

# ALGORITMA NAÏVE BAYES, *RANDOM FOREST* DAN *SUPPORT VECTOR MACHINE* UNTUK KLASIFIKASI PENYAKIT SAPI

Cici Emilia Sukmawati  
Universitas Buana Perjuangan Karawang  
Kab. Karawang, Indonesia  
cici.emilia@ubpkarawang.ac.id

Adi Rizky Pratama  
Universitas Buana Perjuangan Karawang  
Kab. Karawang, Indonesia  
adi.rizky@ubpkarawang.ac.id

Nunung Nurjanah  
Universitas Buana Perjuangan Karawang  
Kab. Karawang, Indonesia  
if20.nunungnurjanah@mhs.ubpkarawang.ac.id

**Abstract**— Data World Health Organization (WHO) menunjukkan bahwa Indonesia memiliki jumlah penduduk terbanyak di dunia, menempatkannya sebagai negara terbesar keempat dalam hal populasi. Dengan jumlah penduduk yang besar, kebutuhan akan protein hewani menjadi meningkat, seiring dengan kesadaran akan pentingnya gizi yang baik. Salah satu sumber protein hewani yang penting adalah daging sapi. Karena itu, menjaga kesehatan sapi menjadi krusial untuk memenuhi kebutuhan gizi masyarakat. Dalam upaya menjaga kesehatan hewan, khususnya sapi, pengamatan terhadap gejala-gejala penyakit menjadi sangat penting. Teknik data mining dapat menjadi solusi yang efektif dalam mengklasifikasikan penyakit pada sapi. Penelitian ini menggunakan metode data mining dengan tiga percobaan yang berbeda. Percobaan pertama menggunakan algoritma *Naive bayes* dengan akurasi, presisi, dan recall sebesar 0.8. Hasil ini menunjukkan bahwa model *Naive bayes* mampu memprediksi dengan benar sekitar 80% dari data yang diproses, dengan tingkat kesalahan yang rendah. Percobaan kedua menggunakan algoritma *Random Forest*, yang menghasilkan hasil yang serupa dengan *Naive bayes*, yaitu akurasi, presisi, dan recall sebesar 0.8. *Random Forest* juga memiliki kinerja yang baik dalam memprediksi penyakit pada sapi. Namun, percobaan ketiga menggunakan algoritma *Support Vector Machine* (SVM) menghasilkan presisi yang lebih rendah, yaitu 0.6. Meskipun memiliki akurasi dan recall yang baik (0.8), presisi yang rendah dapat mengindikasikan adanya false positive yang lebih tinggi dalam prediksi penyakit. Dengan demikian, dapat disimpulkan bahwa *Naive bayes* dan *Random Forest* memiliki kinerja yang lebih baik dalam memprediksi penyakit pada sapi dibandingkan dengan SVM dalam konteks penelitian ini. Hasil ini dapat menjadi dasar untuk pengembangan model prediksi penyakit yang lebih baik untuk membantu manajemen kesehatan hewan dan pemenuhan kebutuhan gizi masyarakat Indonesia.

