



## KLASIFIKASI BUAH JERUK LEMON BERDASARKAN TINGKAT KEMATANGAN MENGGUNAKAN METODE SVM DAN NAIVE BAYES

**Desti Mualfah<sup>\*1)</sup>, Hardi Rivaldi<sup>2)</sup>, Januar Al Amin<sup>3)</sup>, Sunanto<sup>4)</sup>**

<sup>1,2,3,4</sup>Ilmu Komputer, Ilmu Komputer, Universitas Muhammadiyah Riau

email: [destimualfah@umri.ac.id](mailto:destimualfah@umri.ac.id), [200401148@student.umri.ac.id](mailto:200401148@student.umri.ac.id), [januaralamin@umri.ac.id](mailto:januaralamin@umri.ac.id),  
[sunanto@umri.ac.id](mailto:sunanto@umri.ac.id)

### *Abstract*

*This study aims to develop a classification model for determining the ripeness level of lemons (Citrus limon) using digital image analysis. Two methods, namely Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC), were compared to evaluate their performance in terms of accuracy and prediction consistency. The results show that SVM outperformed NBC with an accuracy of 97%, along with precision, recall, and F1-Score of 97% each. The model consistently determined lemon ripeness levels in percentage terms, such as 85% or 95%. In contrast, NBC achieved an accuracy of 82%, with precision, recall, and F1-Score of 83%, 82%, and 83%, respectively. However, NBC was more prone to classification errors, especially in distinguishing between ripe and unripe lemons. In conclusion, the SVM method proved superior to NBC in determining lemon ripeness levels, particularly in handling complex data. SVM's ability to provide accurate and consistent predictions makes it a more effective choice for helping farmers optimize the quality and quantity of lemon production. This study contributes significantly to the application of image processing technology in the agricultural sector.*

**Keywords:** Support Vector Machine (SVM), Naïve Bayes Classifier (NBC), Lemon Ripeness Level, Digital Image Analysis, Classification Accuracy and Consistency.

### **Abstrak**

Penelitian ini bertujuan untuk mengembangkan model klasifikasi tingkat kematangan jeruk lemon (Citrus limon) menggunakan analisis citra digital. Dua metode dibandingkan, yaitu Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC), untuk mengevaluasi performanya dalam akurasi dan konsistensi prediksi. Hasil penelitian menunjukkan bahwa SVM memiliki keunggulan dengan akurasi 97% serta presisi, recall, dan F1-Score masing-masing sebesar 97%. Model ini mampu menentukan tingkat kematangan lemon dalam bentuk persentase, seperti 85% atau 95%, secara konsisten. Sebaliknya, NBC mencatat akurasi 82%, dengan presisi, recall, dan F1-Score sebesar 83%, 82%, dan 83%. Namun, NBC lebih rentan terhadap kesalahan klasifikasi, terutama dalam membedakan lemon matang dan belum matang. Kesimpulannya, metode SVM lebih unggul dibandingkan NBC dalam menentukan tingkat kematangan jeruk lemon, terutama pada data dengan kompleksitas tinggi. Kemampuan SVM untuk memberikan prediksi yang akurat dan konsisten menjadikannya pilihan yang lebih efektif dalam mendukung petani mengoptimalkan kualitas dan kuantitas produksi jeruk lemon. Penelitian ini memberikan kontribusi signifikan terhadap penerapan teknologi pengolahan citra dalam sektor pertanian.

**Keywords:** Support Vector Machine (SVM), Naïve Bayes Classifier (NBC), Tingkat Kematangan Jeruk Lemon, Analisis Citra Digital, Klasifikasi Akurasi dan Konsistensi.

### **PENDAHULUAN**

Dalam era globalisasi kemajuan teknologi sangat pesat, teknologi informasi dan pattern recognition menjadi

bagian yang tidak terpisahkan dari kehidupan sehari-hari. Kemajuan dalam bidang komputer, pengolahan citra, kecerdasan buatan, dan statistik telah membuka peluang dalam memahami dan memanfaatkan pola-

pola variabel yang terkandung dalam data (Salim et al., 2023).

Citra merupakan salah satu bentuk informasi yang di perlukan manusia selain teks, suara dan video. Informasi yang terkandung dalam sebuah citra dapat diinterpretasikan berbeda-beda oleh satu dengan yang lain. Menurut safrizal dalam (Astrianda, 2020). menyebutkan bahwa pengembangan metode pengolahan citra untuk mengklasifikasikan kematangan jeruk lemon menggunakan komputer. Sebelum system pemrosesan gambar diterapkan, identifikasi kematangan jeruk manusia. Identifikasi kematangan jeruk manual memiliki banyak kelemahan yang disebabkan oleh banyak faktor, seperti kelelahan, kurang motivasi, pengalaman, kemahiran dan sebagainya. Untuk mengatasi masalah ini perkembangan dalam pengolahan citra telah melibatkan kemajuan teknologi komputasi terutama dalam pengenalan pola, analisis objek, dan manipulasi gambar. Pengolahan citra warna melibatkan peningkatan teknik analisis dan manipulasi gambar berwarna. Ini berperan penting dalam berbagai aplikasi termasuk fotografi, grafika komputer, identifikasi objek, pencitraan medis, dan pengenalan pola. Teknologi ini berdampak besar pada industri hiburan, kesehatan, otomasi, dan banyak sektor lainnya (Hidayatullah, 2020).

