebih seimbang, yang dapat membantu meningkatkan akurasi dan keandalan model dalam mengklasifikasikan kelas yang berbeda.

DAFTAR PUSTAKA

- [1] Rangga Saputra, A., Panji Sasmito, A., & Rudhistiar, D. (2021). RANCANG BANGUN PAKAN DAN FILTERISASI PADA BUDIDAYA IKAN CHANNA MENGGUNAKAN METODE FUZZY BERBASIS ARDUINO. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 5, Issue 2).
- [2] Jamal, B. F., Umar, N. A., & Budi, S. (2022). ANALISIS KANDUNGAN ALBUMIN IKAN GABUS CHANNA STRIATA PADA HABITAT SUNGAI DAN RAWA DI KABUPATEN MAROWALI. *Journal of Aquaculture and Environment*, 5(1), 14–20. <https://doi.org/10.35965/jae.v5i1.1951>
- [3] Santoso, B. B., Ocsa, P., & Saian, N. (2023). Implementasi Flask Framework pada Development Modul Reporting Aplikasi Sistem Informasi Helpdesk di PT.XYZ). *Jurnal Teknologi Informasi Dan Komunikasi*, 7(2), 2023. <https://doi.org/10.35870/jti>
- [4] Wilyani, F., Nuryan Arif, Q., & Aslimar, F. (n.d.). Pengenalan Dasar Pemrograman Python Dengan Google Colaboratory Basic Introduction to Python Programming With Google Colaboratory. *Jurnal Pengabdian Pada Masyarakat Indonesia*. <https://doi.org/10.55606/jppmi.v3i1.1087>
- [5] Riziq sirfatullah Alfarizi, M., Zidan Al-farish, M., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). PENGGUNAAN PYTHON SEBAGAI BAHASA PEMROGRAMAN UNTUK MACHINE LEARNING DAN DEEP

- LEARNING. In Karimah Tauhid (Vol. 2, Issue 1).
- [6] Utomo, Y. B., Kurniasari, I., & Yanuartanti, I. (2023). PENERAPAN KNOWLEDGE DISCOVERY IN DATABASE UNTUK ANALISA TINGKAT KECELAKAAN LALU LINTAS. *Jurnal Teknik Informatika Kaputama (JTIK)*, 7(1).
- [7] Oktavia, D., & Ramadahan, Y. R. (2023). Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Media Online*, 4(1), 407–417. <https://doi.org/10.30865/klik.v4i1.1040>
- [8] Akira Permata Ramadhani, Eka Dyar Wahyuni, & Amalia Anjani Arifiyanti. (2024). Klasifikasi Cyberbullying pada Komentar Instagram dengan Menggunakan Supervised Learning. *Neptunus: Jurnal Ilmu Komputer Dan Teknologi Informasi*, 2(2), 92–101. <https://doi.org/10.61132/neptunus.v2i2.108>
- [9] Sufardin, Sriwulan, & Anshary, H. (2022). Gyrodactylus (Monogenea: Gyrodactylidae) on marine ornamental fish *Amphiprion percula* from a marine aquaculture facility in Indonesia. *Biodiversitas*, 23(2), 1023–1030. <https://doi.org/10.13057/biodiv/d230245>
- [10] Malick, R. C., Pal, B. B., Bera, A. K., Mohanty, D., & Das, B. K. (2022). Molecular characterisation of *Flavobacterium columnare* associated with columnaris disease in freshwater fishes with emphasis on virulence and antibiotic resistance pattern. *Indian Journal of Fisheries*, 69(3), 91–99. <https://doi.org/10.21077/ijf.2022.69.3.122636-11>