**Kata kunci** — Penyakit sapi, *Naive bayes*, *R andom forest*, *Support Vector Machine*

## I. PENDAHULUAN

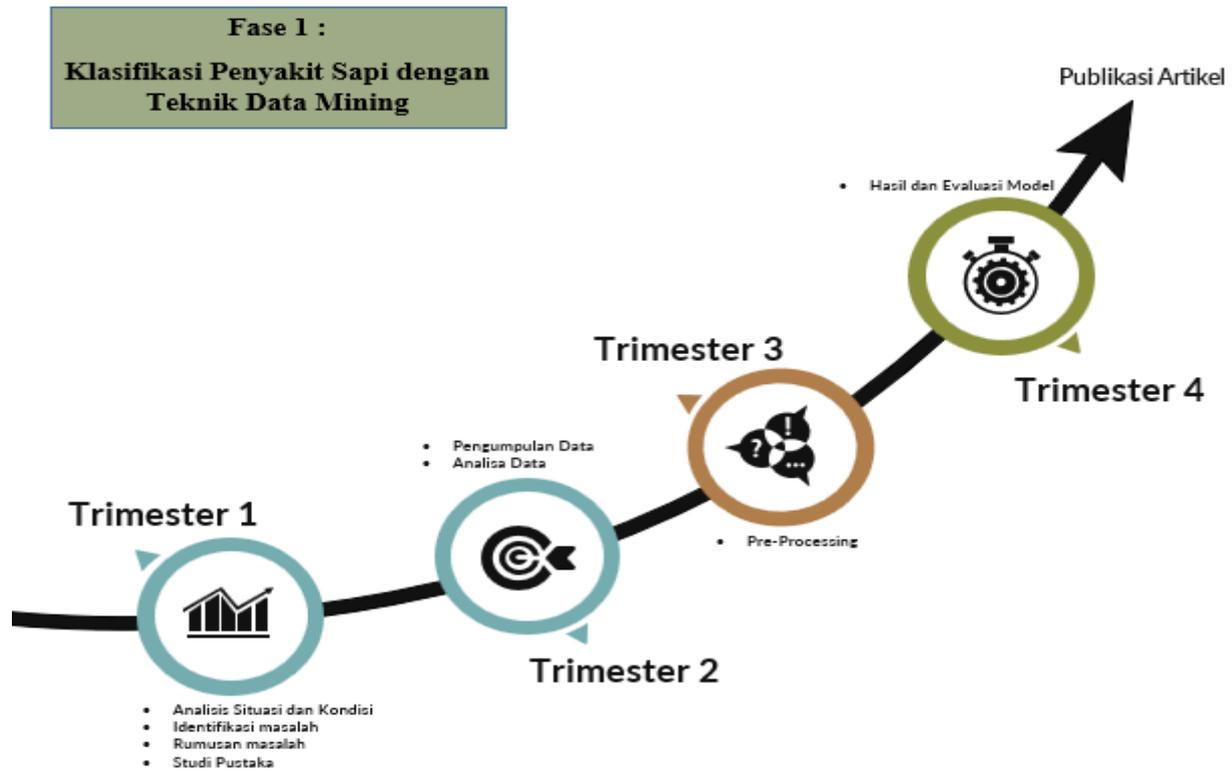
Berdasarkan dari data *World Health Organization* (WHO) menyebutkan bahwa Indonesia menjadi salah satu negara yang memiliki penduduk terbanyak di dunia dengan urutan keempat [1][2]. Karena jumlah penduduk yang banyak, sehingga kebutuhan protein hewani meningkat dengan kesadaran pentingnya akan gizi. Salah satu sumber protein hewani yang bisa diperoleh dari daging sapi [3]–[5]. Hal itu menyebabkan kesehatan sapi menjadi suatu hal yang berpengaruh untuk memenuhi kebutuhan gizi [6]. Dalam menjaga kesehatan sapi, perlu adanya suatu pengamatan terhadap gejala-gejala penyakit pada sapi [7]. Namun, mendiagnosis penyakit pada sapi harus melalui proses yang rumit, karena gejala-gejala yang ditunjukkan oleh sapi seringkali mirip dengan penyakit lain atau bahkan tidak menunjukkan gejala sama sekali [8]. Ada banyak jenis penyakit yang dapat menyerang sapi kapan saja. Gejala-gejala yang muncul pada sapi perlu dideteksi sedini mungkin sebelum berpotensi menyebabkan kematian. Beberapa jenis penyakit sapi membawa dampak yang biasa saja, ada pula penyakit yang sangat parah sehingga bisa berdampak pada kematian [9]. Penyakit pada sapi potong dan penyakit pada sapi perah biasanya memiliki jenis yang sama, sehingga pada penanganannya tidak jauh berbeda. Penyebab serangan penyakit pada sapi diantaranya adalah bakteri, parasite, hingga virus [1], [10], [11]. Penyakit pada sapi bisa memberikan dampak negatif bagi peternak, seperti kerugian akibat kematian ternak, biaya yang dikeluarkan untuk pengobatan, penurunan produksi, dan efisiensi pakan yang menurun [7]. Selain itu, keterbatasan dokter hewan dan pengetahuan peternak mengenai penyakit sapi menyebabkan kesalahan dalam mendiagnosis penyakit sapi, sehingga kondisi tersebut perlu dicegah agar tidak menimbulkan kerugian peternak [6], [11]–[13]. Demikian hal itu, teknik data mining dapat berfungsi untuk mengatasi solusi terkait mengklasifikasikan penyakit pada bidang teknik informatika.

Beberapa penelitian telah membuktikan penggunaan teknik data mining mampu mengembangkan model untuk memprediksi atau mengklasifikasikan dan mendiagnosis penyakit pada sapi [14]–[17]. Data mining, atau yang dikenal sebagai *knowledge discovery in databases* (KDD), merupakan suatu teknik yang menggunakan data historis untuk mengidentifikasi pola dan hubungan dalam kumpulan data yang besar. Dengan pendekatan ini, data yang telah terakumulasi dari masa lalu dapat diolah untuk mengungkap informasi berharga yang dapat membantu dalam pengambilan keputusan di masa depan. Teknik data mining ini mencakup berbagai metode analisis data, seperti clustering, classification, regression, dan association. Dengan memanfaatkan teknik ini, organisasi dapat mengoptimalkan penggunaan data mereka untuk mendapatkan wawasan yang lebih dalam dan membuat keputusan yang lebih baik. [18]. Model-model yang telah dibangun pada penelitian sebelumnya membantu mengidentifikasi dan mendiagnosa penyakit pada sapi. Selain itu juga, model dibangun mampu menggantikan peran pakar dalam mendiagnosis penyakit sapi [6], [15], [16]. Sebelumnya, Algoritma FK-NN telah digunakan untuk menentukan penyakit pada ternak sapi dengan mendapatkan nilai K terbaik. Hasil penelitian tersebut menghasilkan rata-rata presentasi akurasi sebesar 93,8 % dan laju error 6,2 % [7]. Penelitian selanjutnya membangun suatu sistem pakar diagnosis penyakit sapi ternak potong menggunakan metode *Naive bayes - Certainty Factor*. Pada penelitian dilakukan dengan memasukkan gejala-gejala yang muncul pada penyakit sapi oleh pengguna. Kemudian dilakukan suatu perhitungan dengan metode *Naive Bayes- Certainty Factor*. Hasil dari penelitian ini

menghasilkan tingkat akurasi sebesar 92% [19]. Berdasarkan dari penelitian sebelumnya, maka penelitian ini bertujuan untuk menganalisis klasifikasi penyakit sapi menggunakan Teknik data mining.