Jeruk lemon merupakan kelompok buah jeruk yang mempunyai rasa sangat asam. Jeruk lemon memiliki nama ilmiah yaitu Citrus limon. Di Indonesia disebut dengan jeruk sitrun atau jeruk limun. Jeruk lemon (Citrus limon L). Berasal dari daerah Asia khususnya di Birma bagian Utara dan Cina Selatan. Penyebaran jeruk lemon (Citrus limon L) di Indonesia berada di Jawa dan telah dikembangkan. Di dataran rendah hingga ketinggian 800 meter di atas permukaan laut jeruk lemon (Citrus limon L) bisa tumbuh dengan baik (Sihotang & Donna bontry july, n.d.).

Lemon memiliki karakteristik yang membulat atau seperti tabung dengan diameter sekitar 50-60 mm, terdapat tonjolan pada ujungnya serta mempunyai rasa masam dan aroma yang khas (Arifandi et al., 2021). Tumbuhan lemon merupakan perdu atau pohon yang daunnya berbentuk oval, sayap daun sempit/marginal, warna bunganya kemerahan disertai dengan stamens yang banyak, buahnya berwarna kuning dengan bentuk membundar (panjang 8-9 cm), kulitnya kasar dan rasanya asam, bijinya kecil dengan bentuk ovoid, serta permukaan biji yang halus. Buah lemon memiliki kulit yang mengkilap, lemon yang masih muda akan berwarna hijau, lemon yang sudah matang akan berwarna kuning cerah, dan lemon yang sudah tua akan berwarna kuning tua, keriput dan lembek ketika ditekan (Kartikawati & Yudi, 2019). Lemon mengandung berbagai senyawa kimia penting yang dapat dimanfaatkan dalam dunia kesehatan. Senyawa yang terkandung dalam lemon antara lain asam sitrat, asam askorbat, mineral, dan flavonoid (Harahap et al., 2021). Menurut Dr. Michael Petani

sering kali mengandalkan pengalaman dan intuisi dalam memanen lemon, tetapi metode ini tidak selalu konsisten, yang dapat mengakibatkan buah yang tidak matang sempurna atau yang kehilangan kualitasnya jika dipanen terlalu lama. Dari pernyataan tersebut menunjukkan pentingnya pengembangan metode klasifikasi yang lebih sistematis dan akurat untuk menentukan tingkat kematangan lemon. Oleh karena itu, penelitian ini berfokus pada pengembangan model klasifikasi yang dapat membantu petani mengoptimalkan mana warna lemon yang sudah siap panen, sehingga kualitas dan kuantitas produksi lemon dapat ditingkatkan. Pemanfaatan teknologi citra pada buah jeruk lemon ini dapat membandingkan dengan metode Naïve Bayes dan Support Vector Machine karena metode ini baik untuk menentukan tingkat kematangan buah jeruk lemon berdasarkan warna dalam akurasi yang akurat. Beberapa penelitian sebelumnya pernah membahas klasifikasi buah jeruk berdasarkan Tingkat kematangan metode SVM. Dengan metode SVM untuk buah jeruk terdiri dari 100 citra jeruk dan hasil klasifikasi SVM menggunakan fitur warna mendapatkan nilai hasil akurasi sebesar 100%. Maka peneliti ingin menerapkan klasifikasi buah jeruk lemon berdasarkan Tingkat kematangan menggunakan metode SVM dan Naive Bayes dengan cara memabnndingkan antara 2 metode dengan hasil akurasi yang (Hughes, 2023). Naïve bayes classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistic yang dikemukakan oleh ilmuwan inggris Thomas Bayes. Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian (Arifin & Sasongko, 2019).

Support Vector Machine merupakan salah satu metode Klasifikasi yang baik dalam pemecahan masalah untuk dua kelas, penelitian yang melibatkan kasus nyata cenderung mencakup kasus multiclass. Sehingga metode SVM untuk dapat menyelesaikan masalah pada penelitian ini terutama pada pangan, khususnya klasifikasi tingkat kematangan pada buah jeruk lemon (Amrozi et al., 2022a). Penelitian ini bertujuan untuk menentukan kematangan buah jeruk menggunakan metode SVM dan NBC, dua metode ini akan ditentukan akurasi yang mana akurasi paling tinggi akan menentukan arah kematangan buah jeruk lemon (D. Muallfah, dkk. 2019).

