



Analisis Klasifikasi Hewan Menggunakan Metode *K-Nearest Neighbor*, *Decision Tree*, dan *Naïve Bayes*

Mukhlisah Syaukati Robbi

Teknik Informatika, Program Pascasarjana, Universitas Pamulang, Tangerang Selatan, Banten

Email: mukhlisahsr@gmail.com

ABSTRACT

Animal classification is an important topic in biology and conservation science, appropriate data analysis methods can help in identifying, classifying, and understanding animal characteristics. This research aims to analyze and compare various animal classification methods using the K-Nearest Neighbor, Decision Tree, and Naïve Bayes methods by implementing them in the Orange Data Mining environment. This study uses a dataset that includes a wide range of biological, morphological, and behavioral attributes of animals. Through the implementation of the orange tools, three different classification methods were developed and evaluated based on their performance in classifying animals into groups according to their characteristics. The research results show that the performance of the K-Nearest Neighbor method is superior to the others. From 101 data tested using the K-Nearest Neighbor method, an accuracy value of 94.1% was obtained. Comparative analysis reveals differences in accuracy and predictive ability between K-Nearest Neighbor, Decision Tree, and Naïve Bayes. These results provide insight into the effectiveness of each method in the context of animal classification and can potentially serve as a basis for selecting appropriate methods in biological research, conservation, or other animal science studies. This research enriches understanding of orange tools in the context of biology and animal science.

Keywords: Classification; Data Mining; Decision Tree; K-Nearest Neighbor; Naive Bayes

ABSTRAK

Klasifikasi hewan adalah topik penting dalam ilmu biologi dan konservasi, metode analisis data yang tepat dapat membantu dalam mengidentifikasi, mengklasifikasikan, dan memahami karakteristik hewan. Tujuan dari penelitian ini adalah untuk menganalisis dan membandingkan berbagai metode klasifikasi spesies hewan menggunakan metode *K-Nearest Neighbor* (KNN), *Decision Tree*, dan *Naïve Bayes* dengan mengimplementasikannya ke dalam lingkungan *Orange Data Mining*. Studi ini menggunakan dataset yang mencakup berbagai atribut biologis, morfologis, dan perilaku hewan. Melalui implementasi *tools Orange*, tiga metode klasifikasi yang berbeda dikembangkan dan dievaluasi berdasarkan kinerja dalam mengklasifikasikan hewan ke dalam kelompok yang sesuai dengan karakteristik mereka. Hasil penelitian menunjukkan bahwa kinerja metode *K-Nearest Neighbor* lebih unggul dibanding yang lainnya. Dari 101 data yang diuji menggunakan metode *K-Nearest Neighbor* dihasilkan nilai akurasi sebesar 94,1%. Analisis komparatif mengungkapkan perbedaan dalam akurasi dan kemampuan prediksi antara *K-Nearest Neighbor*, *Decision Tree*, dan *Naïve Bayes*. Hasil ini memberikan wawasan tentang keefektifan masing-masing metode dalam konteks klasifikasi hewan dan berpotensi menjadi dasar pemilihan metode yang sesuai dalam penelitian biologi, konservasi, atau studi ilmu hewan lainnya. Penelitian ini memperkaya pemahaman tentang *tools Orange* dalam konteks ilmu biologi dan ilmu hewan.

Kata kunci: *Data Mining, Decision Tree, K-Nearest Neighbor, Klasifikasi, Naive Bayes*

1. PENDAHULUAN

Data mining adalah rangkaian suatu proses dengan tujuan untuk mengekstraksi nilai tambah secara manual dari sekumpulan data berupa pengetahuan yang saat ini tidak diketahui [1]. Salah satu aspek penting dari penelitian ini adalah klasifikasi hewan berdasarkan ciri dan atribut yang terkait. Taksonomi hewan diperlukan untuk memahami perilaku, ciri morfologi, dan ekologi berbagai jenis hewan, mendukung upaya konservasi dan ilmu biologi [2]. Dalam upaya untuk mengklasifikasi hewan dengan lebih akurat dan efisien, berbagai metode dan alat bantu analisis data telah dikembangkan.