## II. METODE PENELITIAN

Penelitian ini mengadopsi pendekatan metodologi yang terstruktur dan komprehensif untuk mengatasi kompleksitas masalah yang terkait dengan prediksi penyakit pada sapi. Pendekatan ini melibatkan empat tahap utama yang saling terkait: perumusan masalah dan studi literatur, pengumpulan data dan analisis data, pre-processing, serta hasil dan evaluasi model. Adapun alur dari penelitian ini ditunjukkan pada Gambar 2.1 berikut.



Gambar 2. 1 Alur penelitian

Berdasarkan Gambar 2.1 menunjukkan bahwa alur penelitian ini terdiri dari 4 trimester. Adapun trimester 1 ada pada tahap perumusan masalah dan studi literatur, penelitian ini melakukan analisis mendalam terhadap literatur yang relevan untuk memahami konteks masalah dan mengidentifikasi kerangka kerja serta metodologi yang sesuai untuk mengatasi permasalahan yang dihadapi. Pendekatan ini membantu dalam merumuskan pertanyaan penelitian yang tepat serta memastikan bahwa pendekatan yang digunakan akan memberikan solusi yang efektif. Selanjutnya trimester 2 yaitu tahap pengumpulan data dan analisa data yang melibatkan proses pengumpulan data dari Dinas Pertanian dan Ketahanan Pangan dan analisis data untuk mengidentifikasi pola dan hubungan antara atribut yang ada dalam dataset. Analisis data ini merupakan langkah penting dalam memahami karakteristik data serta mempersiapkan data untuk tahap selanjutnya. Langkah berikutnya adalah trimester 3 yaitu *pre-processing*. Tahap pre-processing dilakukan untuk meningkatkan kualitas data sebelum dilakukan analisis lebih lanjut. Teknik pre-processing yang digunakan meliputi cleaning data untuk menghapus missing value dan duplikat data. Langkah ini penting untuk memastikan bahwa data yang digunakan dalam analisis adalah data yang valid dan dapat diandalkan. Kemudian dilanjutkan dengan Trimester 4 yaitu hasil dan evaluasi model tahap hasil dan evaluasi model melibatkan implementasi model menggunakan teknik data mining yang telah dipilih, seperti *Naive bayes*, *Random Forest*, dan *SVM*, untuk memprediksi penyakit pada sapi. Evaluasi model dilakukan menggunakan *confusion matrix* untuk mengukur kinerja dan keefektifan model yang telah dibangun.

Dalam penelitian ini, digunakan tiga algoritma utama untuk prediksi penyakit pada sapi: *Naive Bayes*, *Random Forest*, dan *Support Vector Machine (SVM)*. Setiap algoritma memiliki prinsip dan rumus dasar yang digunakan untuk melakukan prediksi. Berikut tahapan yang dilakukan untuk setiap metode yang digunakan :

1. Dalam tahap persiapan data, langkah pertama yang dilakukan adalah memisahkan atribut-atribut yang digunakan sebagai fitur dan target. Dataset yang digunakan mengandung informasi mengenai gejala-gejala penyakit pada sapi, dimana kolom 'Penyakit' merupakan target yang menunjukkan jenis penyakit seperti PMK, LSD, Brucellosis, dan Anthrax. Sementara itu, atribut-atribut lain seperti 'nervous system disorders', 'blister-like lesions on the skin', dan sebagainya digunakan sebagai fitur yang menjadi input untuk model. Selanjutnya, karena model *Naive Bayes*, *Random Forest* dan *SVM* tidak dapat menangani nilai NaN, maka perlu dilakukan imputasi nilai untuk menggantikan nilai NaN dengan nilai rata-rata dari