## METODOLOGI PENELITIAN



Gambar 1. Metodologi Penelitian

Penelitian ini dilakukan dengan menyusun langkah-langkah yang diambil secara sistematis. Hal ini dapat digunakan sebagai panduan yang jelas untuk menyelesaikan masalah, menganalisis hasil penelitian, dan mengatasi masalah yang muncul.

### 2.1. Literature Review

Pada tahap ini, dilakukan pencarian referensi penelitian terkait studi klasifikasi Tingkat kematangan buah jeruk lemon dengan menggunakan model Support Vector Machine dan Naïve Bayes. Sumber referensi berasal dari sejumlah jurnal penelitian sebelumnya. Dari referensitersebut, penulis mengekstraksi teori dasar dan juga mengambil kutipan yang dianggap relevan dengan konteks penelitian yang sedang dilakukan.

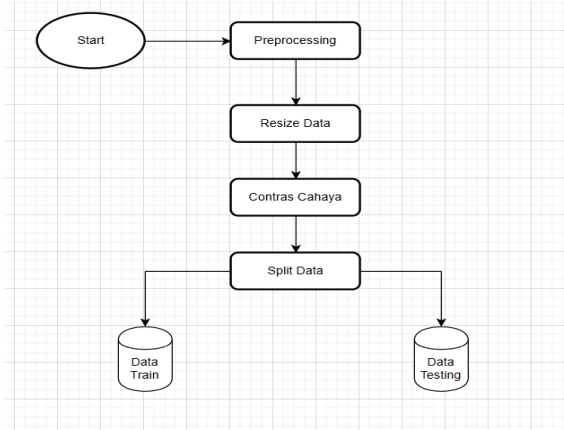
### 2.2. Pengumpulan Dataset



Gambar 2 Sample Dataset

Dalam tahap ini dilakukan pengumpulan dataset pada penelitian ini bersumber dari data publik <https://www.kaggle.com/datasets/yusufemir/lemon-quality-dataset>. Dataset yang digunakan dalam penelitian ini yaitu tiga jenis tingkat kematangan pada buah jeruk lemon yang diambil dari sample data public kaggle berupa pengambilan foto citra buah jeruk lemon untuk menentukan tingkat kematangannya melalui warna buah jeruk lemon yang terdiri yaitu 500 buah jeruk lemon matang, 500 buah jeruk lemon belum matang, dan 500 buah jeruk lemon terlalu matang dengan total dataset 1.500 buah jeruk lemon.

### 2.3. Preprocessing Data



Gambar 3 Alur Preprocessing Data

Proses ini mencakup serangkaian langkah untuk membersihkan, mentransformasi, dan mempersiapkan data agar sesuai dengan kebutuhan analisis yang akan dilakukan. Tujuannya adalah untuk meningkatkan kualitas data, mengurangi noise, dan mempersiapkan dataset agar sesuai dengan persyaratan model yang akan digunakan. Dengan melakukan preprocessing yang tepat, data menjadi lebih siap untuk digunakan dalam proses Machine Learning, yang pada gilirannya dapat menghasilkan model yang lebih akurat.

### 2.4. Klasifikasi Metode SVM dan Naïve Bayes

#### A. Support Vector Machine (SVM)

SVM merupakan metode pembelajaran linier yang menemukan hyperline optimal untuk memisahkan dua kelas (positif dan negatif). SVM memiliki kemampuan dalam menerapkan pemisah non-linier berdimensi tinggi yang membuat SVM menjadi istimewa dengan menggunakan fungsi kernelnya. Terdapat beberapa kernel SVM yang dapat digunakan, namun didalam suatu penelitian, kernel yang sering digunakan ialah linear, polynominal, radial basis fuction (RBF). Proses latih algoritma pembelajaran adalah proses optimasi dimana learning bias diimplementasikan. SVM memiliki untuk kerja yang baik untuk diimplementasikan pada bioinformatics, pengenalan, tulisan tangan, klasifikasi dan lain sebagainya.

#### B. Naïve Bayes

Naïve Bayes Clasisifier (NBC) merupakan pengklasifikasian probabilitas sederhana berdasarkan pada teorema bayes. Teorema “bayes” dikombinasikan dengan “Naïve” yang berarti setiap atribut atau variabel bersifat bebas (independent). Naïve Bayes Clasisifier dapat dilatih dengan efisien dalam pembelajaran terawasi, proses Naïve Bayes Classifier mengamsumsikan bahwa ada atau tidaknya suatu fitur pada suatu kelas tidak berhubungan atau tidaknya fitur lain dikelas yang sama. Naïve Bayes Classifier

termasuk ke dalam pembelajaran supervised, sehingga pada tahapan pembelajaran dibutuhkan data awal berupa data pelatihan untuk bisa dilakukan pengambilan data Keputusan. Rumus dari perhitungan Naïve bayes dalam Data Mining Concept and Techniques yaitu :