Dalam lingkungan ilmu data, teknik-teknik *data mining* telah membuka peluang baru dalam pemahaman dan analisis spesies-spesies hewan. Salah satu *tools* yang cukup populer untuk tujuan ini adalah "*Orange Data Mining*". *Tools* ini merupakan aplikasi *data mining* yang dapat menghitung secara otomatis untuk membantu penelitian dalam menganalisa suatu data [3]. *Tools orange* memiliki *widget* yang berfungsi sebagai unit komputasi untuk memproses, membaca, memvisualisasikan, menganalisis, menjelajahi data, dan lain sebagainya. *Widget* ini dirangkai sedemikian rupa hingga membentuk alur kerja [4]. Dalam rangka mengkaji efektivitas dalam konteks klasifikasi hewan, dilakukan penelitian untuk membandingkan tiga metode klasifikasi yang berbeda, yaitu *K-Nearest Neighbor*, *Decision Tree*, dan *Naïve Bayes*.

Dalam penelitian yang dilakukan oleh [5] menggunakan metode *Naïve Bayes* mampu mengklasifikasikan secara akurat 101 spesies hewan yang ditemukan dalam dataset zoo.tab. Secara spesifik, dengan akurasi klasifikasi 0,911. Klasifikasinya dibagi menjadi 7 kategori, antara lain amfibi, burung, ikan, serangga, invertebrata, mamalia, dan reptil. Tingkat prediksi menunjukkan terdapat tiga klasifikasi prediksi dengan nilai 100%, satu prediksi dengan nilai 92,9%, satu prediksi dengan nilai 80%, satu prediksi dengan nilai 66,7%, dan satu prediksi bernilai 33,3%.

Penelitian ini menggabungkan konsep ilmu data dengan pengetahuan biologi dan konservasi, memberikan pemahaman yang lebih baik tentang cara menggunakan *tools Orange Data Mining* untuk mengklasifikasikan hewan. Metode klasifikasi dapat berdampak signifikan terhadap hasil dan kesalahan klasifikasi, sehingga perbandingan antara metode yang berbeda menjadi diperlukan.

2. METODE

2.1. Tahap Pengumpulan Data

Penelitian ini menggunakan dataset kebun binatang atau zoo.tab, yang merupakan dataset bawaan dari *tools Orange Data Mining*. Dataset berisikan enam belas atribut dengan 101 data spesies hewan yang akan diklasifikasi ke dalam kelas hewan seperti *amphibian, bird, fish, insect, invertebrate, mammal, dan reptile*.

2.2. Teknik Data Mining

Teknik yang digunakan dalam penelitian ini adalah teknik algoritma *K-Nearest Neighbor, Naïve Bayes, dan Decission Tree*. Mengutip dari jurnal [6], untuk menentukan kinerja nilai akurasi ketiga metode *data mining* yang digunakan antara lain *K-Nearest Neighbor, Naïve Bayes, dan Decission Tree* antara lain sebagai berikut.

2.2.1. K- Nearest Neighbor

K-Nearest Neighbor (KNN) merupakan metode *supervised* klasifikasi objek baru yang artinya memerlukan informasi pelatihan untuk mengklasifikasikan objek dengan jarak terdekat. Prinsip pengoperasian metode ini adalah mencari jarak terpendek antara informasi yang akan dievaluasi dan nilai k tetangga dalam data pelatihan. Rumus yang dipakai untuk menghitung jarak terdekat (*neighbor*) secara umum sebagai berikut.

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

Dimana d adalah jarak tetangga terdekat. x adalah data latih ke- i . y adalah data uji ke- i . n adalah jumlah data.

Langkah-langkah melakukan klasifikasi dengan algoritma KNN yaitu, langkah pertama menentukan nilai k (jumlah tetangga terdekat), langkah kedua melakukan perhitungan nilai jarak (*euclidean distance*), langkah ketiga mengurutkan hasil perhitungan jarak secara *ascending* (dari nilai rendah ke tinggi), langkah keempat melakukan pengelompokkan data sesuai dengan nilai k , lalu dengan menggunakan kategori paling banyak akan dapat diprediksi kategori objeknya [7].

