

Klasifikasi Kualitas Buah Pisang Berdasarkan Waktu Panen dan Tingkat Kematangan Menggunakan Metode SVM & KNN

Doni Andriansyah¹, Elly Mufida², Nandang Iriadi³, Rian Septian Anwar⁴, Luci Kanti Rahayu⁵

¹Universitas Nusa Mandiri, doni.dad@nusamandiri.ac.id, Jl. Jatiwaringin Raya, Jakarta, Indonesia

²Universitas Bina Sarana Informatika, elly.elm@bsi.ac.id, Jl. Kramat 98, Jakarta, Indonesia

³Universitas Bina Sarana Informatika, nandang.ndi@bsi.ac.id, Jl. Kramat 98, Jakarta, Indonesia

⁴Universitas Bina Sarana Informatika, rian.ptn@bsi.ac.id, Jl. Kramat 98, Jakarta, Indonesia

⁵Universitas Bina Sarana Informatika, luci.lkr@bsi.ac.id, Jl. Kramat 98, Jakarta, Indonesia

Informasi Makalah

Submit : Apr 20, 2024
Revisi : May 15, 2024
Diterima : May 30, 2024

Kata Kunci :

Klasifikasi;
KNN;
SVM;
Exploratory Data Analysis;
Kualitas Buah Pisang.

Abstrak

Tanaman pisang atau Musa Paradisiaca merupakan tanaman yang masuk kedalam golongan klimakterik, sehingga memerlukan perhatian khusus pasca panen. Tingkat kematangan buah pisang saat dipanen sangat mempengaruhi daya simpan dan kualitas buah. Waktu panen sangat penting untuk mendapatkan buah yang matang dan berkualitas. Penelitian menggunakan algoritma SVM dan KNN dengan tujuan untuk mengetahui algoritma terbaik dalam klasifikasi kualitas buah pisang. Kumpulan data yang digunakan merupakan data publik mengenai kualitas buah pisang dengan jumlah data sebanyak 8.000 baris data, dan dengan delapan atribut kolom. Dalam pengolahan data hanya menggunakan atribut kolom waktu panen dan tingkat kematangan serta dilakukan proses pengacakan terhadap kumpulan data agar model dapat belajar lebih baik dan mencegah data dari bias. Hasil penelitian menunjukkan klasifikasi dengan SVM memiliki nilai akurasi sebesar 73,4%, lebih baik dari hasil klasifikasi KNN yang hanya mencapai 69,6%.

Abstract

Banana or Musa Paradisiaca is a plant that falls into the climacteric group, so it requires special attention after harvest. The level of ripeness of bananas when harvested greatly influences the shelf life and quality of the fruit. Harvest time is very important to get ripe and quality fruit. The research uses the SVM and KNN algorithms with the aim of finding out the best algorithm for classifying banana fruit quality. The data set used is public data regarding the quality of bananas with a total of 8,000 rows of data, and with eight column attributes. In data processing, we only use the column attributes of harvest time and ripeness level and a randomization process is carried out on the data set so that the model can learn better and prevent the data from being biased. The research results show that classification with SVM has an accuracy value of 73.4%, better than the results of KNN classification which only reached 69.6%.

1. Pendahuluan

Tanaman pisang atau yang dikenal dengan nama latin *Musa Paradisiaca* merupakan salah satu tanaman herba yang terdiri dari akar, batang, daun, dan buah (Ekayanti et al., 2023). Tanaman ini termasuk kedalam suku Musaceae, dapat tumbuh baik di dataran rendah maupun di dataran tinggi sehingga tanaman ini banyak ditanam oleh masyarakat sebagai bahan pangan. Pisang memiliki tekstur yang lunak dan mudah dicerna oleh tubuh (Wulandari et al., 2018). Umumnya pisang dianggap sebagai sumber vitamin dan mineral, dalam daging buah pisang terdapat 70% air, 27% karbohidrat, 0,5% serat, 1,2% protein, dan 0,31% lemak.

Di Indonesia, pusat produksi pisang terdapat di beberapa wilayah yaitu di Jawa Barat (Sukabumi, Cianjur, Bogor), Jawa Tengah (Demak, Pati, Banyumas), Jawa Timur (Banyuwangi dan Malang), Sumatera Utara (Padangsidempuan, Tarutung), Sumatera Selatan (Ogan Komeling Ilir), Lampung (Kayu Agung, Metro), Kalimantan, Sulawesi, Maluku, Bali, dan Nusa Tenggara Barat (Suryalita, 2019). Indonesia juga merupakan salah satu pemasok pisang segar ke berbagai negara seperti Singapura, Cina, Arab, Australia, Amerika Serikat, dan beberapa negara lainnya. Sebagai salah satu komoditi ekspor, kualitas pisang perlu diperhatikan. Salah satu cara untuk mengetahui kualitas buah pisang dapat dilihat dari tingkat kematangan buah tersebut (Wulansari & Muslih, 2023).