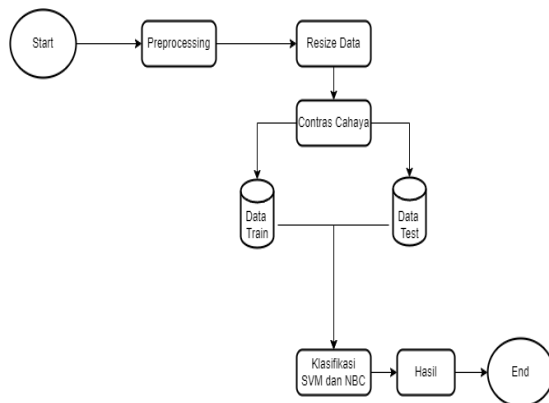
$$P(X) = \frac{P(C)P(C)}{P(X)}$$

Atau dapat ditulis dengan:

$$Posterior = \frac{likelihood \times prior\ probability}{evidence}$$

Keterangan :  $P(C_i|X)$  = Peluang kategori  $C_i$  jika diberikan fakta atau bukti  $X$  (posterior)  $P(X|C_i)$  = Peluang pada kategori  $C_i$ , dimana fakta atau bukti  $X$  muncul pada kategori tersebut (likelihood)  $P(C_i)$  = Peluang dari kategori yang diberikan, dibandingkan dengan kategori lainnya yang dianalisa (prior probability)  $P(X)$  = Jumlah peluang dari fakta atau bukti  $X$  (evidence).

Alur kerja dari naïve bayes adalah sebagai berikut : Baca data training Hitung jumlah peluang untuk setiap variabel, cari nilai likelihood dan nilai probabilitas. Mendapatkan nilai dalam tabel mean, standar deviasi, dan likelihood serta probabilitas



Gambar 4 Alur Kerja Naïve Bayes

## HASIL DAN PEMBAHASAN

### 3.1. Pengumpulan Dataset

Berdasarkan perancangan yang telah dijelaskan, pengumpulan data dalam penelitian ini dilakukan dengan mengunduh dataset dari sumber data publik yang tersedia di Kaggle <https://www.kaggle.com/datasets/yusufemir/lemon-quality-dataset>. Dataset tersebut terdiri dari gambar buah jeruk lemon dengan tiga kategori tingkat kematangan yang berbeda, yaitu matang, setengah matang, dan tidak matang.

Tabel 1 Dataset

| Citra                                                                              | Label          | Jumlah |
|------------------------------------------------------------------------------------|----------------|--------|
|  | Belum Matang   | 500    |
|  | Matang         | 500    |
|  | Terlalu Matang | 500    |

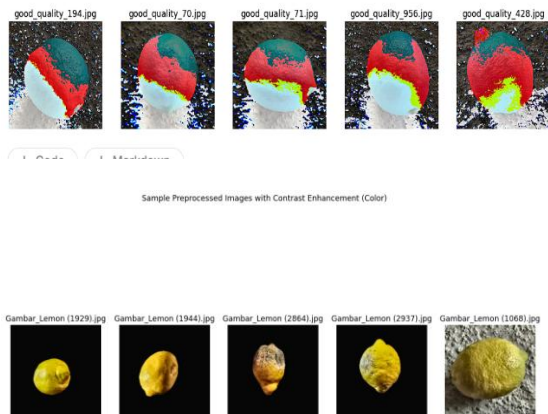
### 3.2. Preprocessing data

Setelah dataset dikumpulkan, langkah selanjutnya adalah melakukan preprocessing data untuk memastikan bahwa data tersebut siap digunakan dalam analisis dan pengembangan model Machine Learning. Proses preprocessing ini mencakup beberapa tahap penting, termasuk pembersihan data dari anomali atau noise, normalisasi nilai piksel pada citra agar konsisten, serta transformasi data ke dalam format yang sesuai dengan kebutuhan model. Data citra lemon yang diperoleh kemudian diubah ukurannya agar seragam, serta dilakukan augmentasi data seperti rotasi, flipping, atau perubahan kecerahan untuk memperkaya variasi dalam dataset. Preprocessing ini juga melibatkan pembagian dataset ke dalam set pelatihan, validasi, dan pengujian guna memastikan bahwa model yang dikembangkan dapat mengevaluasi performanya secara objektif. Dengan melakukan preprocessing data yang komprehensif dan tepat, kualitas dataset ditingkatkan, sehingga dapat mendukung pengembangan model yang lebih akurat dan robust.

### 3.3 Resize Data

Salah satu langkah penting dalam preprocessing data adalah mengubah ukuran (resize) citra agar seragam dan sesuai dengan persyaratan input model Machine Learning yang akan digunakan. Dalam penelitian ini, semua citra buah lemon diubah ukurannya menjadi dimensi yang telah ditentukan, dengan ukuran 224x224 piksel. Ukuran ini dipilih karena sesuai dengan input standar pada arsitektur model yang akan digunakan dan memungkinkan pemrosesan data yang efisien tanpa mengorbankan terlalu banyak detail visual yang penting untuk prediksi tingkat kematangan.