- setiap kolom menggunakan SimpleImputer dari scikit-learn. Ini bertujuan untuk menjaga konsistensi data dan memastikan bahwa model dapat melakukan prediksi dengan baik tanpa terpengaruh oleh nilai yang hilang.
- Setelah persiapan data, langkah selanjutnya adalah membagi dataset menjadi dua subset: data latih dan data uji. Pembagian ini dilakukan dengan menggunakan fungsi `train_test_split` dari library scikit-learn, dimana data latih memiliki proporsi 80% dari total data dan data uji memiliki proporsi 20%. Data latih digunakan untuk melatih model machine learning, sedangkan data uji digunakan untuk menguji kinerja model yang telah dilatih.
  - Dalam tahap inialisasi model, tiga algoritma machine learning yang digunakan yaitu Naive Bayes, Random Forest, dan Support Vector Machine (SVM) dengan kernel RBF. Pertama, untuk model Naive Bayes, digunakan GaussianNB yang cocok untuk data kontinu. Karena dataset memiliki beberapa kelas target, model diubah menjadi multi-kelas menggunakan `OneVsRestClassifier` untuk penanganan multi-kelas. Kedua, untuk model Random Forest, digunakan `RandomForestClassifier` dengan parameter `n_estimators=100`, yang berarti hutan acak akan terdiri dari 100 pohon keputusan. Penggunaan `random_state=42` bertujuan untuk memastikan hasil yang konsisten dan dapat direproduksi. Ketiga, untuk model SVM dengan kernel RBF, digunakan SVC dengan `kernel='rbf'` yang menunjukkan penggunaan Radial Basis Function (RBF) untuk kernel SVM. Kernel RBF dipilih karena kemampuannya dalam menangani data yang tidak dapat dipisahkan secara linear. Penggunaan `random_state=42` pada model ini juga digunakan untuk memastikan hasil yang konsisten.
  - Melatih model. Model Naive Bayes dilatih dengan menggunakan data latih (`X_train` dan `y_train`), di mana proses pelatihan melibatkan perhitungan probabilitas posterior untuk setiap kelas penyakit berdasarkan fitur yang diberikan. Sementara itu, model Random Forest juga dilatih dengan data latih, dengan proses pelatihan yang melibatkan pembentukan beberapa pohon keputusan (dengan sampel acak dari data latih) dan penggabungan hasil mereka untuk membuat prediksi yang lebih akurat. Sedangkan untuk model SVM, proses pelatihannya melibatkan pencarian hyperplane yang memaksimalkan margin antara kelas yang berbeda dalam ruang fitur yang ditransformasikan oleh kernel RBF. Dengan proses pelatihan ini, ketiga model siap untuk diuji dan dievaluasi menggunakan data uji.
  - Prediksi dan evaluasi. Model memprediksi kelas untuk data uji (`X_test`). Akurasi, presisi, dan recall dihitung untuk mengukur kinerja model.
  - Visualisasi Kurva ROC. Kurva ROC untuk model *Naive Bayes* diplot bersama dengan model lain yaitu *Random Forest* dan *SVM* untuk membandingkan kinerja.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan hasil pengumpulan data dari dinas pertanian dan ketahanan pangan. Data disajikan dalam format .csv dengan jumlah total 100 data dan terdiri dari 44 atribut. Salah satu atribut yang menarik dalam dataset ini adalah atribut penyakit yang mengindikasikan jenis penyakit yang diderita oleh sapi. Atribut ini terdiri dari beberapa jenis penyakit, antara lain PMK, LSD, Brucellosis, dan Anthrax. Adapun dataset tersebut ditunjukkan pada Tabel 3.1

Tabel 3. 1 Dataset penyakit sapi

Penyakit	A	B	C	...	AO	AP	AQ
PMK	1	1	1		0	0	0
PMK	0	1	1		0	0	0
PMK	1	1	0		0	0	0
...							
Anthrax	1	1	0		1	0	0
Anthrax	0	1	0		1	1	1
Anthrax	1	1	0		1	0	0

Berdasarkan Tabel 3.1 menunjukkan dataset penyakit sapi dengan 44 atribut. Adapun penjelasan dari 44 atribut tersebut ditunjukkan pada Tabel 3.2.

Tabel 3. 2 Deskripsi Atribut

Atribut	Deskripsi
A	Demam
B	nafsu makan turun
C	hipersaliva
D	radang pada mulut dan lidah
E	lepuh-lepuh pada kuku, puting dan ambing (sapi betina)
F	kekakuan kaki belakang dan pincang
G	nafas cepat dan terengah-engah
H	hipersativasi
I	kehilangan berat badan
J	pembengkakan kelenjar submandibular

Atribut	Deskripsi
K	lesi ekstensif pada kulit berupa nodul dengan ukuran 5–20 mm
L	pembesaran nodus limfa tikus
M	peningkatan opasitas kornea
N	gangguan respirasi disertai dengan leleran hidung/ingus
O	Depresi
P	penurunan produksi susu
Q	ingusan
R	konjungtivitis
S	pembengkakan limfoglandula yaitu Lgl. subscapularis dan Lgl. prefemoral
T	terdapat nodul pada kulit yang berbatas, jelas dan menonjol di bawah kulit atau di bawah otot dengan diameter antara 2-5 cm
U	lesi kulit berupa nodul berukuran 1-7 cm yang biasanya ditemukan pada daerah leher, kepala, kaki, ekor dan ambing.
V	adanya leleran mata
W	pembengkakan limfonodus subscapula dan prefemoralis, serta dapat terjadi oedema pada kaki
X	abortus
Y	infertilitas
Z	Lemah
AA	Hygroma
AB	Mastitis
AC	Retensi uterus
AD	keguguran pada fase akhir kebuntingan
AE	keluarnya cairan keluron
AF	menimbulkan gangguan pertumbuhan
AG	orchitis
AH	epididimitis
AI	seminal vesikulitis
AJ	Abses testis
AK	metritis
AL	retensio plasenta
AM	edema di sekitar leher, hidung, kepala dan scrotum
AN	hewan terlihat sempoyongan
AO	gemetar
AP	Lemas
AQ	Kepala diletakan ditanah