Di Indonesia, waktu panen buah pisang umumnya ditentukan oleh kebutuhan ekonomi dan keamanan, bukan berdasarkan tingkat kematangan atau usia petiknya (Radiena, 2016). Tak jarang sering kali ditemukan buah pisang yang belum matang dipasarkan.

Pisang termasuk kedalam golongan buah klimakterik dan bersifat mudah rusak setelah dipanen (S. E. Widodo et al., 2023). Tingkat

kematangan buah pisang saat dipanen sangat mempengaruhi daya simpan dan kualitas buah. Waktu panen sangat penting untuk mendapatkan buah yang matang dan berkualitas (Irhamni et al., 2023). Buah yang dipanen terlalu muda memiliki daya simpan yang rendah dan kualitas yang kurang baik ketika matang (W. D. Widodo et al., 2019). Umumnya usia panen terbaik dalam penanganan pasca panen untuk memperpanjang umur pisang antara 95 hari sampai dengan 110 hari setelah antesis (HSA) disesuaikan dengan varietasnya masing-masing, dan usia penyimpanan berkisar antara 10-11 hari (Safitri et al., 2023).

Dalam penelitian yang dilakukan (Shudiq et al., 2020) menyatakan bahwa masyarakat sering kali merasa tidak puas terhadap pohon pisang yang mereka tanam dengan alasan tidak seperti yang diharapkan. Hal itu dapat dimaklumi karena pohon pisang memerlukan waktu cukup lama untuk dapat tumbuh dan akhirnya memiliki buah. Metode yang digunakan dalam penelitian adalah klasifikasi SVM dan KNN, dengan penelitian ini akan diketahui pendekatan mana yang memiliki hasil klasifikasi terbaik. Hasil penelitian menunjukkan metode KNN dengan nilai akurasi 74,00% lebih baik dari metode SVM yang mencapai 67,89%.

Penelitian oleh (Muhammad et al., 2021) menjelaskan bahwa didalam menentukan kematangan buah, warna biasanya menjadi indikator dalam menentukan tingkat kematangan buah yang dapat dilihat dengan mata manusia. Namun mata manusia tidak selalu tepat dalam menentukan kematangan buah. Penelitian menggunakan metode SVM dengan delapan puluh gambar berbagai warna. Untuk memudahkan klasifikasi, gambar tersebut diubah menjadi grayscale (hitam putih). Hasil penelitian menunjukkan tingkat kematangan yang diperoleh mencapai 75% dari 80 citra pisang.

(Adenugraha et al., 2022) dalam penelitiannya memaparkan bahwa kualitas

buah pisang dipengaruhi oleh waktu panen, tingkat kematangannya pun dapat mempengaruhi waktu penyimpanan sehingga berdampak pada jangkauan pemasaran. Petani biasanya menggunakan metode manual untuk mengukur kematangan pisang ambon, sehingga terdapat faktor yang dapat membuat hasilnya kurang akurat. Penelitian menggunakan klasifikasi KNN dengan memanfaatkan fitur warna RGB dan HSV dengan jumlah kumpulan data sebanyak 41 yang terbagi menjadi 30 data pelatihan dan 11 data pengujian. Dari penelitian didapatkan akurasi sebesar 90,9% dimana 10 data pengujian menunjukkan hasil yang akurat dan 1 data menunjukkan hasil yang salah.

Penelitian yang dilakukan oleh (Wicaksono, 2022) memaparkan permasalahan yang sering ditemukan yaitu kurang tepat dan kurangnya pengetahuan didalam membedakan jenis, kualitas dan kematangan buah pisang. Tujuan dari penelitian yang dilakukan untuk mengklasifikasikan tingkat kematangan, kualitas dan jenis buah pisang berdasarkan ciri warna, ukuran dan bentuk. Dataset yang digunakan merupakan gambar pisang dengan 9 jenis yaitu pisang ambon, pisang raja, pisang cavendish, pisang kirana, pisang barangan, pisang nangka, pisang mas, pisang susu, dan pisang kepok. Adapun tingkat kematangan yang diukur yaitu mentah, matang, dan terlalu matang. Aplikasi dibuat menggunakan TensorFlow Python dengan menghasilkan akurasi sebesar 98,7%.