### 3.4. Contras Cahaya



Gambar 5 *Preprocessing Data*

Peningkatan kontras cahaya merupakan langkah penting dalam *preprocessing* data gambar, khususnya dalam aplikasi *machine learning* yang bertujuan untuk mendeteksi pola atau fitur spesifik pada citra. Dalam penelitian ini, teknik Contrast Limited Adaptive Histogram Equalization (CLAHE) digunakan untuk meningkatkan kontras pada citra buah lemon. CLAHE bekerja dengan memperbaiki perbedaan intensitas piksel dalam gambar secara adaptif, sehingga detail-detail penting yang mungkin tidak terlihat pada gambar asli menjadi lebih jelas dan terdefinisi. Dengan menerapkan CLAHE, citra buah lemon yang memiliki variasi intensitas cahaya atau bayangan yang tidak merata dapat diperbaiki, sehingga pola kematangan yang ingin dideteksi oleh model menjadi lebih mudah diidentifikasi. Proses ini tidak hanya meningkatkan kualitas visual gambar, tetapi juga memperkaya informasi yang dapat diolah oleh model machine learning, sehingga hasil analisis dan prediksi menjadi lebih akurat.

### 3.5. Split data

Setelah data citra buah lemon selesai melalui proses preprocessing, langkah berikutnya adalah membagi dataset tersebut menjadi dua bagian utama: data training dan data testing. Pembagian ini penting untuk memastikan bahwa model machine learning dapat dilatih dengan baik dan diuji secara efektif. Data training adalah bagian dari dataset yang digunakan untuk melatih model. Dalam penelitian ini, sekitar 80% dari total dataset digunakan sebagai data training. Data ini mencakup gambar-gambar lemon dalam berbagai tingkat kematangan, yang memungkinkan model untuk belajar mengenali pola dan fitur penting yang terkait dengan setiap kategori kematangan. Sebaliknya, data testing merupakan 20% dari dataset yang tidak digunakan dalam proses pelatihan model. Data ini disimpan untuk menguji kinerja model setelah

pelatihan selesai. Dengan menggunakan data testing yang belum pernah dilihat oleh model selama pelatihan, kita dapat mengevaluasi seberapa baik model dapat menggeneralisasi dan mengklasifikasi citra buah lemon yang baru. Proses pembagian ini dilakukan menggunakan fungsi `'train_test_split'`, yang secara acak membagi data dengan rasio 80:20, memastikan distribusi yang seimbang antara data training dan data testing.

```
# Split dataset into training and testing sets
X_train, X_test, y_train, y_test, paths_train, paths_test = train_test_split(
    features, encoded_labels, image_paths, test_size=0.2, random_state=42
)
```

### 3.6. Klasifikasi SVM dan Naïve Bayes

#### 1. Support vector machine

Metode ini sangat efektif dalam situasi di mana data memiliki dimensi tinggi dan dapat menggunakan berbagai fungsi kernel untuk menangani non-linearitas dalam data. Dalam penelitian ini, kernel linear digunakan karena kesederhanaannya dan kemampuannya dalam menangani data yang relatif tidak kompleks. Model SVM dilatih menggunakan 80% dari dataset sebagai data training, dan hasil klasifikasinya dievaluasi menggunakan data testing yang tersisa. Hasil evaluasi menunjukkan akurasi, confusion matrix, dan loss dari model SVM, memberikan gambaran mengenai efektivitas model dalam mengklasifikasikan tingkat kematangan buah lemon. Berikut adalah kode implementasi SVM:

# Klasifikasi menggunakan SVM

```
svm_classifier = SVC(kernel='linear',
    probability=True, random_state=42)
```

```
svm_classifier.fit(X_train, y_train)
```

```
y_pred_svm = svm_classifier.predict(X_test)
```

```
y_pred_svm_proba = svm_classifier.predict_proba(X_test)
```

2. Metode ini menggunakan teorema Bayes untuk menghitung probabilitas kelas berdasarkan fitur yang ada. Naïve Bayes terkenal dengan kemampuannya untuk melakukan klasifikasi dengan cepat dan efisien, meskipun dengan asumsi independensi fitur. Dalam penelitian ini, model Naïve Bayes dilatih menggunakan data training yang sama dengan model SVM, dan dievaluasi menggunakan data testing. Evaluasi mencakup akurasi, confusion matrix, dan loss dari model NBC, memberikan indikasi mengenai performa model dalam klasifikasi. Berikut adalah kode implementasi Naïve Bayes:

# Klasifikasi menggunakan Naïve Bayes

```
nb_classifier = GaussianNB()
```

```
nb_classifier.fit(X_train, y_train)
```

```
y_pred_nb = nb_classifier.predict(X_test)
```

```
y_pred_nb_proba
```

```
nb_classifier.predict_proba(X_test)
```