2.2.2. Decision Tree

Pohon keputusan adalah teknik mengklasifikasi yang biasa digunakan untuk mengekstraksi hubungan yang bersangkutan dalam data. Sebuah metode yang membuat pohon keputusan berdasarkan sekumpulan data masukan berlabel. Atribut dengan

frekuensi tinggi dianggap memisahkan data berdasarkan informasi yang tersedia dalam kumpulan data. Hitung nilai *entropy* dan nilai *gain* dengan rumus berikut [8].

$$Entropy(S) = \sum_{i=1}^n -p_i \log_2 p_i \quad (2)$$

Dimana S adalah himpunan kasus. n adalah jumlah partisi. p_i adalah proporsi dari S_i terhadap S . Lalu tentukan nilai *Gain* dengan rumus menghitung informasi *Gain* sebagai berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (3)$$

Dimana S adalah himpunan kasus. A adalah atribut. n jumlah partisi atribut A . S_i adalah jumlah kasus pada partisi ke- i . S jumlah kasus dalam S .

2.2.3. Naïve Bayes

Naïve Bayes adalah prosedur klasifikasi probabilistik sederhana. Metode ini menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan campuran nilai dari kumpulan data tertentu. Metode *Naïve Bayes* percaya bahwa semua atribut dari setiap kategori tidak bergantung satu sama lain. Berikut rumus yang digunakan pada teorema bayes [9].

$$P(H|X) = \frac{P(X|H)*P(H)}{P(X)} \quad (4)$$

Dimana X adalah data uji dengan kelas yang belum diketahui. H adalah hipotesis data X yang merupakan suatu kelas spesifik. $P(H|X)$ adalah nilai probabilitas hipotesis H berdasarkan kondisi X (*posteriori probability*). $P(H)$ adalah nilai probabilitas pada hipotesis H (*prior probability*). $P(X|H)$ adalah nilai probabilitas X berdasarkan kondisi H (*likelihood*). $P(X)$ adalah nilai probabilitas X (*predictor prior probability*).

2.3. Pengujian Algoritma

Pengujian algoritma *K-Nearest Neighbor*, *Naïve Bayes*, dan *Decision Tree* dilakukan untuk menilai kinerja algoritma dan mengetahui tingkat keberhasilan algoritma dalam mengklasifikasikan jenis hewan. Pengujian dalam penelitian ini menerapkan *Stratified K-Fold Cross Validation* yang diproses dalam *tools Orange* dengan menentukan nilai k secara manual menggunakan *widget test and score*. Dari *widget test and score* dapat memilih metode pengujian apa yang akan dipakai. *Widget* ini juga menampilkan nilai-nilai hasil klasifikasi metode algoritma yang digunakan seperti nilai *accuracy*, *precision*, *recall*, *f-measure*, dan nilai-nilai lainnya.

Di dalam *tool orange* untuk hasil klasifikasi hewan dapat ditampilkan atau dilihat dengan menggunakan *widget confusion matrix*. Dalam penggunaan *confusion matrix* terdapat empat istilah untuk menggambarkan hasil klasifikasi, yaitu *True Positive*, *True Negative*, *False Positive*, serta *False Negative* [10]. Berikut tampilan bentuk tabel dari *confusion matrix* [11].

Tabel 1. *Confusion Matrix*

<i>Classification</i>		<i>Predicted Class</i>	
		<i>True</i>	<i>False</i>
<i>Actual Class</i>	<i>True</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
	<i>False</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Berdasarkan tabel 1, nilai *true positive* adalah jumlah data positif yang terdeteksi dengan benar atau data perkiraan sesuai dengan data yang sesungguhnya. Sedangkan nilai *false positive* adalah jumlah data negatif yang terdeteksi tidak sesuai atau terdeteksi sebagai positif, yaitu data perkiraan tidak sesuai dengan data sesungguhnya.

Sementara nilai *true negative* adalah jumlah data negatif yang terdeteksi dengan sesuai, data perkiraan sama dengan data sesungguhnya. Dan nilai *false negative* adalah jumlah data negatif yang terdeteksi tidak sesuai atau terdeteksi sebagai positif, data perkiraan tidak sama dengan data sesungguhnya.

Dari istilah-istilah tersebut didapat nilai *accuracy*, *precision*, *recall* dan *f-measure*. Nilai *accuracy* untuk mengevaluasi suatu model algoritma agar diketahui ketepatan nilai prediksi dengan nilai aktualnya. Nilai *precision* untuk mengetahui seberapa besar tingkat keberhasilan suatu model algoritma dalam memberikan jawaban yang tepat. Nilai *recall* untuk mengetahui seberapa besar tingkat keberhasilan suatu model algoritma dalam menemukan kembali informasi yang akurat. Sedangkan *f1-score* atau *f-measure* untuk mengetahui hasil perbandingan antara nilai *precision* dengan nilai *recall*.

