

Klasifikasi Performa Pokémon Menggunakan Algoritma Naïve Bayes Berdasarkan Statistika Basis dan Tipe Elemen

Adam Ibnu Widiyansyah¹, Eko Hari Rachmawanto²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Kota Semarang

Artikel Info

Kata kunci:

Pokémon
Machine Learning
Naïve Bayes
Klasifikasi
Supervised Learning

ABSTRAK

Game Pokémon memiliki berbagai karakter dengan statistika basis dan tipe elemen yang berbeda-beda sehingga memengaruhi performa dalam pertempuran. Penelitian ini bertujuan untuk mengklasifikasikan performa Pokémon menggunakan algoritma Gaussian Naïve Bayes berdasarkan atribut statistika basis yang terdiri dari HP, Attack, Defense, Special Attack, Special Defense, dan Speed. Penelitian menggunakan pendekatan supervised learning dengan tahapan pengumpulan data, preprocessing, implementasi model, dan evaluasi performa. Dataset diperoleh dari sumber terpercaya seperti Kaggle dan Pokémon Database, kemudian dilakukan normalisasi data menggunakan StandardScaler sebelum proses pelatihan model. Evaluasi dilakukan menggunakan Stratified 5-Fold Cross Validation dan confusion matrix. Hasil penelitian menunjukkan bahwa model Gaussian Naïve Bayes memperoleh akurasi rata-rata sebesar 93% pada proses cross-validation dan 88% pada data uji. Hasil tersebut menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan performa Pokémon dengan cukup baik meskipun masih dipengaruhi oleh ketidakseimbangan data antar kelas.

Penulis Korespondensi :

Adam Ibnu Widiyansyah,
Program Studi Teknik Informatika
Fakultas Ilmu Komputer
Universitas Dian Nuswantoro, Semarang, 50131
Email: adam.ibnuw@gmail.com

1. PENDAHULUAN

Game Pokémon merupakan salah satu permainan populer yang telah berkembang sejak tahun 1996 dan memiliki jutaan pemain di seluruh dunia [1], [2]. Dalam permainan ini, setiap Pokémon memiliki karakteristik yang berbeda-beda, terutama pada atribut statistika basis seperti HP, Attack, Defense, Special Attack, Special Defense, dan Speed [3], [4], [5]. Selain itu, setiap Pokémon juga memiliki tipe elemen tertentu yang memengaruhi kekuatan dan kelemahan saat bertarung [6]. Dalam permainan Pokémon, pemilihan karakter yang tepat menjadi salah satu faktor penting dalam membangun strategi pertempuran [7]. Namun, proses pemilihan Pokémon sering kali dilakukan berdasarkan pengalaman pribadi atau referensi sederhana dari pemain. Dengan semakin banyaknya data Pokémon yang tersedia, diperlukan pendekatan yang lebih sistematis untuk menganalisis performa Pokémon berdasarkan atribut yang dimiliki [8].

Salah satu pendekatan yang dapat digunakan adalah machine learning [9], [10], [11]. Machine learning merupakan cabang kecerdasan buatan yang memungkinkan sistem melakukan proses pembelajaran berdasarkan data untuk menghasilkan prediksi atau klasifikasi tertentu [10], [12]. Dalam penelitian ini, algoritma yang digunakan adalah Naïve Bayes, yaitu algoritma klasifikasi berbasis probabilitas yang dikenal memiliki proses komputasi sederhana namun tetap mampu menghasilkan performa yang baik pada berbagai kasus klasifikasi [13], [14], [15].

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma Naïve Bayes memiliki performa yang baik dalam berbagai kasus klasifikasi. Penelitian yang dilakukan oleh Abdullah [16] pada tahun 2023 mengenai klasifikasi tingkat keparahan penyakit menunjukkan bahwa algoritma Naïve Bayes mampu menghasilkan