### 3.7. Hasil

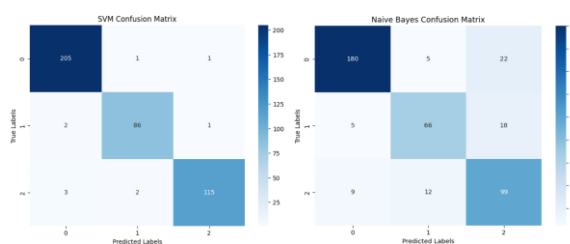
Setelah menerapkan model Support Vector Machine (SVM) dan Naïve Bayes pada dataset citra buah lemon dari Kaggle, hasil klasifikasi dari kedua model dievaluasi untuk menentukan kinerjanya. Pengujian dilakukan dengan mengukur tingkat akurasi serta menganalisis confusion matrix untuk masing-masing model. Hasil pengujian menunjukkan seberapa baik setiap model dalam mengklasifikasikan tingkat kematangan buah lemon. Selain itu, visualisasi dari hasil pemodelan memberikan wawasan yang berharga mengenai performa model, memungkinkan identifikasi area untuk perbaikan dan penyempurnaan. Dengan analisis ini, kita dapat mengevaluasi efektivitas masing-masing metode dan memilih model yang paling sesuai untuk aplikasi praktis dalam klasifikasi kematangan buah lemon. Berikut adalah Hasil dari SVM dan NBC:

#### A. Akurasi, Presisi, Recall, Dan F-1 Score

Tabel 2. Pencapaian Hasil SVM dan Naïve Bayes

| Algoritma | Akurasi | Loss | Presisi | Recall | F-1 Score |
|-----------|---------|------|---------|--------|-----------|
| SVM       | 97%     | 0.06 | 97%     | 97%    | 97%       |
| NBC       | 82%     | 6.15 | 83%     | 82%    | 83%       |

#### B. Confusion Matrix

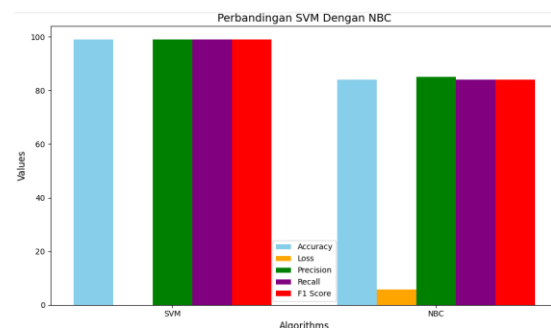


Gambar 6 Confusion Matrix NBC dan SVM

Untuk model SVM, confusion matrix menunjukkan hasil  $\begin{bmatrix} 205 & 1 & 1 \\ 2 & 86 & 1 \\ 3 & 2 & 115 \end{bmatrix}$ . Matrix ini menunjukkan bahwa SVM hampir sempurna dalam mengklasifikasikan buah lemon ke dalam kategori yang benar. Model ini berhasil mengidentifikasi 205

citra lemon matang, 86 citra lemon setengah matang, dan 115 citra lemon tidak matang tanpa melakukan kesalahan klasifikasi yang berarti pada dua kategori pertama. Hanya ada 3 kesalahan dalam mengklasifikasikan lemon matang sebagai lemon tidak matang dan 2 kesalahan dalam mengklasifikasikan lemon tidak matang sebagai lemon matang. Kesalahan yang sangat sedikit menunjukkan bahwa SVM sangat efektif dalam menangani dataset ini, memberikan akurasi dan presisi yang sangat tinggi dalam klasifikasi. Sebaliknya, confusion matrix untuk model Naïve Bayes adalah  $\begin{bmatrix} 180 & 5 & 22 \\ 5 & 66 & 18 \\ 9 & 12 & 99 \end{bmatrix}$ . Meskipun model ini juga mampu mengklasifikasikan citra lemon dengan cukup baik, terdapat lebih banyak kesalahan dibandingkan dengan SVM. Model ini mengalami kesulitan yang signifikan dalam mengklasifikasikan lemon matang dan lemon tidak matang, dengan 22 citra lemon matang yang salah diklasifikasikan sebagai lemon tidak matang dan 66 citra lemon tidak matang yang salah diklasifikasikan sebagai lemon matang. Selain itu, terdapat 9 kesalahan dalam mengklasifikasikan lemon setengah matang, yang menunjukkan bahwa model ini kurang efisien dalam membedakan antara kategori yang lebih halus. Perbedaan ini mengindikasikan bahwa SVM memiliki kemampuan yang jauh lebih baik dalam mengidentifikasi dan memisahkan kelas-kelas yang berbeda secara akurat, sementara Naïve Bayes menghadapi kesulitan yang lebih besar dalam menghadapi data yang kompleks dan sering kali salah mengklasifikasikan citra di antara kategori yang serupa. Dengan demikian, SVM terbukti lebih efektif dalam mengklasifikasikan tingkat kematangan buah lemon dibandingkan dengan Naïve Bayes, mendemonstrasikan kekuatan SVM dalam menangani dataset yang memiliki variabilitas tinggi dan kompleksitas yang lebih besar.