3. HASIL DAN PEMBAHASAN

3.1. Dataset

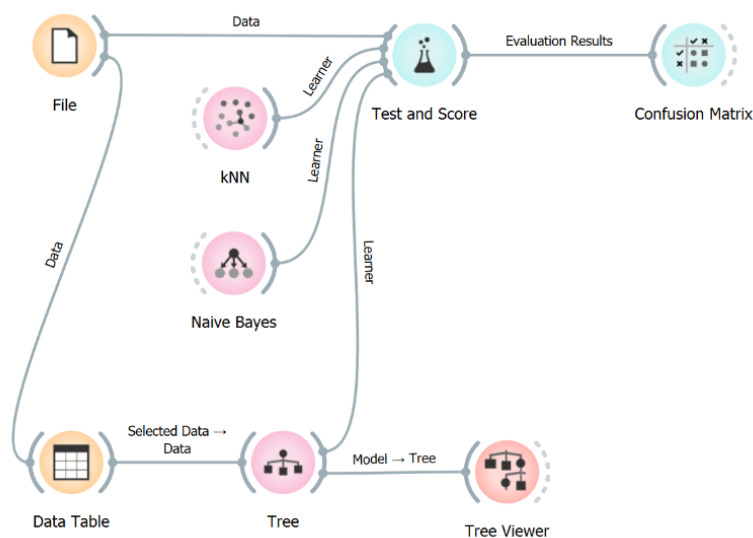
Dataset yang digunakan merupakan bawaan dari *tools Orange Data Mining*. Nama dataset yang digunakan adalah *zoo.tab*. Pada dataset ini terdapat 101 data *instances* dan 16 atribut. Dengan target 7 nilai yaitu, *mammal*, *fish*, *bird*, *invertebrate*, *insect*, *amphibian*, dan *reptile*. Berikut adalah tampilan data tabel dari dataset *zoo.tab* yang sudah diimplementasikan pada *tools orange*.

	type	name	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed
1	mammal	aardvark	1	0	0	1	0	0	1	1
2	mammal	antelope	1	0	0	1	0	0	0	1
3	fish	bass	0	0	1	0	0	1	1	1
4	mammal	bear	1	0	0	1	0	0	1	1
5	mammal	boar	1	0	0	1	0	0	1	1
6	mammal	buffalo	1	0	0	1	0	0	0	1
7	mammal	calf	1	0	0	1	0	0	0	1
8	fish	carp	0	0	1	0	0	1	0	1
9	fish	catfish	0	0	1	0	0	1	1	1
10	mammal	cavy	1	0	0	1	0	0	0	1
11	mammal	cheetah	1	0	0	1	0	0	1	1
12	bird	chicken	0	1	1	0	1	0	0	0
13	fish	chub	0	0	1	0	0	1	1	1
14	invertebrate	clam	0	0	1	0	0	0	1	0
15	invertebrate	crab	0	0	1	0	0	1	1	0
16	invertebrate	crayfish	0	0	1	0	0	1	1	0
17	bird	crow	0	1	1	0	1	0	1	0
18	mammal	deer	1	0	0	1	0	0	0	1
19	fish	dogfish	0	0	1	0	0	1	1	1
20	mammal	dolphin	0	0	0	1	0	1	1	1
21	bird	dove	0	1	1	0	1	0	0	0
22	bird	duck	0	1	1	0	1	1	0	0
23	mammal	elephant	1	0	0	1	0	0	0	1
24	bird	flamingo	0	1	1	0	1	0	0	0
25	insect	flea	0	0	1	0	0	0	0	0
26	amphibian	frog	0	0	1	0	0	1	1	1

Gambar 1. Data Tabel Zoo

3.2. Proses Data Mining

Dengan menganalisis kinerja beberapa metode klasifikasi, perbandingan dilakukan pada *tools Orange* untuk memilih metode terbaik dengan akurasi tinggi.