(Utami et al., 2021) dalam penelitiannya menjelaskan sering ditemukannya kualitas buah yang tidak baik ketika sudah berada ditangan konsumen, ini menyebabkan membuat citra produsen menurun. Penelitian menggunakan metode Convolutional Neural Network (CNN) yang mampu melakukan pembelajaran mandiri untuk mengenal objek, mengekstraksi dan mengklasifikasi dengan empat kelompok gambar dan menggunakan jarak. Kelompok pertama adalah objek pisang dengan kondisi mentah. Kelompok

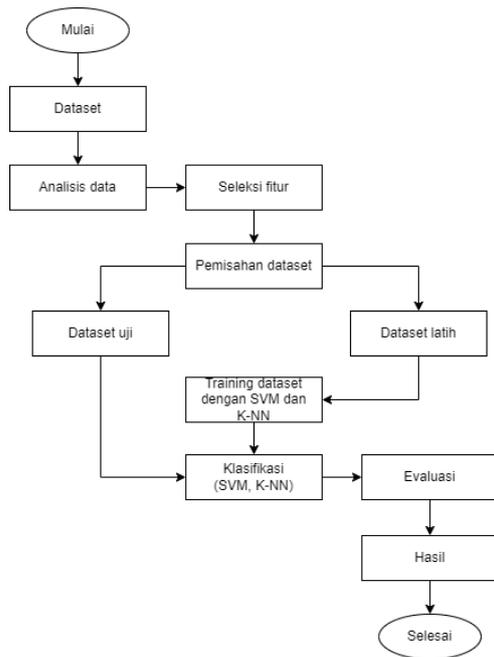
kedua adalah pisang dengan kondisi matang. Kelompok ketiga pisang dengan kondisi busuk, dan kelompok terakhir gambar selain buah pisang. Hasil deteksi oercobaan kelompok pertama mencapai 100%, hasil deteksi kelompok kedua mencapai 78%, kelompok ketiga mendapatkan hasil deteksi sebesar 89%, dan pada kelompok keempat sistem tidak mendeteksi adanya pola buah pisang. Hasil penelitian menunjukkan bahwa metode CNN berpotensi mengenal objek secara otomatis dalam membedakan jenis pola buah pisang.

Penelitian ini menggunakan algoritma SVM dan KNN dengan maksud mengetahui algoritma terbaik dalam mengklasifikasikan kualitas pisang berdasarkan tingkat kematangan dan waktu panen. SVM merupakan salah satu metode klasifikasi dengan tingkat akurasi tinggi (Muhammad et al., 2021). Metode klasifikasi KNN menggunakan data dari objek paling dekat. Ada atau tidaknya fitur yang relevan mempengaruhi ketepatan algoritma KNN (Shudiq et al., 2020).

2. Metode Penelitian

2.1 Alur Penelitian

Penelitian ini menggunakan bahasa pemrograman Python yang berisi berbagai pustaka untuk membantu pemrosesan pembelajaran mesin. Berikut adalah alur penelitian yang digunakan:



Gambar 1. Alur Penelitian.

Proses penelitian diawali dengan mengunduh kumpulan data dari website penyedia materi publik, kemudian dilanjutkan dengan analisis materi untuk lebih memahami data. Selanjutnya, mengimplementasikan fitur berdasarkan literatur yang diperoleh kedalam proses pembelajaran mesin. Kumpulan data kemudian dibagi menjadi dua kelompok, yaitu data pelatihan yang digunakan dalam proses pembuatan model dan data pengujian yang digunakan dalam evaluasi model. Langkah terakhir adalah tahap evaluasi model untuk mencapai tingkat akurasi klasifikasi.

2.2 Kumpulan Data

Kumpulan data yang digunakan untuk penelitian adalah data publik berisi informasi berupa angka mengenai kualitas buah pisang yang dapat di akses dengan bebas melalui Kaggle. Data terdiri dari 8.000 baris data dengan delapan atribut yaitu ukuran (*size*), berat (*weight*), tingkat kemanisan (*sweetness*), tingkat kelembutan (*softness*), waktu panen (*harvest time*), tingkat kematangan (*ripeness*), keasaman (*acidity*), dan kualitas (*quality*).

Terdapat satu kolom yang teridentifikasi sebagai label dengan tipe data String, yaitu

kolom Quality dengan nilai “Good” dan “Bad”.