tingkat akurasi yang baik dengan efisiensi komputasi yang tinggi. Selanjutnya, penelitian oleh Widodo et al. [17] pada tahun 2025 terkait klasifikasi sentimen ulasan produk e-commerce menyatakan bahwa metode Naïve Bayes mampu menangani data dalam jumlah besar dengan tingkat akurasi yang tinggi, meskipun algoritma tersebut menggunakan asumsi bahwa setiap fitur bersifat independen. Selain itu, penelitian yang dilakukan oleh Padilah et al. [18] pada tahun 2024 mengenai klasifikasi karakter game MOBA menunjukkan bahwa pendekatan berbasis Naïve Bayes sesuai digunakan untuk proses klasifikasi karakter berdasarkan atribut tertentu. Berdasarkan beberapa penelitian tersebut, dapat diketahui bahwa algoritma Naïve Bayes memiliki kemampuan yang cukup baik dalam menangani permasalahan klasifikasi pada berbagai bidang, sehingga metode ini dinilai sesuai untuk diterapkan dalam penelitian klasifikasi performa Pokémon berdasarkan statistika basis dan tipe elemen.

Penelitian ini bertujuan untuk membangun model klasifikasi performa Pokémon menggunakan algoritma Gaussian Naive Bayes serta mengevaluasi tingkat akurasi model dalam memprediksi kategori performa Pokémon. Selain itu, penelitian ini juga bertujuan untuk mengidentifikasi atribut yang paling berpengaruh terhadap hasil klasifikasi.

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode supervised learning menggunakan algoritma Gaussian Naive Bayes. Tahapan penelitian terdiri dari pengumpulan data, preprocessing data, implementasi model, serta evaluasi performa model.

2.1. Pengumpulan Data

Tahap pengumpulan data merupakan langkah awal yang dilakukan dalam penelitian untuk memperoleh dataset yang akan digunakan pada proses pelatihan dan pengujian model klasifikasi. Dataset Pokémon diperoleh dari beberapa sumber terpercaya, yaitu platform Kaggle dan situs Pokémon Database yang menyediakan informasi lengkap mengenai karakteristik setiap Pokémon. Pemilihan sumber data tersebut dilakukan karena memiliki data yang cukup lengkap, terstruktur, dan sering digunakan dalam berbagai penelitian maupun analisis berbasis machine learning. Dataset yang digunakan dalam penelitian ini memuat berbagai atribut statistika basis Pokémon, antara lain HP (Health Points), Physical Attack, Physical Defense, Special Attack, Special Defense, dan Speed. Atribut-atribut tersebut dipilih karena merupakan komponen utama yang memengaruhi performa Pokémon dalam sistem pertarungan pada game Pokémon. Selain statistika basis, dataset juga memuat informasi mengenai tipe elemen Pokémon, seperti Fire, Water, Grass, Electric, dan tipe lainnya yang menjadi identitas karakter setiap Pokémon. Dalam penelitian ini, atribut statistika basis digunakan sebagai fitur utama dalam proses klasifikasi. Setiap atribut memiliki peran yang berbeda dalam menentukan kemampuan Pokémon. Sebagai contoh, atribut HP menunjukkan daya tahan Pokémon, atribut Attack dan Special Attack menunjukkan kekuatan serangan, sedangkan atribut Defense dan Special Defense menunjukkan kemampuan bertahan terhadap serangan lawan. Sementara itu, atribut Speed digunakan untuk menentukan kecepatan Pokémon dalam melakukan serangan.

Untuk mempermudah proses klasifikasi, performa Pokémon dikategorikan ke dalam tiga kelas, yaitu kuat, sedang, dan lemah. Kategori tersebut ditentukan berdasarkan total nilai statistika basis yang dimiliki oleh masing-masing Pokémon. Nilai total diperoleh dari hasil penjumlahan seluruh atribut statistika basis, sehingga dapat merepresentasikan kemampuan keseluruhan Pokémon secara umum. Pokémon dengan total statistika tinggi dikategorikan sebagai Pokémon kuat, sedangkan Pokémon dengan nilai statistika rendah dikategorikan sebagai Pokémon lemah. Adapun Pokémon dengan nilai statistika menengah dimasukkan ke dalam kategori sedang. Proses pengumpulan data dilakukan dengan memperhatikan konsistensi dan kelengkapan data agar dataset yang digunakan dapat mendukung proses pelatihan model secara optimal. Setelah seluruh data berhasil dikumpulkan, tahap selanjutnya adalah preprocessing data untuk membersihkan dan mempersiapkan dataset sebelum digunakan dalam implementasi algoritma Gaussian Naïve Bayes.