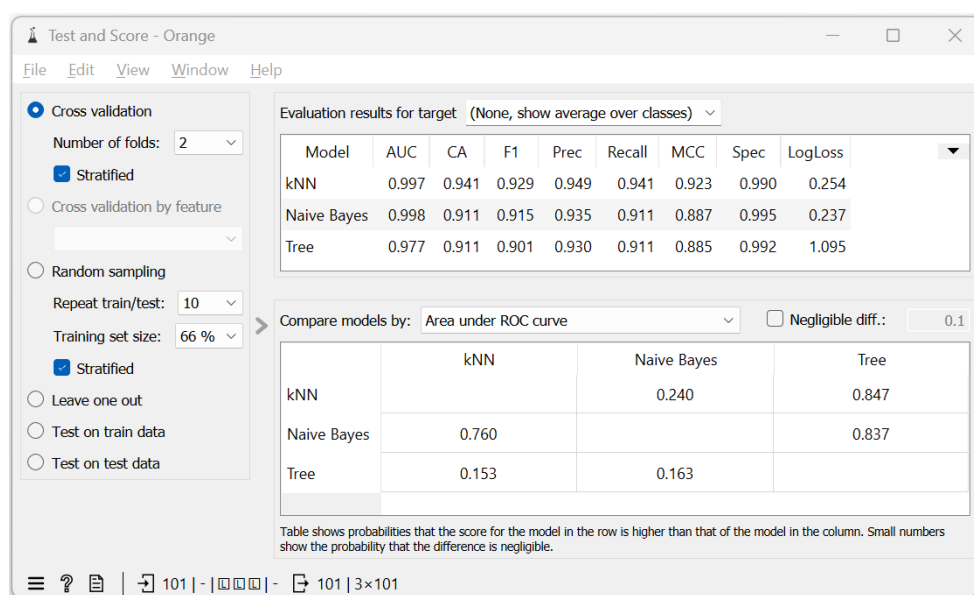


Gambar 2. Design Model Klasifikasi Hewan

Pada Gambar 2 adalah rancangan model klasifikasi hewan yang diterapkan ke dalam *tools orange*, digunakan *widget K-Nearest Neighbor*, *Decision Tree*, *Naive Bayes*, *test and score*, *confussion matrix* dan lainnya. Begitu juga *file* yang digunakan adalah dataset *zoo.tab*. Kemudian dataset tersebut diproses dengan model-model klasifikasi secara otomatis dan hasil sudah dapat dilihat secara langsung.

3.3. Hasil Simulasi

Hasil simulasi dari 3 metode klasifikasi yaitu *Decision Tree*, *K-Nearest Neighbor*, dan *Naive Bayes* dilakukan pada dataset *zoo* diperoleh hasil yang ditampilkan oleh *widget test and score* seperti pada gambar dibawah.



Gambar 3. Hasil Pengujian Model

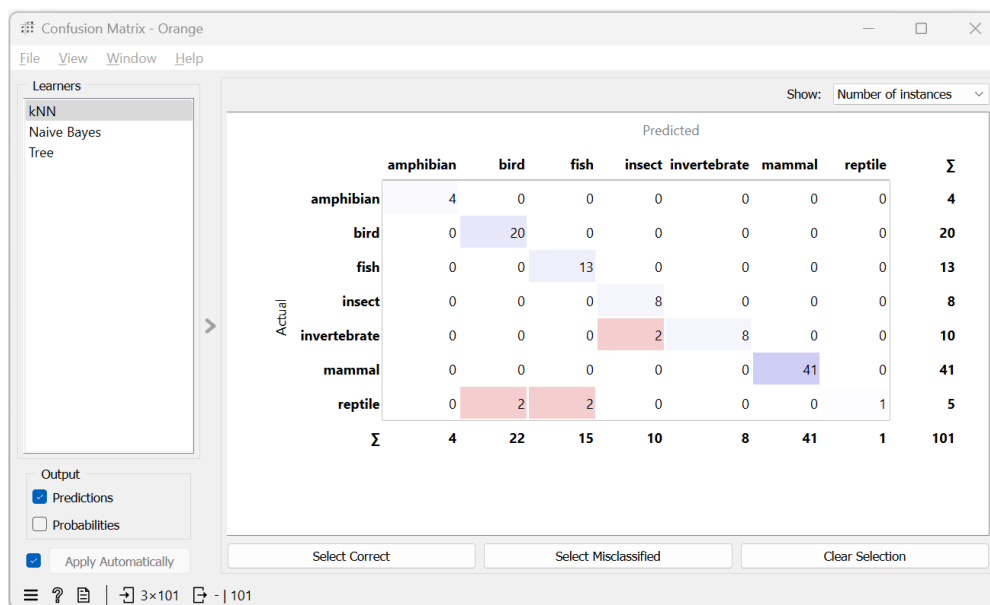
Pada Gambar 3, dari 101 data hewan yang telah diuji, didapat hasil *accuracy*, *precision*, *recall*, dan *f1-score* dari masing-masing metode. Hasil klasifikasi metode *Decision Tree*, *K-Nearest Neighbor* dan *Naive Bayes* menunjukkan bahwa nilai akurasi *K-Nearest Neighbor* adalah yang tertinggi yaitu 0,941 atau 94,1%. Berarti metode ini memiliki ketepatan yang tinggi antara nilai perkiraan dengan nilai sebenarnya. Nilai *precision* juga yang tertinggi sebesar 0,949 atau 94,9% menandakan bahwa metode ini dapat memberikan jawaban dengan tepat. Lalu nilai *recall* juga merupakan yang tertinggi berarti keberhasilan model dalam mengambil informasi dengan akurat sebesar 0,941 atau 94,1%. Sedangkan *f1-score* sebesar 0,929 atau 92,9% yang mengartikan bahwa model memiliki *precision* dan *recall* yang baik karena nilai yang terbaik adalah yang mendekati 100%.