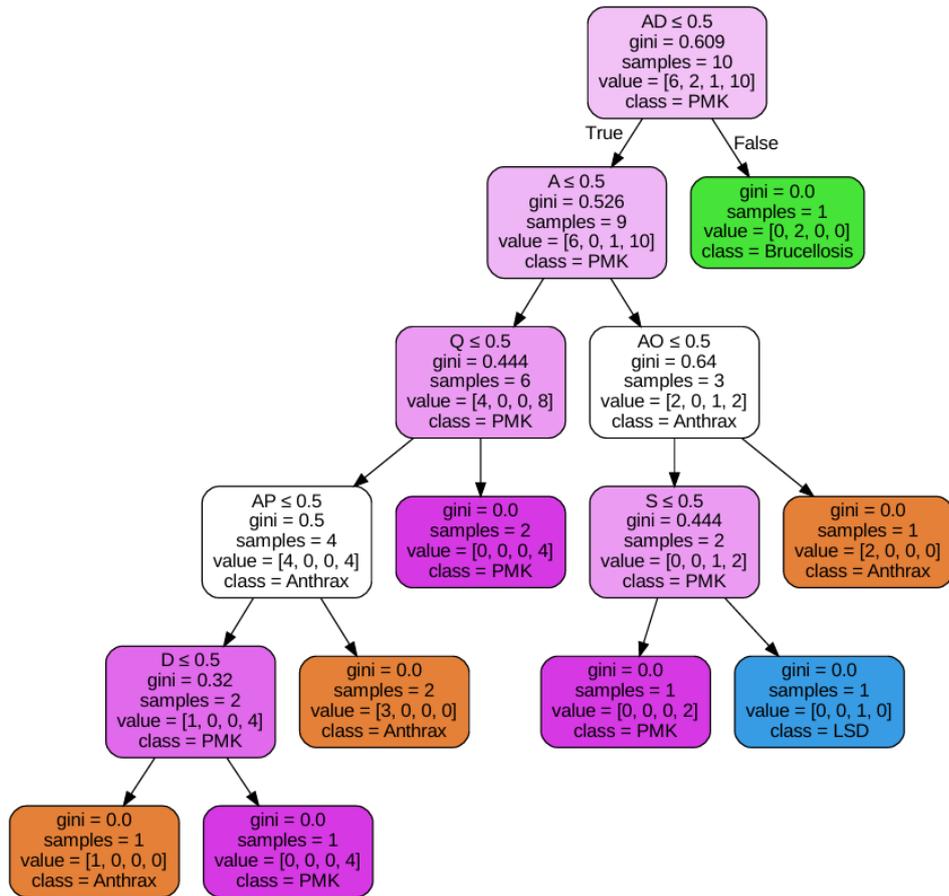
### B. Hasil *Preprocessing*

Pada tahap ini dilakukan tahap *Preprocessing* dengan Langkah-langkah sebagai berikut:

- Pembersihan Data:** Data telah diperiksa untuk mengidentifikasi dan mengatasi nilai yang hilang atau tidak valid. Hasilnya tidak ditemukan data yang kosong maupun duplikasi data.
- Pemilihan Fitur:** Dalam kasus ini, fokus utama adalah pada atribut penyakit sapi (PMK, LSD, Brucellosis, Anthrax). Fitur-fitur lain yang tidak relevan atau tidak memberikan informasi yang berguna untuk analisis telah dihilangkan dari dataset.
- Normalisasi:** Atribut yang memiliki skala yang berbeda-beda telah dinormalisasi untuk memastikan bahwa setiap atribut memiliki pengaruh yang seimbang dalam analisis.
- Pengelompokan Data:** Data telah dikelompokkan berdasarkan atribut penyakit untuk memudahkan analisis. Hal ini akan memungkinkan untuk melihat pola dan hubungan antara atribut penyakit dengan lebih baik.
- Encoding:** Atribut kategori seperti jenis penyakit telah diubah menjadi representasi numerik agar dapat diproses oleh algoritma machine learning.

### C. Hasil implementasi dan evaluasi model

Sebelum tahap implementasi model, maka tahap yang perlu dilakukan yaitu partisi dataset. Dataset dipartisi menjadi data uji dan data latih. Dataset tersebut dipartisi menggunakan Teknik Split data dengan perbandingan 80:20. Adapun data uji terdiri dari 5 data dan data latih sebanyak 19 data. Selanjutnya dilakukan implementasi model dengan 3 kali percobaan dengan *algoritma Naïve bayes*, *Random Forest* dan *Support Vector Machine*. Adapun hasil dari implementasi menggunakan *algoritma Random forest* ditunjukkan pada Gambar 3.1.