2.3 EDA (Exploratory Data Analysis)

Analisis data merupakan cara untuk mengetahui bagaimana data digambarkan, hubungannya, semantiknya, dan keterbatasannya dalam sistem informasi (Husni et al., 2022). EDA merupakan metode untuk menganalisis dan memvisualisasikan data dengan tujuan meningkatkan pemahaman tentang informasi yang ada didalamnya (Patricia et al., 2023). Tugas utama EDA adalah membersihkan data, mendeskripsikan data, melihat sebaran data, membandingkan hubungan antar data dan menyimpulkan data (Sinaga et al., 2023).

Python memiliki beberapa pustaka yang sangat berguna dan dapat membantu proses analisis data (Hermanto et al., 2023), diantaranya:

2.3.1. Numpy

Numpy merupakan pustaka yang digunakan untuk komputasi numerik. Numpy menyediakan array multidimensi yang efektif dan efisien untuk melakukan operasi matematika.

2.3.2. Pandas

Pandas merupakan *packages* didalam Python yang menyediakan struktur data dengan cepat, fleksibel, dan ekspresif yang dirancang untuk bekerja dengan relasional dan label data. Dengan Pandas, data akan tersaji kedalam *DataFrame* sehingga pengguna akan lebih mudah memahami dan menganalisis data.

2.3.3. Matplotlib

Salah satu pustaka yang dapat digunakan untuk memvisualisasikan data di Python. Matplotlib menyediakan fungsi dan alat untuk membuat grafik, plot, dan visualisasi lainnya.

2.3.4. Seaborn

Seaborn merupakan pustaka yang dibangun diatas Matplotlib, digunakan secara khusus untuk visualisasi data statistik. Seaborn menyediakan fungsi dan tampilan yang lebih kaya untuk menghasilkan visualisasi yang menarik dan informatif.

2.4 SVM (Support Vector Machine)

SVM merupakan salah satu cara dalam pembelajaran mesin yang bekerja atas prinsip SRM (*Structural Risk Minimization*) dan bertujuan menemukan *hyperlane* terbaik yang memisahkan dua kelompok data (Shudiq et al., 2020). *Hyperlane* merupakan batasan dan membantu mengklasifikasikan titik data (Andriansyah & Fridayanthie, 2023). SVM sudah banyak digunakan dalam memecahkan masalah klasifikasi dan salah satu metode *supervised learning* populer. SVM memiliki beberapa kernel yang dapat digunakan dalam proses klasifikasi, yaitu Sigmoid, Polynomial, RBF, dan Linear. Perhitungan algoritma SVM menggunakan kernel RBF dapat dilihat pada persamaan (1).

$$f(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|^2}{2\sigma^2}} \quad (1)$$

Keterangan:

σ : *variance hyperparameter*

$\|x_1 - x_2\|$: *Euclidean distance* antara dua titik x_1 dan x_2

2.5 KNN (K-Nearest Neighbors)

KNN adalah salah satu cara untuk menempatkan objek ke dalam kategori atas dasar data pembelajaran yang terdekat atau memiliki persamaan karakteristik terbanyak dengannya (Adenugraha et al., 2022). Persamaan jarak geometris dapat digunakan untuk menghitung nilai jarak metode KNN. Metode ini sederhana dan dapat memberikan hasil klasifikasi yang sangat akurat.

$$d(xy) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Keterangan:

d: jarak kedekatan, x: data latih, y: data uji,

n: jumlah atribut antara 1 s.d n

i: atribut individu antara 1 s.d n.

2.6 Evaluasi

Confusion Matrix merupakan salah satu metode yang dapat digunakan dalam mengevaluasi kinerja model dengan menghasilkan nilai *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN).

Dengan matrik ini akan dihitung nilai presisi, akurasi dan *recall* atau sensitivitasnya. Akurasi adalah sejauh mana nilai prediksi mendekati nilai sebenarnya.

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) <i>Type I Error</i>
	0 (Negative)	FN (False Negative) <i>Type II Error</i>	TN (True Negative)

Gambar 2. Confusion Matrix.

Persamaan untuk menghitung nilai akurasi (3).

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Keterangan:

TP = *True Positive*

FP = *False Positive*

TN = *True Negative*

FN = *False Negative*

Presisi menggambarkan indeks akurasi hasil perkiraan yang diberikan oleh data dan model yang diminta adalah rasio perkiraan positif aktual terhadap jumlah total hasil prediksi positif. Persamaan untuk menghitung nilai presisi (4).

$$precision = \frac{TP}{TP+FP} \quad (4)$$

Keterangan:

TP = *True Positive*

FP = *False Positive*

Sensitivitas atau *recall* menggambarkan tingkat keberhasilan model dalam mengambil data, sensitivitas dapat diartikan juga sebagai rasio prediksi positif yang sebenarnya

terhadap semua positif yang sebenarnya. Berikut persamaan untuk menghitung sensitivitas (5).