Berdasarkan Gambar 3 juga memperlihatkan hasil AUC dari perbandingan ketiga metode, didapat nilai AUC yang paling tinggi ada pada metode *K-Nearest Neighbor* dan *Naïve Bayes* yaitu 0,997 dengan 0,998. AUC digunakan untuk mengukur kinerja diskriminasi dengan memperkirakan kemungkinan keluaran ilustrasi yang dipilih acak dari kumpulan positive atau negative, semakin tinggi AUC, semakin baik hasil klasifikasi.

Dari Gambar 3 juga menampilkan hasil *logloss*, yang mana nilai terendah ada pada model *Naïve Bayes* dan *K-Nearest Neighbor* dengan nilai sebesar 0,237 dan 0,254. Nilai yang mendekati nol menunjukkan akurasi yang tinggi, sedangkan yang menjauh dari nol menunjukkan akurasi yang rendah, itu artinya metode *Naïve Bayes* dan *K-Nearest Neighbor* menunjukkan hasil akurasi yang baik atau lebih akurat dibandingkan dengan metode *Decision Tree*.

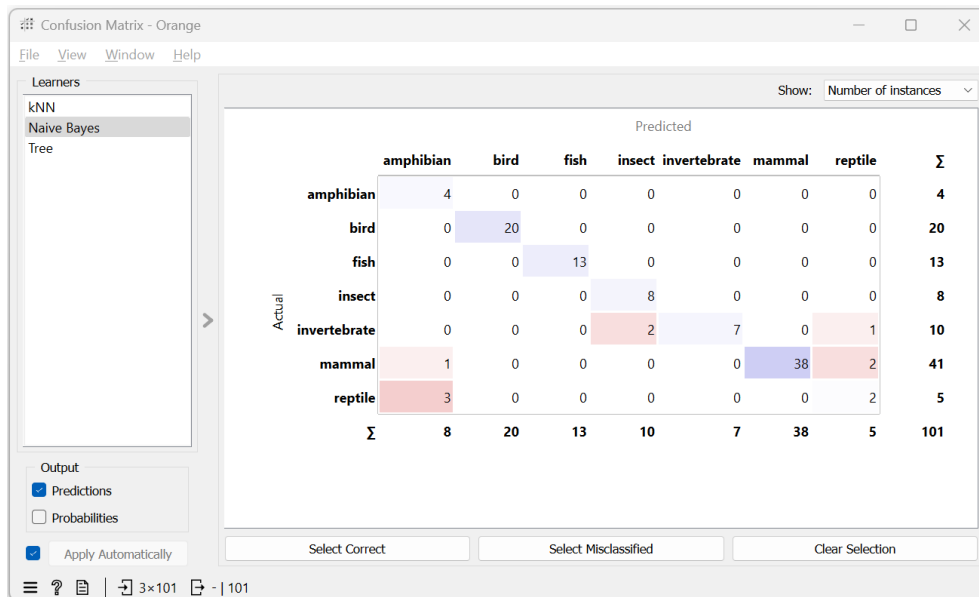
3.4. Confusion Matrix

Ukuran kinerja pada *machine learning* yang hasilnya berupa dua kelas atau lebih. Berupa tabel dengan 4 kombinasi nilai perkiraan dan nilai sebenarnya yang berbeda. Hasil penilaian masing-masing metode klasifikasi dapat dilihat pada tampilan *confusion matrix* dibawah ini.