2.2. Preprocessing Data

Tahap preprocessing data dilakukan untuk meningkatkan kualitas dataset sebelum digunakan dalam proses pelatihan model klasifikasi. Tahapan ini memiliki peran penting dalam machine learning karena kualitas data yang baik akan sangat memengaruhi performa model yang dihasilkan. Preprocessing bertujuan untuk memastikan bahwa data berada dalam kondisi bersih, konsisten, dan siap digunakan pada proses pelatihan maupun pengujian algoritma Gaussian Naïve Bayes. Pada tahap awal preprocessing, dilakukan pemeriksaan terhadap dataset untuk mengidentifikasi adanya data duplikat dan nilai kosong (missing value). Data duplikat dihapus agar tidak menyebabkan bias pada proses pelatihan model, sedangkan penanganan nilai kosong dilakukan untuk memastikan bahwa setiap atribut memiliki nilai yang valid sehingga tidak mengganggu proses

perhitungan probabilitas pada algoritma Naïve Bayes. Proses pembersihan data ini penting dilakukan agar model dapat mempelajari pola data secara lebih akurat dan stabil. Setelah proses pembersihan data selesai, dilakukan normalisasi data menggunakan metode StandardScaler. Normalisasi bertujuan untuk menyamakan rentang nilai pada setiap atribut statistika basis Pokémon, seperti HP, Attack, Defense, Special Attack, Special Defense, dan Speed. Tahap ini diperlukan karena setiap atribut memiliki skala nilai yang berbeda-beda. Tanpa normalisasi, atribut dengan nilai yang lebih besar dapat memberikan pengaruh yang lebih dominan terhadap proses klasifikasi. Dengan menggunakan StandardScaler, data diubah sehingga memiliki rata-rata mendekati nol dan standar deviasi sebesar satu, sehingga distribusi data menjadi lebih seimbang.

Selain itu, proses preprocessing juga dilakukan untuk memastikan bahwa distribusi data antar kelas tetap terjaga selama proses pelatihan dan pengujian model. Setelah seluruh proses preprocessing selesai dilakukan, dataset dibagi menjadi dua bagian, yaitu data pelatihan (training data) sebesar 80% dan data pengujian (testing data) sebesar 20%. Pembagian data dilakukan menggunakan metode stratified sampling agar proporsi jumlah data pada setiap kelas tetap seimbang baik pada data pelatihan maupun data pengujian. Penggunaan stratified sampling bertujuan untuk mengurangi kemungkinan terjadinya bias akibat ketidakseimbangan distribusi kelas. Dengan demikian, model dapat dilatih menggunakan representasi data yang lebih merata sehingga hasil evaluasi yang diperoleh menjadi lebih akurat dan representatif terhadap kondisi dataset secara keseluruhan. Setelah tahap preprocessing selesai dilakukan, dataset siap digunakan pada proses implementasi algoritma Gaussian Naïve Bayes.

2.3. Implementasi Algoritma Naïve Bayes

Model klasifikasi dalam penelitian ini dibangun menggunakan algoritma Gaussian Naïve Bayes. Algoritma ini dipilih karena memiliki proses komputasi yang sederhana, efisien, serta mampu memberikan performa klasifikasi yang baik pada data numerik. Gaussian Naïve Bayes bekerja berdasarkan pendekatan probabilistik dengan menerapkan Teorema Bayes dan asumsi bahwa setiap atribut bersifat independen terhadap atribut lainnya. Tahap awal dalam implementasi algoritma adalah menghitung probabilitas prior untuk setiap kelas. Probabilitas prior menunjukkan kemungkinan awal suatu data termasuk ke dalam kelas tertentu sebelum mempertimbangkan atribut lainnya. Probabilitas prior dihitung menggunakan persamaan (1) berikut:

$$P(C) = \frac{\text{Jumlah data dalam Kelas } C}{\text{Total Jumlah data}} \dots \dots (1)$$

Pada persamaan (1) diatas, $P(C)$ menyatakan probabilitas awal dari suatu kelas C . Nilai ini diperoleh berdasarkan proporsi jumlah data pada masing-masing kelas terhadap total keseluruhan data. Selanjutnya dilakukan perhitungan probabilitas likelihood atau $P(X/C)$. Pada penelitian ini, atribut statistika basis Pokémon seperti HP, Physical Attack, Physical Defense, Special Attack, Special Defense, dan Speed merupakan data kontinu, sehingga probabilitas likelihood dihitung menggunakan distribusi Gaussian. Persamaan Gaussian yang digunakan adalah sebagai berikut:

$$P(X_i|C) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right) \dots \dots (2)$$