$$recall = \frac{TP}{TP+FN} \quad (5)$$

Keterangan:

TP = True Positive

FN = False Negative

Nilai F1 mewakili rata-rata harmonik dari sensitivitas dan presisi dengan nilai terbaik 1 dan terburuk 0. Proporsi relatif presisi dan sensitivitas pada nilai F1 sama. Berikut persamaan untuk menghitung nilai F1 (6).

$$F1\ score = 2 * \frac{precision*recall}{precision+recall} \quad (6)$$

Keterangan:

Precision = nilai precision

Recall = nilai recall

3. Hasil dan Pembahasan

3.1. Import Library

Untuk dapat menyajikan data maka perlu memasukkan *library*/pustaka yang dibutuhkan agar data dapat dengan mudah di analisis dan di olah. Terdapat beberapa pustaka yang sering digunakan dalam menganalisis data menggunakan Python diantaranya Numpy, Pandas, Matplotlib, dan Seaborn.

Pustaka lainnya yang perlu ditambahkan adalah pustaka pembelajaran mesin untuk SVM dan KNN, pustaka ini akan digunakan untuk menghasilkan akurasi dari model.

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.preprocessing import LabelEncoder
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix, classification_report
from sklearn.metrics import f1_score
from sklearn.metrics import precision_score
from sklearn.metrics import recall_score
```

Gambar 3. Import Library.

3.2. Loading Dataset

Untuk lebih memudahkan didalam mengolah dan menganalisis data, kumpulan data kemudian di unggah ke Google Drive dalam format (*.csv), Python kemudian mengaksesnya menggunakan pustaka *DataFrame*.

	Size	Weight	Sweetness	Softness	HarvestTime	Ripeness	Acidity	Quality
0	-1.924968	0.468078	3.077832	-1.472177	0.294799	2.435570	0.271290	Good
1	-2.409751	0.486870	0.346921	-2.495099	-0.892213	2.067549	0.307325	Good
2	-0.357607	1.483176	1.568452	-2.645145	-0.647267	3.090643	1.427322	Good
3	-0.868524	1.566201	1.889605	-1.273761	-1.006278	1.873001	0.477862	Good
4	0.651825	1.319199	-0.022459	-1.209709	-1.430692	1.076345	2.812442	Good
...
7995	-6.414403	0.723565	1.134953	2.952763	0.297928	-0.156946	2.398091	Bad
7996	0.851143	-2.217875	-2.812175	0.489249	-1.323410	-2.316883	2.113136	Bad
7997	1.422722	-1.907665	-2.532364	0.964976	-0.562375	-1.834765	0.697361	Bad
7998	-2.131904	-2.742600	-1.008029	2.126946	-0.802632	-3.580266	0.423569	Bad
7999	-2.660879	-2.044666	0.159026	1.499706	-1.581856	-1.605859	1.435644	Bad

Gambar 4. Dataset Kualitas Buah Pisang.

Dengan *DataFrame*, data akan menjadi lebih mudah untuk dipahami dan dianalisis lebih lanjut. Dari gambar diatas, dapat diketahui terdapat 8.000 baris data dan 8 kolom dimana terdapat satu kolom berisi data string, yaitu kolom Quality.

Kolom Quality dapat diidentifikasi sebagai label didalam kumpulan data tersebut yang memiliki nilai “Good” dan “Bad”.

3.3. Analisis Data

Tahap berikutnya adalah melihat kondisi data untuk mengetahui apakah data dapat dilanjutkan ke proses berikutnya atau masih terdapat nilai pada data yang perlu disesuaikan.

Hal yang pertama kali dilakukan adalah mencari tahu apakah terdapat nilai yang kosong didalam kumpulan data. Apabila terdapat nilai yang kosong, maka dapat menimbulkan anomali, sehingga data perlu diperbaiki.