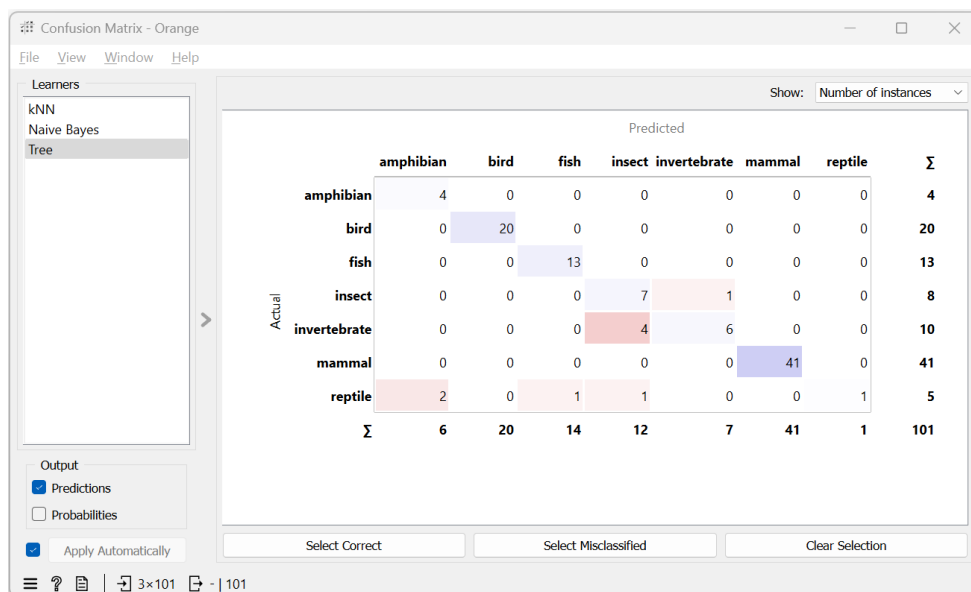
Gambar 4. Confusion Matrix *K-Nearest Neighbor*

Pada Gambar 4. dari 101 jumlah data jenis hewan yang diklasifikasi dengan metode *K-Nearest Neighbor*, terdapat 95 data hewan dengan prediksi sesuai dengan data aktual, sedangkan 6 data tidak sesuai dengan data aktual.



Gambar 5. Confusion Matrix Naïve Bayes

Pada Gambar 5. dari 101 jumlah data jenis hewan yang diklasifikasi dengan metode *Naïve Bayes*, terdapat 92 data hewan dengan prediksi sesuai dengan data aktual, sedangkan 9 data hewan dengan prediksi yang tidak sesuai dengan data aktual.



Gambar 6. Confusion Matrix Decision Tree

Pada Gambar 6. dari 101 jumlah data jenis hewan yang diklasifikasi dengan metode *Decision Tree*, terdapat 92 data hewan dengan prediksi sesuai dengan data aktual, sedangkan 9 data dengan prediksi yang tidak sesuai dengan data aktual.

4. KESIMPULAN

Hasil penelitian dari membandingkan 3 model yaitu, *Decision Tree*, *K-Nearest Neighbor* (KNN), dan *Naïve Bayes* untuk mengklasifikasi jenis hewan menjadi 7 *values* yaitu *mammal*, *fish*, *bird*, *invertebrate*, *insect*, *amphibian*, dan *reptile*, diperoleh hasil bahwa kinerja metode *K-Nearest Neighbor* (KNN) lebih unggul dibanding *Decision Tree* dan *Naïve Bayes*. Dari 101 data yang diuji menggunakan model *K-Nearest Neighbor* (KNN) dihasilkan nilai akurasi sebesar 94,1%, nilai presisi sebesar 94,9%, nilai *recall* sebesar 94,1%, dan nilai *f-measure* sebesar 92,9%. Sedangkan hasil dari metode *Naive Bayes* dengan nilai akurasi sebesar 91,1%, nilai presisi sebesar 93,5%, nilai *recall* sebesar 91,1%, dan nilai *f-measure* sebesar 91,5%. Dan hasil dari metode *Decission Tree* didapat nilai akurasi sebesar 91,1%, nilai presisi sebesar 93%, nilai *recall* sebesar 91,1%, dan nilai *f-measure* sebesar 90%.