Gambar 3. 1 pohon keputusan

Berdasarkan Gambar 3.1 menunjukkan hasil implementasi model random forest yang divisualisasikan dengan pohon keputusan. Oleh karena itu, visualisasi tersebut menghasilkan beberapa keputusan. Adapun keputusan tersebut sebagai berikut:

1. Jika  $AD \leq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,609 dari total 10 sampel
2. Jika  $AD \geq 0.5$  dan  $A \leq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,526 dari total 9 sampel. Namun ada satu kondisi yang menunjukkan klasifikasi Brucellosis dengan nilai gini 0,0 dari satu sampel.
3. Jika  $A \geq 0.5$  dan  $Q \leq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,444 dari total 6 sampel, Tapi jika  $A \geq 0.5$  dan  $AO \leq 0.5$  maka klasifikasi menunjukkan Anthrax dengan nilai gini 0,64 dari total 3 sampel.
4. Jika  $AO \geq 0.5$  dan  $S \leq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,444 dari total 2 sampel. Namun ada satu kondisi yang menunjukkan klasifikasi Anthrax dengan nilai gini 0,0 dari satu sampel.
5. Jika  $AO \geq 0.5$  dan  $S \geq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,0 dari 1 sampel. Namun ada satu kondisi yang menunjukkan klasifikasi LSD dengan nilai gini 0,0 dari satu sampel.
6. Jika  $Q \geq 0.5$  dan  $AP \leq 0.5$  maka klasifikasi menunjukkan Anthrax dengan nilai gini 0,0 dari 4 sampel. Namun ada kondisi yang menunjukkan klasifikasi PMK dengan gini 0,0 dari 1 sampel.
7. Jika  $AP \geq 0.5$  dan  $D \leq 0.5$  maka klasifikasi menunjukkan PMK dengan nilai gini 0,32 dari 2 sampel. Namun ada kondisi yang menunjukkan klasifikasi Anthrax dengan gini 0,0 dari 1 sampel.
8. Jika  $D \geq 0.5$  maka klasifikasi menunjukkan Anthrax dengan gini 0,0 dari 1 sampel dan kondisi lain menunjukkan klasifikasi PMK dari 1 sampel.

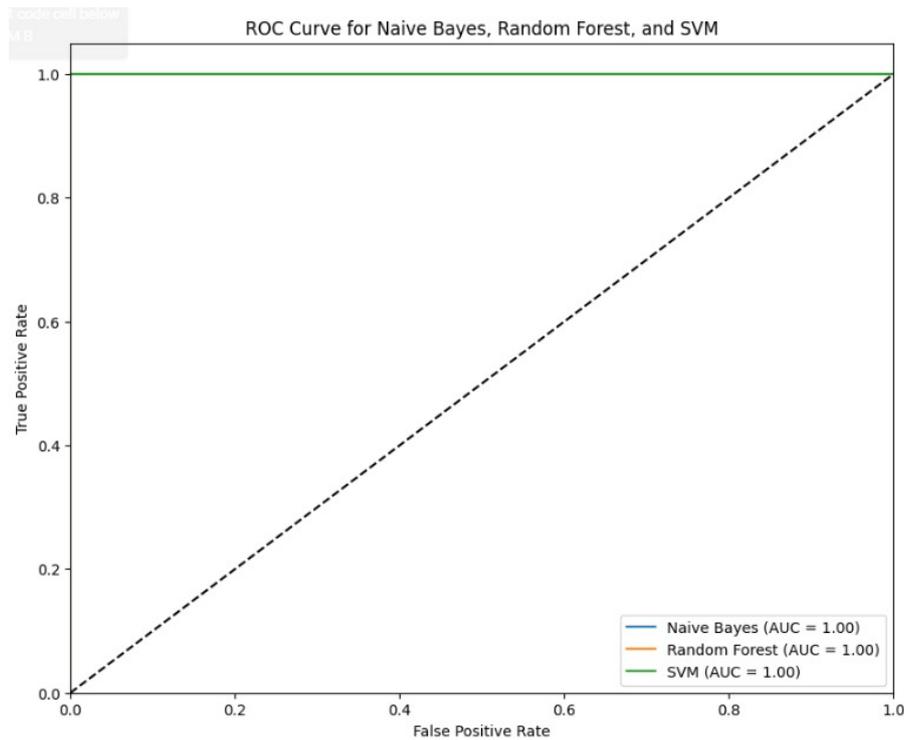
Kemudian evaluasi model yang digunakan yaitu *confusion matrix* yang terdiri dari *accuracy*, *presicion* dan *recall*. Adapun hasil tersebut ditunjukkan pada Tabel 3.2

Tabel 3. 3 Hasil evaluasi model

Metode	Accuracy	Precision	Recall
Naive Bayes	0.8	0.8	0.8
Random Forest	0.8	0.8	0.8
Support Vector Machine	0.8	0.6	0.8

Berdasarkan Tabel 3.1 menunjukkan bahwa implementasi model dengan algoritma *naive bayes* menghasilkan akurasi sebesar 0,8, presisi 0,8 dan *recall* sebesar 0,8. Sedangkan, implemenetasi model dengan menggunakan algoritma *Support Vector Machine* menghasilkan akurasi sebesar 0,8, pesisi sebesar 0,6 dan *recall* sebesar 0,8. Berdasarkan hal itu bahwa implementasi model dengan

3 kali percobaan menunjukkan tidak adanya perbedaan yang signifikan antara akurasi dan *recall* dengan algoritma yang digunakan. Namun, untuk *precision* menunjukkan adanya perbedaan yang signifikan antara *Naïve Bayes*, *Random Forest* dengan *Support Vector Machine* yaitu sebesar 0.2. Adapun nilai ROC dari setiap model dapat dilihat pada Gambar 3.2.