```
Size      0
Weight    0
Sweetness 0
Softness  0
HarvestTime 0
Ripeness  0
Acidity   0
Quality   0
dtype: int64
```

Gambar 5. Analisis Data Null.

Berdasarkan informasi gambar 5 diatas dapat diketahui tidak ada data yang kosong (*null value*) didalam kumpulan data.

	count	mean	std	min	25%	50%	75%	max
Size	8000.000000	-0.747802	2.136023	-7.998074	-2.277851	-0.897514	0.854216	7.970800
Weight	8000.000000	-0.761019	2.015934	-8.283002	-2.223574	-0.888659	0.775491	5.679692
Sweetness	8000.000000	-0.770224	1.948455	-6.434022	-2.107329	-1.020673	0.311048	7.539374
Softness	8000.000000	-0.014441	2.065216	-6.959320	-1.590458	0.202644	1.547120	8.241555
HarvestTime	8000.000000	-0.751288	1.996661	-7.570008	-2.120659	-0.934192	0.507326	6.293280
Ripeness	8000.000000	0.781098	2.114289	-7.423155	-0.574226	0.964952	2.261650	7.249034
Acidity	8000.000000	0.008725	2.293467	-8.226977	-1.629450	0.098735	1.682063	7.411633

Gambar 6. Kondisi Data Kualitas Buah Pisang.

Melihat kondisi data yang ditunjukkan gambar 6 diatas, dapat diketahui nilai mean, standar deviasi, min dan max. Terlihat data cukup bagus dengan memiliki skala yang sama sehingga data dapat dilanjutkan ke proses berikutnya tanpa perlu dilakukan *scalling technique*. Sebelum melanjutkan ke tahap selanjutnya, data pada kolom Quality perlu diubah kedalam format numerik agar data dapat dengan mudah diproses.

	Size	Weight	Sweetness	Softness	HarvestTime	Ripeness	Acidity	Quality
0	-1.924968	0.468078	3.077832	-1.472177	0.294799	2.435570	0.271290	1
1	-2.409751	0.486870	0.346921	-2.495099	-0.892213	2.067549	0.307325	1
2	-0.357607	1.483176	1.568452	-2.645145	-0.647267	3.090643	1.427322	1
3	-0.868524	1.566201	1.889605	-1.273761	-1.006278	1.873001	0.477862	1
4	0.651825	1.319199	-0.022459	-1.209709	-1.430692	1.078345	2.812442	1
...
7995	-6.414403	0.723565	1.134953	2.952763	0.297928	-0.156946	2.398091	0
7996	0.851143	-2.217875	-2.812175	0.489249	-1.323410	-2.316883	2.113136	0
7997	1.422722	-1.907665	-2.532364	0.964976	-0.562375	-1.834765	0.697361	0
7998	-2.131904	-2.742600	-1.008029	2.126946	-0.802632	-3.580266	0.423569	0
7999	-2.660879	-2.044666	0.159026	1.499708	-1.581856	-1.605859	1.435644	0

Gambar 7. Ubah Format Data Kolom Quality.

Sebelum dilakukan pemilihan fitur, dilakukan pengacakan (*shuffle*) terhadap kumpulan data dengan tujuan agar model dapat belajar lebih baik dan mencegah data dari bias.

3.4. Seleksi Fitur

Pada tahap ini dilakukan pemilihan fitur atau atribut kolom yang akan digunakan dalam pengolahan data. Kolom yang akan digunakan yaitu kolom *HarvestTime* (waktu panen) dan *Ripeness* (tingkat kematangan).

Pemilihan fitur tidak melalui proses pembelajaran mesin, melainkan berdasarkan data dan referensi dengan langsung memilih kolom yang akan dijadikan landasan pengukuran dalam penelitian.

	HarvestTime	Ripeness
1870	-1.442473	3.523150
2103	-2.323390	0.829444
4817	3.178717	3.212256
2380	-0.799835	1.966106
588	-1.945009	1.748981
...
3124	-2.284062	0.245692
2107	0.273149	1.295956
6047	-2.013535	0.803934
978	-1.861185	0.361241
7068	-4.650889	0.021511

Gambar 8. Pemilihan Fitur *HarvestTime* dan *Ripeness*.

3.5. Pemisahan Kumpulan Data

Langkah selanjutnya adalah memisahkan kumpulan data menjadi dua kelompok yaitu data pelatihan sebesar 80% dengan 6.400 data, dan data pengujian sebesar 20% dengan 1.600 data dari total keseluruhan sebanyak 8.000 data.

3.6. Training dan Evaluasi Model

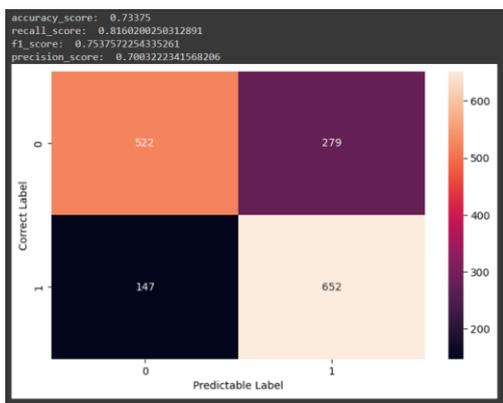
Dari hasil pemisahan kumpulan data, dibuatkanlah model klasifikasi menggunakan algoritma SVM dengan pemilihan kernel *Radial Basis Function* (RBF) dan KNN, model akan menggunakan data uji dan fitur yang telah dipilih untuk memperoleh nilai akurasi yang terbaik.