#### C. Grafik



Gambar 7 Grafik Perbandingan SVM dan NBC

Grafik batang ini menyajikan perbandingan kinerja dua algoritma pembelajaran mesin, yaitu Support Vector Machine (SVM) dan Naive Bayes Classifier (NBC), berdasarkan lima metrik evaluasi utama: Accuracy, Loss, Precision, Recall, dan F1 Score.

#### D. Hasil Klasifikasi



Gambar 8 Hasil Klasifikasi

Setelah menerapkan model Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC) pada dataset citra untuk mengklasifikasikan tiga kategori berbeda, hasil klasifikasi dari kedua model dianalisis untuk mengevaluasi kinerjanya. Pengujian dilakukan dengan memanfaatkan hasil visualisasi prediksi dan confusion matrix guna memahami efektivitas masing-masing algoritma dalam melakukan klasifikasi. Berdasarkan hasil pengujian, SVM menunjukkan performa yang jauh lebih unggul dibandingkan NBC.

#### E. Perbandingan Hasil Uji Split Data

Perbandingan antara penggunaan data dengan ukuran hasil praproses resize sebesar 224x224 pada pembagian data uji sebesar 80:20, dan data dengan ukuran resize 299x299 pada pembagian data uji sebesar 90:10 menghasilkan hasil evaluasi sebagai berikut:

Tabel 2 Hasil Perbandingan uji split data 80:20 dan 90:10 svm dan naive bayes

| Metrik   | SVM (224x224, 80:20) | SVM (299x299, 90:10) | Naive Bayes (224x224, 80:20) | Naive Bayes (299x299, 90:10) |
|----------|----------------------|----------------------|------------------------------|------------------------------|
| Akurasi  | 0.97596 (97.60%)     | 0.80529 (80.53%)     | 0.82933 (82.93%)             | 0.67548 (67.55%)             |
| Loss     | 0.0618               | 0.4456               | 6.1517                       | 11.6526                      |
| Presisi  | 0.9761 (97.61%)      | 0.8034 (80.34%)      | 0.8373 (83.73%)              | 0.6729 (67.29%)              |
| Recall   | 0.97596 (97.60%)     | 0.80529 (80.53%)     | 0.82933 (82.93%)             | 0.67548 (67.55%)             |
| F1 Score | 0.97592 (97.59%)     | 0.8042 (80.42%)      | 0.8314 (83.14%)              | 0.6725 (67.25%)              |

Berdasarkan hasil pengujian didapatkan beberapa output berupa:

##### 1. SVM:

Pada eksperimen kedua (299x299, 90:10), performa SVM menurun dibandingkan dengan eksperimen pertama (224x224, 80:20) di semua metrik (akurasi, presisi, recall, dan F1 Score). Penurunan ini dapat disebabkan oleh perubahan rasio data training/testing

atau ukuran gambar yang lebih besar, yang mungkin meningkatkan kompleksitas data. Penurunan Akurasi: 97.60% → 80.53% (penurunan 17.07%).

##### 2. Naive Bayes:

Naive Bayes juga mengalami penurunan performa yang signifikan pada eksperimen kedua. Akurasi menurun dari 82.93% menjadi 67.55%, dengan F1 Score turun dari 83.14% menjadi 67.25%. Penurunan Akurasi: 82.93% → 67.55% (penurunan 15.38%).

##### 3. Perbandingan Antar Model:

SVM consistently outperforms Naive Bayes dalam kedua eksperimen, menunjukkan bahwa SVM lebih mampu menangani data dengan dimensi lebih tinggi (224x224 atau 299x299) dan lebih kompleks. Perbedaan performa antara SVM dan Naive Bayes semakin signifikan pada eksperimen kedua, di mana Naive Bayes menunjukkan kesulitan yang lebih besar dalam menangani data yang di-resize menjadi 299x299.

#### SIMPULAN

Berdasarkan hasil evaluasi dan analisis dari model klasifikasi Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC) memiliki keterbatasan data set yang bersumber dari kagle dengan sample tingkat kematangan pada jeruk lemon, beberapa kesimpulan dapat diambil:

1. SVM menunjukkan performa yang sangat superior dibandingkan dengan Naïve Bayes dalam mengklasifikasikan tingkat kematangan buah jeruk lemon berdasarkan citra digital. Dengan akurasi mencapai 97%, serta presisi, recall, dan F1-Score masing-masing sebesar 97%, SVM membuktikan kemampuannya dalam mengidentifikasi dan memisahkan kategori dengan sangat tepat. Selain itu, model ini juga berhasil menunjukkan tingkat kematangan tiap citra dalam bentuk persentase, seperti tingkat kematangan 80%, 95%, atau 100% berdasarkan analisis warna dominan dan distribusi piksel pada citra. Sebagai contoh, untuk citra buah jeruk lemon tertentu, model mampu menentukan bahwa tingkat kematangan berada pada 85%, sementara untuk citra lain tingkat kematangan mencapai 98%.