Pada persamaan tersebut, X_i merupakan nilai atribut ke- i , μ merupakan nilai rata-rata atribut pada kelas tertentu, sedangkan σ^2 menyatakan variansi atribut dalam kelas tersebut. Distribusi Gaussian digunakan untuk memperkirakan probabilitas kemunculan suatu atribut kontinu berdasarkan distribusi normal data. Karena algoritma Naïve Bayes mengasumsikan bahwa setiap atribut bersifat independen, maka probabilitas likelihood keseluruhan diperoleh dari hasil perkalian probabilitas masing-masing atribut. Perhitungan likelihood total dilakukan menggunakan persamaan (3) berikut:

$$P(X | C) = P(HP | C) \cdot P(Phy. Atk|C) \cdot P(Phy. Def |C) \cdot P(Sp. Atk |C) \\ \cdot P(Sp. Def |C) \cdot P(Speed |C) \dots \dots (3)$$

Hasil likelihood tersebut kemudian digunakan untuk menghitung probabilitas posterior atau probabilitas akhir suatu kelas terhadap data yang diuji. Probabilitas posterior dihitung menggunakan persamaan (4) Bayes sebagai berikut:

$$P(C|X) = P(X|C) \cdot P(C) \dots\dots\dots (4)$$

Nilai posterior yang diperoleh menunjukkan tingkat kemungkinan suatu data Pokémon termasuk ke dalam kelas tertentu berdasarkan atribut statistika basis yang dimiliki. Tahap terakhir dalam proses klasifikasi adalah menentukan kelas prediksi. Kelas yang dipilih merupakan kelas dengan nilai probabilitas posterior tertinggi. Penentuan kelas prediksi dilakukan menggunakan persamaan (5) berikut:

$$C^* = \operatorname{argmax}_C P(C|X) \dots\dots\dots (5)$$

Berdasarkan persamaan tersebut, sistem akan memilih kelas dengan nilai probabilitas maksimum sebagai hasil akhir klasifikasi performa Pokémon. Implementasi algoritma dilakukan menggunakan bahasa pemrograman Python dengan bantuan pustaka pandas, NumPy, scikit-learn, matplotlib, dan seaborn untuk proses pengolahan data, pelatihan model, dan visualisasi hasil evaluasi.

2.4. Evaluasi Model

Evaluasi model dilakukan untuk mengukur tingkat performa algoritma Gaussian Naïve Bayes dalam mengklasifikasikan performa Pokémon berdasarkan atribut statistika basis. Proses evaluasi dilakukan menggunakan metode Stratified 5-Fold Cross Validation agar distribusi data pada setiap kelas tetap seimbang selama proses pelatihan dan pengujian model. Selain itu, digunakan beberapa metrik evaluasi untuk mengetahui tingkat akurasi dan kualitas prediksi model secara lebih detail. Metrik pertama yang digunakan adalah accuracy. Accuracy digunakan untuk mengukur persentase jumlah prediksi yang benar dibandingkan dengan keseluruhan data yang diuji. Semakin tinggi nilai accuracy, maka semakin baik kemampuan model dalam melakukan klasifikasi. Perhitungan accuracy dilakukan menggunakan persamaan (6) berikut:

$$\text{Accuracy} = \frac{\text{Jumlah prediksi benar}}{\text{Total jumlah data}} \dots\dots\dots (6)$$

Selain accuracy, penelitian ini juga menggunakan precision dan recall untuk mengevaluasi performa model pada setiap kelas. Precision digunakan untuk mengukur tingkat ketepatan prediksi model terhadap suatu kelas tertentu. Nilai precision menunjukkan seberapa banyak data yang diprediksi positif benar-benar termasuk dalam kelas tersebut. Precision dihitung menggunakan persamaan (7) berikut:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \dots\dots\dots (7)$$

Sementara itu, recall digunakan untuk mengukur kemampuan model dalam mendeteksi seluruh data yang benar-benar termasuk ke dalam suatu kelas tertentu. Nilai recall dihitung menggunakan persamaan (8) berikut:

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \dots\dots\dots (8)$$

Untuk memperoleh keseimbangan antara precision dan recall, digunakan metrik F1-Score. F1-Score merupakan rata-rata harmonis dari precision dan recall, sehingga dapat memberikan gambaran performa model secara lebih menyeluruh, terutama pada kondisi data yang tidak seimbang. Perhitungan F1-Score dilakukan menggunakan persamaan (9) berikut:

$$\text{F1 Score} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \dots\dots\dots (9)$$

Selain menggunakan metrik evaluasi numerik, penelitian ini juga memanfaatkan confusion matrix untuk menganalisis hasil klasifikasi model. Confusion matrix digunakan untuk menunjukkan distribusi prediksi benar dan salah pada setiap kelas, sehingga dapat diketahui jenis kesalahan klasifikasi yang dilakukan oleh model. Melalui confusion matrix, performa model dapat dianalisis secara lebih detail berdasarkan hubungan antara nilai actual class dan predicted class. Implementasi seluruh proses evaluasi dilakukan menggunakan bahasa pemrograman Python dengan bantuan beberapa pustaka pendukung. Pustaka NumPy dan pandas digunakan untuk proses manipulasi dan analisis data, sedangkan scikit-learn digunakan untuk proses pembagian dataset,

normalisasi data, implementasi algoritma Gaussian Naïve Bayes, serta evaluasi model. Selain itu, matplotlib dan seaborn digunakan untuk visualisasi hasil evaluasi, termasuk confusion matrix dan grafik performa model.

3. PEMBAHASAN HASIL

Hasil preprocessing menunjukkan bahwa dataset Pokémon berhasil dibersihkan dan dinormalisasi sebelum digunakan dalam proses pelatihan model. Tahap normalisasi dilakukan menggunakan StandardScaler untuk memastikan bahwa seluruh atribut statistika basis, seperti HP, Attack, Defense, Special Attack, Special Defense, dan Speed, memiliki rentang nilai yang seimbang sehingga tidak memberikan pengaruh dominan terhadap proses klasifikasi. Dengan demikian, model dapat mempelajari pola data secara lebih optimal dan stabil. Setelah tahap preprocessing selesai dilakukan, model Gaussian Naïve Bayes dilatih dan dievaluasi menggunakan metode Stratified 5-Fold Cross Validation. Penggunaan metode ini bertujuan untuk menjaga proporsi distribusi kelas pada setiap proses pelatihan dan pengujian sehingga hasil evaluasi menjadi lebih representatif. Berdasarkan hasil pengujian, model memperoleh performa klasifikasi yang cukup baik sebagaimana ditunjukkan pada Tabel 1.

Tabel 1 Hasil Evaluasi Model

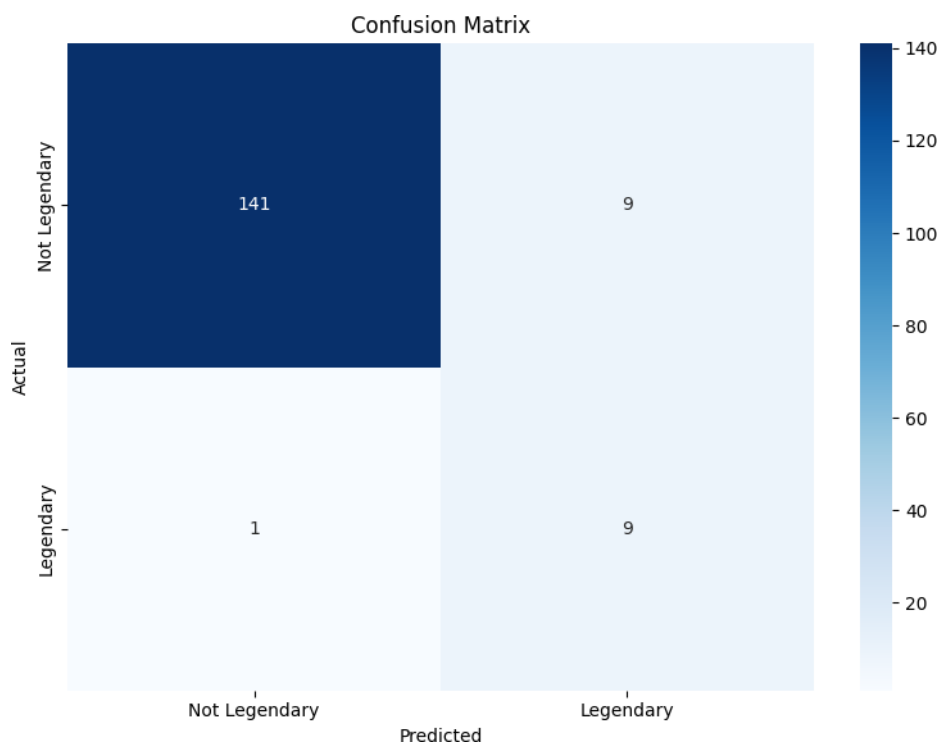
Metrik	Nilai
Akurasi	0.93
Precision	0.57
Recall	0.86
F1-Score	0.68