```
Model klasifikasi SVM dengan kernel RBF
svcmodel = SVC(kernel='rbf', C=10)
svcmodel.fit(X_train, y_train)
y_pred_svm = svcmodel.predict(X_test)
```

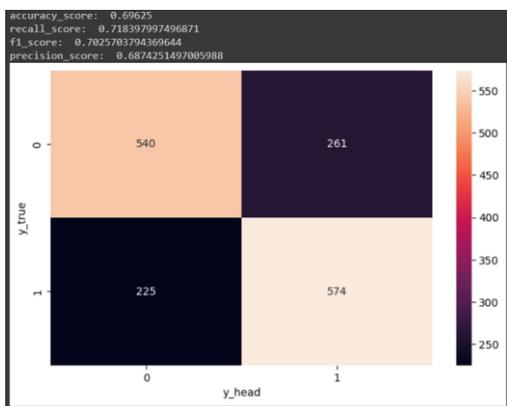
Model klasifikasi KNN

```
knn_classifier =
KNeighborsClassifier(n_neighbors = 5, metric
= 'minkowski', p = 2)
knn_classifier.fit(X_train, y_train)
y_pred_knn = knn_classifier.predict(X_test)
```

Selanjutnya model di evaluasi untuk memperoleh hasil kinerja model. Gambar 9 dan gambar 10 adalah hasil evaluasi model yang diperoleh dari SVM dan KNN secara berurutan.



Gambar 9. Hasil Evaluasi SVM.



Gambar 10. Hasil Evaluasi KNN.

Ringkasan klasifikasi dapat dilihat pada tabel 1.

Tabel 1. Hasil klasifikasi.

Algoritma	Akurasi	Presisi	Re call	F1 Score
SVM	73,4 %	70,0%	81,6%	75,4%
KNN	69,6 %	68,7%	71,8%	70,2%

4. Simpulan

Dari hasil penelitian dengan metode pengacakan (*shuffle*) terhadap kumpulan data dan pemilihan fitur diketahui bahwa klasifikasi kualitas buah pisang berdasarkan waktu panen (*harvest time*) dan tingkat kematangan (*ripeness*) menggunakan algoritma SVM dengan kernel RBF memiliki nilai akurasi sebesar 73,4%, dan algoritma KNN dengan nilai akurasi sebesar 69,6%. SVM menunjukkan hasil yang lebih baik dan mampu mengklasifikasi kualitas buah pisang berdasarkan waktu panen dan tingkat kematangan. Untuk penelitian berikutnya pada algoritma SVM dapat menggunakan kernel lain seperti kernel Linear atau Polynomial untuk mengetahui kernel mana yang memiliki hasil klasifikasi terbaik.

5. Referensi

Adenugraha, S. P., Arinal, V., & Mulyana, D. I. (2022). Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(1), 9. <https://doi.org/10.30865/mib.v6i1.3287>

Andriansyah, D., & Fridayanthie, E. W. (2023). Optimization of Support Vector Machine and XGBoost Methods Using Feature Selection to Improve Classification Performance. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, 6(2), 484–493. <https://doi.org/10.31289/jite.v6i2.8373>

Ekayanti, N. L. F., Megawati, F., & Dewi, N. L. K. A. A. (2023). PEMANFAATAN TANAMAN PISANG (Musa Paradisiaca L.) SEBAGAI SEDIAAN KOSMETIK: ARTIKEL REVIEW: PEMANFAATAN TANAMAN PISANG (Musa Paradisiaca L.) SEBAGAI SEDIAAN KOSMETIK. *Usadha*, 2(2), 19–24. <https://doi.org/10.36733/usadha.v2i2.6217>

Hermanto, K., Salim, D., Wu, B., Salim, O. R., & Gunadi, R. B. (2023). Penggunaan Python Untuk Menganalisis Pola Penyebaran Covid-19 Di Masa Pandemi.

- Journal of Student Development Information System (JoSDIS)*, 3(2), 62–75.
- Husni, D. T., Sitompul, D. R. H., Sinurat, S. H., Ruben, R., Situmorang, A., Ziegel, D. J., Rahmad, J., & Indra, E. (2022). ANALISIS BIG DATA PENJUALAN VIDEO GAMES MENGGUNAKAN EDA. *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, 5(1), 43. <https://doi.org/10.37600/tekinkom.v5i1.517>
- Irhanni, D., Hayati, R., & Hasanuddin, H. (2023). Pengaruh Tingkat Kematangan dan Lama Penyimpanan terhadap Kualitas Pisang Mas (*Musa acuminata* Colla). *JURNAL AGROTROPIKA*, 22(2), 145. <https://doi.org/10.23960/ja.v22i2.7883>