2. Sebaliknya, Naïve Bayes memiliki akurasi sebesar 82%, dengan presisi, recall, dan F1-Score masing-masing 83%, 82%, dan 83%. Meskipun masih memberikan hasil yang dapat diterima, NBC menunjukkan kesulitan dalam membedakan kategori yang memiliki kesamaan visual. Confusion matrix untuk NBC menunjukkan lebih banyak kesalahan klasifikasi, terutama dalam membedakan antara kategori lemon matang dan tidak matang. Namun, kemampuan model ini dalam mendeteksi tingkat

kematangan dalam bentuk persentase masih dapat ditingkatkan melalui optimasi fitur. Sebagai contoh, prediksi tingkat kematangan untuk citra tertentu oleh NBC berada pada rentang 70% hingga 90%, namun cenderung kurang konsisten dibandingkan SVM. Dari hasil evaluasi tersebut, dapat disimpulkan bahwa metode SVM lebih unggul dibandingkan dengan Naïve Bayes dalam menentukan tingkat kematangan buah jeruk lemon. SVM lebih efektif dalam menangani data dengan kompleksitas tinggi dan menghasilkan performa yang lebih konsisten dibandingkan NBC. Kemampuan SVM dalam mendeteksi tingkat kematangan dalam bentuk persentase memberikan nilai tambah yang signifikan untuk aplikasi praktis.

3, Penelitian ini juga membantu para petani dalam mengoptimalkan warna lemon yang sudah siap panen.

#### DAFTAR PUSTAKA

- Amrozi, Y., Yuliati, D., Susilo, A., Novianto, N., & Ramadhan, R. (2022a). Klasifikasi Jenis Buah Pisang Berdasarkan Citra Warna dengan Metode SVM. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 11(3), 394–399. <https://doi.org/10.32736/sisfokom.v11i3.1502>.
- Arifandi, M. R., Zahra, N., & Juliastuti, E. (2021). Rancang Bangun Sistem Penentuan Mutu Buah Lemon Berbasis Pengolahan Citra. *Jurnal Otomasi Kontrol Dan Instrumentasi*, 13(2), 117–123. <https://doi.org/10.5614/joki.2021.13.2.7>
- Astrianda, N. (2020). Klasifikasi Kematangan Buah Tomat Dengan Variasi Model Warna Menggunakan Support Vector Machine. *VOCATECH: Vocational Education and Technology Journal*, 1(2), 45–52. <https://doi.org/10.38038/vocatech.v1i2.27>
- D. Mualfah, A. Prihatin, R. Firdaus, and Sunanto, “Analisis Sentimen Masyarakat terhadap Kasus Pembobolan Data Nasabah Bank BSI pada Twitter Menggunakan Metode Random Forest dan Naïve Bayes,” *J. FASILKOM*, vol. 13, no. 3, pp. 614–620, 2023, Accessed: Mar. 06, 2024. [Online]. Available: <https://ejurnal.umri.ac.id/index.php/JIK/article/view/6478>
- D. Mualfah, Ramadhoni, R. Gunawan, dan D. Mulyadipa Suratno, “Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM,” *J. Fasilkom*, vol. 13, no. 01, hal. 72–80, 2023, doi: 10.37859/jf.v13i01.4920.
- Harahap, I. S., Halimatussakdiah, H., & Amna, U. (2021). Skrining Fitokimia Ekstrak Daun Jeruk Lemon (*Citrus limon L.*) dari Kota Langsa, Aceh. *QUIMICA: Jurnal Kimia Sains Dan Terapan*, 3(1), 19–23. <https://doi.org/10.33059/jq.v3i1.3492>
- Hidayatullah, P. (2020). Pengolahan Citra Digital.
- Salim, M. R., Daulay, A. A. A., & Sinaga, G. V. (2023). Pengenalan Pola Garis Telapak Tangan Menggunakan Metode Fuzzy K-Nearest Neighbor. *Bulletin of Computer Science Research*, 3(4), 270–275.
- Sanjaya, C. B., & Rosadi, M. I. (2019). Klasifikasi buah jeruk lemon berdasarkan tingkat kematangan menggunakan least-squares support vector machine. *Explore IT: Jurnal Keilmuan Dan Aplikasi Teknik Informatika*, 10(2), 1–8.
- Talib, S., Sudin, S., & Suratin, M. D. (n.d.). PENERAPAN METODE SUPPORT VECTOR MACHINE (SVM) PADA KLASIFIKASI JENIS CENGKEH BERDASARKAN FITUR TEKSTUR DAUN.
- Wijaya, H. D., & Dwiasnati, S. (2020). Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat. *Jurnal Informatika*, 7(1), 1–7.