Dari hasil yang dicapai, klasifikasi jenis hewan menggunakan *tools Orange Data Mining* mendapatkan nilai-nilai akurasi yang tinggi dengan 3 metode algoritma yang digunakan yaitu *Decision Tree*, *K-Nearest Neighbor* (KNN), dan *Naïve Bayes*. Saran pada penelitian selanjutnya untuk melakukan perbandingan antara metode klasifikasi (*classification*) dan pengelompokan (*clustering*) untuk menentukan manfaat yang tepat dari penggunaan kedua metode algoritma tersebut.

5. DAFTAR PUSTAKA

- [1] K. Annisa, B. Serasi Ginting, and M. A. Syari, "PENERAPAN DATA MINING PENGELOMPOKAN DATA PENGGUNA AIR BERSIH BERDASARKAN KELUHANNYA MENGGUNAKAN METODE CLUSTERING PADA PDAM LANGKAT," *Jurnal Sistem Informasi Kaputama (JSIK)*, vol. 6, no. 2, 2022, [Online]. Available: www.kaputama.ac.id
- [2] Sri Diantika, Hiya Nalatissifa, Riki Supriyadi, Nurlaelatul Maulidah, and Ahmad Fauzi, "IMPLEMENTATION OF MULTI-CLASS GRADIENT BOOSTING TO CLASSIFY ANIMAL SPECIES IN ZOOS," *Antivirus : Jurnal Ilmiah Teknik Informatika*, vol. 17, no. 1, pp. 33–40, Jun. 2023, doi: 10.35457/antivirus.v17i1.2812.
- [3] N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "IMPLEMENTASI ORANGE DATA MINING UNTUK PREDIKSI HARGA BITCOIN," *JURNAL*

- RESPONSIF*, vol. 4, no. 2, pp. 118–125, 2022, [Online]. Available: <https://investing.com/crypto/bitcoin/historical->
- [4] S. Hartono, H. Sujaini, and A. Perwitasari, “Komparasi Algoritma Nonparametrik untuk Klasifikasi Citra Wajah Berdasarkan Suku di Indonesia,” *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 6, no. 3, 2020.
- [5] W. Irmayani, “VISUALISASI DATA PADA DATA MINING MENGGUNAKAN METODE KLASIFIKASI NAÏVE BAYES,” *JURNAL KHATULISTIWA INFORMATIKA*, vol. IX, no. 1, pp. 68–72, 2021, [Online]. Available: www.bsi.ac.id
- [6] S. Alim, “IMPLEMENTASI ORANGE DATA MINING UNTUK KLASIFIKASI KELULUSAN MAHASISWA DENGAN MODEL K-NEAREST NEIGHBOR, DECISION TREE SERTA NAIVE BAYES,” 2021.
- [7] H. A. Dwi Fasnuari, H. Yuana, and M. T. Chulkamdi, “PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS,” *Antivirus : Jurnal Ilmiah Teknik Informatika*, vol. 16, no. 2, pp. 133–142, Oct. 2022, doi: 10.35457/antivirus.v16i2.2445.
- [8] A. H. Nasrullah, “IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PRODUK LARIS,” *Jurnal Ilmiah Ilmu Komputer*, vol. 7, no. 2, 2021, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>
- [9] M. Asfi and N. Fitrianiingsih, “Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi,” *InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 5, no. 1, 2020, doi: 10.30743/infotekjar.v5i1.2536.
- [10] V. Alvian, D. Hidayatullah, A. Nilogiri, H. Azizah, and A. Faruq, “Klasifikasi Siswa Berprestasi Menggunakan Metode K-Nearest Neighbor (KNN) Pada SMA Negeri 2 Situbondo,” 2022. [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST>
- [11] Ainurrohmah, “Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka,” *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 4, pp. 493–499, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>