- Muhammad, A. A., Arkadia, A., NaufalRifqi, S., Prasvita, D. S., & Trianto, T. (2021). *Klasifikasi Kematangan Buah Pisang Berdasarkan Fitur Warna dengan Metode SVM*. Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA), Jakarta-Indonesia.
- Patricia, S., Marpaung, 2Christopher Parulian, Wijaya, L. R., Paramartha, M. A., Atmadja, W. D., & Ningsih, R. Y. (2023). Implementasi Exploratory Data Analysis (EDA) Untuk Menganalisis Berbagai Faktor Risiko Penyakit Jantung Di Amerika Serikat. *Journal of Student Development Information System (JoSDIS)*, 3(2).
- Radiena, M. S. Y. (2016). UMUR OPTIMUM PANEN PISANG KEPOK (*Musa paradisiaca*, L) TERHADAP MUTU TEPUNG PISANG. *Majalah Biam*, 12(2), 27–33.
- Safitri, D. S., Arti, I. M., Miska, M. E. E., & Kalsum, U. (2023). KARAKTERISTIK BUAH PISANG MAS KIRANA PADA BERBAGAI UMUR PANEN DAN TEKNIK PENYIMPANAN. *JURNAL TEKNOLOGI PANGAN*, 17(2), 70–82.
- Shudiq, W. J., As, A. H., & Rahman, M. F. (2020). Penentuan Metode Terbaik Dalam Menentukan Jenis Pohon Pisang Menurut Tekstur Daun (Metode K-NN dan SVM). *Jurnal Teknologi dan Manajemen Informatika*, 6(2), 128–136. <https://doi.org/10.26905/jtmi.v6i2.5156>
- Sinaga, P. T., Purba, S. S., Wiranto, D., Maharja, O. J., & Indra, E. (2023). EXPLORATORY DATA ANALYSIS OF CLINICAL HEART FAILURE USING A SUPPORT VECTOR MACHINE. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, 7(1), 142–154. <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v7i1.4100>
- Suryalita, S. (2019, July 20). *Review Beraneka Ragam Jenis Pisang dan Manfaatnya*. Prosiding Seminar Nasional Biodiversitas Indonesia, Gowa.
- Utami, M., Andika, J., & Attamimi, S. (2021). Artificial Intelligence For Banana's Ripeness Detection Using Conventional Neural Network Algorithm. *Jurnal Teknologi Elektro*, 12(2), 73. <https://doi.org/10.22441/jte.2021.v12i2.005>
- Wicaksono, A. D. P. (2022). Klasifikasi Tingkat Kematangan, Kualitas dan Jenis Buah Pisang Berdasarkan Ciri Warna dan Bentuk Menggunakan Artificial Neural Networks. *JTII*, 07(02).
- Widodo, S. E., Waluyo, S., Karyanto, A., Zulferiyeni, Z., Febrianingrum, N., Latansya, R., & Putri, M. D. (2023). APLIKASI THERMAL IMAGE PENDETEKSI TINGKAT KEMATANGAN BUAH PISANG DAN APOKAT. *Jurnal Agrotek Tropika*, 11(2), 165. <https://doi.org/10.23960/jat.v11i2.6168>
- Widodo, W. D., Suketi, K., & Rahardjo, R. (2019). Evaluasi Kematangan Pascapanen Pisang Barangan untuk Menentukan Waktu Panen Terbaik Berdasarkan Akumulasi Satuan Panas. *Buletin Agrohorti*, 7(2), 162–171. <https://doi.org/10.29244/agrob.7.2.162-171>
- Wulandari, R. T., Widyastuti, N., & Ardiaria, M. (2018). PERBEDAAN PEMBERIAN PISANG RAJA DAN PISANG AMBON TERHADAP VO₂max PADA REMAJA DI SEKOLAH SEPAK BOLA. *Journal of Nutrition College*, 7(1), 8. <https://doi.org/10.14710/jnc.v7i1.20773>
- Wulansari, N. W., & Muslih, M. (2023). Implementasi Metode Convolutional Neural Network Untuk Deteksi Kematangan Buah Pisang Menggunakan Inception V3. *JOINS (Journal of*

Information System), 8(2), 147–155.
<https://doi.org/10.33633/joins.v8i2.9074>