Gambar 3. 2 ROC Curve for Naive Bayes, Random Forest dan SVM

Berdasarkan Gambar 3.2 menunjukkan bahwa nilai ROC dari setiap model tidak ada perbedaan yang signifikan yaitu 1.00. Demikian hal itu, dari setiap metode yang digunakan untuk klasifikasi penyakit sapi cenderung memiliki performa yang baik.

#### D. Interpretasi Hasil

Berdasarkan hasil mining data penyakit sapi dengan algoritma *Naïve Bayes*, *Random Forest*, dan *Support Vector Machine* (SVM) dengan data sebagai berikut:

- Naïve Bayes*: Menghasilkan akurasi, presisi, dan *recall* sebesar 0.8. Hal ini menunjukkan bahwa model *Naïve Bayes* mampu memprediksi dengan benar sekitar 80% dari semua data yang diproses. Presisi 0.8 menunjukkan bahwa dari semua prediksi yang dilakukan oleh model yang diklasifikasikan sebagai positif, sekitar 80% benar-benar positif. *Recall* 0.8 menunjukkan bahwa dari semua kasus yang sebenarnya positif, model mampu mengidentifikasi sekitar 80% dari mereka.
- Random Forest*: Percobaan kedua menggunakan algoritma *Random Forest* dan menghasilkan nilai akurasi, presisi, dan *recall* yang sama dengan *Naïve Bayes*, yaitu 0.8. Ini menunjukkan bahwa *Random Forest* memiliki kinerja yang setara dengan *Naïve Bayes* dalam hal prediksi penyakit sapi.
- Support Vector Machine* (SVM): Percobaan ketiga menggunakan algoritma *Support Vector Machine* (SVM) menghasilkan akurasi sebesar 0.8, presisi 0.6, dan *recall* 0.8. Meskipun akurasi sama dengan *Naïve Bayes* dan *Random Forest*, presisi yang lebih rendah (0.6) menunjukkan bahwa SVM cenderung memberikan lebih banyak false positive dalam prediksinya, yaitu sekitar 40% dari prediksi positif yang sebenarnya negatif. *Recall* yang tinggi (0.8) menunjukkan bahwa SVM mampu mengidentifikasi sebagian besar kasus penyakit positif dengan benar.

Berdasarkan hasil ini, dapat disimpulkan bahwa *Naïve Bayes* dan *Random Forest* memiliki kinerja yang setara dan lebih baik daripada SVM dalam memprediksi jenis penyakit pada sapi. Meskipun SVM memiliki *recall* yang tinggi, presisi yang rendah menunjukkan bahwa model cenderung memberikan lebih banyak false positive. Dalam konteks manajemen kesehatan hewan sapi, penting untuk memilih model yang dapat memberikan prediksi yang akurat dan presisi yang tinggi untuk menghindari kesalahan dalam pengambilan keputusan terkait kesehatan hewan.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa model klasifikasi yang dibangun dengan menggunakan algoritma *Naive Bayes*, *Random Forest*, dan *Support Vector Machine* mampu mengklasifikasikan data dengan baik. Ketiga algoritma tersebut menghasilkan akurasi, presisi, dan *recall* yang baik dalam memprediksi jenis penyakit pada sapi.

Percobaan pertama dengan algoritma *Naive bayes* menghasilkan akurasi, presisi, dan recall sebesar 0.8. Percobaan kedua dengan algoritma *Random Forest* menghasilkan hasil yang sama dengan *Naive bayes*. Sementara itu, percobaan ketiga dengan algoritma *Support Vector Machine* menghasilkan akurasi 0.8, presisi 0.6, dan recall 0.8.

Untuk penelitian lanjutan, disarankan untuk melakukan analisis lebih lanjut terhadap faktor-faktor yang mempengaruhi kinerja algoritma. Beberapa faktor yang perlu diperhatikan adalah pemilihan fitur, preprocessing data, dan tuning parameter. Analisis ini dapat membantu meningkatkan performa dan ketepatan model klasifikasi yang telah dibangun sehingga dapat memberikan prediksi yang lebih akurat dan handal dalam mendiagnosis penyakit pada sapi.

#### DAFTAR PUSTAKA

- [1] Marwan Hakim, "Sistem Pakar Mengidentifikasi Penyakit Pada Sapi Menggunakan Metode Forward Chaining," *Tek. Teknol. Inf. dan Multimed.*, vol. 3, no. 2, pp. 106–111, 2022, doi: 10.46764/teknimedia.v3i2.81.
- [2] R. Rahmiati and N. A. Temesveri, "Hubungan Dimensi Kualitas Pelayanan Dengan Minat Kunjungan Ulang Pasien Di Instalasi Rawat Jalan Rumah Sakit Umum Kabupaten Tangerang Tahun 2019," *J. Kesehat.*, vol. 13, no. 1, pp. 13–21, 2020, doi: 10.23917/jk.v13i1.11097.
- [3] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.
- [4] F. Sani and A. Annisa, "肖沉 1, 2, 孙莉 1, 2Δ, 曹杉杉 1, 2, 梁浩 1, 2, 程焱 1, 2," *Tjyybjb.Ac.Cn*, vol. 27, no. 2, pp. 635–637, 2019.
- [5] Bunga Ihda Norra, T. Putri Hendrika, A. Auliyaur Rohmah, and Ila nabinya, "IDENTIFIKASI PEMAHAMAN UMUM AYAM (*Gallus gallus*) DAN IKAN MUJAIR (*Oreochromis Mossambicus*) PADA MAHASISWA UIN WALISONGO SEMARANG," *Bio-Lectura*, vol. 8, no. 1, pp. 29–36, 2021, doi: 10.31849/bl.v8i1.5763.
- [6] W. D. Prasetyo and R. Wahyudi, "Sistem Pakar Diagnosis Penyakit Ternak Sapi Menggunakan Metode Forward Chaining Berbasis Website Responsif," *J. Teknol. dan Terap. Bisnis*, vol. 2, no. 1, pp. 13–21, 2019.
- [7] D. M. Nuraini, S. Sunarto, N. Widyas, A. Pramono, and S. Prastowo, "Peningkatan Kapasitas Tata Laksana Kesehatan Ternak Sapi Potong di Pelemrejo, Andong, Boyolali," *PRIMA J. Community Empower. Serv.*, vol. 4, no. 2, p. 102, 2020, doi: 10.20961/prima.v4i2.42574.
- [8] imam, "1633-2002-3-Pb," vol. 9, no. 2, pp. 75–80, 2020.
- [9] G. B. Kristi, N. Hidayat, and E. Santoso, "Sistem Diagnosis Penyakit Pada Sapi Potong menggunakan Metode Bayesian Network," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 1, p. 5, 2019.
- [10] A. Subhan, S. Nurawaliah, and M. Syarif, *Inovasi Teknologi Pengembangan Usaha Ternak Sapi*. 2022. [Online]. Available: [http://repository.pertanian.go.id/handle/123456789/18695%0Ahttp://repository.pertanian.go.id/bitstream/handle/123456789/18695/bk103\\_kalsel.pdf?sequence=1&isAllowed=y](http://repository.pertanian.go.id/handle/123456789/18695%0Ahttp://repository.pertanian.go.id/bitstream/handle/123456789/18695/bk103_kalsel.pdf?sequence=1&isAllowed=y)
- [11] M. D. O. Amaral Barreto and R. D. Indahsari, "Cattle Disease Diagnosis Expert System in Banded Village, Ainaro District," *Procedia Eng. Life Sci.*, vol. 1, no. 2, 2021, doi: 10.21070/pels.v1i2.1009.
- [12] M. Agarina and A. S. Karim, "Rancang Bangun Sistem Informasi Kegiatan Seminar Nasional Berbasis Web Pada Institut Informatics Dan Bisnis Darmajaya," *Explor. J. Sist. Inf. dan Telemat.*, vol. 10, no. 1, 2019, doi: 10.36448/jsit.v10i1.1215.
- [13] N. Anggraini and R. F. F. Afidh, "Sistem Pakar Diagnosa Penyakit Sapi Menggunakan Metode CBR Dan Algoritma Similarity Sorgenfrei," vol. 2, no. 1, pp. 1–10, 2023.
- [14] Aristoteles, K. Adhianto, R. Andrian, and Y. N. Sari, "Comparative analysis of cow disease diagnosis expert system using Bayesian network and Dempster-Shafer method," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 4, pp. 227–235, 2019, doi: 10.14569/ijacsa.2019.0100427.
- [15] Y. Malelak and J. H. Tomasoey, "Penerapan Metode Fuzzy K-Nearest Neighbor (Fk-Nn) Untuk Menentukan Penyakit Pada Ternak Sapi Potong," *High Educ. Organ. Arch. Qual. J. Teknol. Inf.*, vol. 10, no. 2, pp. 66–72, 2018, doi: 10.52972/hoaq.vol10no2.p66-72.
- [16] H. Mursalan and Sumijan, "Akurasi dalam Identifikasi Penyakit Sapi Pesisir Menggunakan Metode Forward Chaining," *J. Inf. dan Teknol.*, vol. 3, pp. 115–120, 2021, doi: 10.37034/jidt.v3i3.117.
- [17] N. V. Kishan, S. T. Y, and S. Kavalur, "Cattle disease identification using Prediction Techniques," no. 3, pp. 2764–2770, 2021.
- [18] M. A. K-medoids, "Clustering Vaksinasi Penyakit Mulut dan Kuku Di Provinsi Riau," vol. 5, pp. 90–98, 2023.
- [19] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, "Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM," *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.