Berdasarkan Tabel 1, model memperoleh nilai akurasi sebesar 0,93 yang menunjukkan bahwa sebagian besar data Pokémon berhasil diklasifikasikan dengan benar oleh sistem. Nilai recall sebesar 0,86 menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mendeteksi Pokémon pada kelas tertentu, terutama kategori Legendary. Sementara itu, nilai F1-Score sebesar 0,68 menunjukkan bahwa model memiliki keseimbangan performa yang cukup baik antara precision dan recall. Namun, nilai precision yang relatif lebih rendah, yaitu sebesar 0,57, menunjukkan bahwa masih terdapat beberapa kesalahan prediksi ketika model mengidentifikasi Pokémon ke dalam kategori Legendary. Selain pengujian menggunakan cross-validation, model juga diuji menggunakan data testing sebesar 20% dari total dataset. Hasil pengujian menunjukkan bahwa model memperoleh nilai akurasi sebesar 0,88. Nilai tersebut menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik terhadap data baru yang belum pernah dipelajari sebelumnya. Untuk menganalisis hasil klasifikasi secara lebih rinci, penelitian ini menggunakan confusion matrix. Struktur confusion matrix yang digunakan ditunjukkan pada Tabel 2.

Tabel 2 Struktur Confusiom Matrix

	Pred: Not Legendary	Pred: Legendary
Actual: Not Legendary	TN (benar)	FP (salah)
Actual: Legendary	FN (salah)	TP (benar)

Tabel 2 menunjukkan struktur dasar confusion matrix yang digunakan untuk mengevaluasi hasil klasifikasi model. True Negative (TN) menunjukkan jumlah data Not Legendary yang berhasil diprediksi dengan benar, sedangkan True Positive (TP) menunjukkan jumlah data Legendary yang berhasil diklasifikasikan secara tepat. False Positive (FP) menunjukkan kesalahan prediksi ketika Pokémon Not Legendary diprediksi sebagai Legendary, sedangkan False Negative (FN) menunjukkan kesalahan ketika Pokémon Legendary diprediksi sebagai Not Legendary. Hasil implementasi confusion matrix ditunjukkan pada Gambar 1.



Gambar 1 Hasil Confusion Matrix Gaussian Naïve Bayes

Berdasarkan Gambar 1, model berhasil mengklasifikasikan sebanyak 141 data Pokémon Not Legendary dengan benar dan 9 data Pokémon Legendary dengan benar. Akan tetapi, masih terdapat 9 data Pokémon Non-Legendary yang diprediksi sebagai Legendary serta 1 data Legendary yang diprediksi sebagai Non-Legendary. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mendeteksi Pokémon Legendary, namun masih mengalami kesalahan klasifikasi pada beberapa Pokémon dengan atribut statistika basis yang mirip antar kelas. Selain confusion matrix, performa model juga divisualisasikan menggunakan classification report seperti yang ditunjukkan pada Gambar 2.

Classification Report:				
	precision	recall	f1-score	support
False	0.99	0.94	0.97	150
True	0.50	0.90	0.64	10
accuracy			0.94	160
macro avg	0.75	0.92	0.80	160
weighted avg	0.96	0.94	0.95	160

Gambar 2 Classification Report Model Gaussian Naïve Bayes

Berdasarkan Gambar 2, kelas Not Legendary memperoleh nilai precision sebesar 0,99 dan recall sebesar 0,94, yang menunjukkan bahwa model sangat baik dalam mengenali Pokémon Non-Legendary. Sementara itu, kelas Legendary memperoleh nilai precision sebesar 0,50 dan recall sebesar 0,90. Nilai recall yang tinggi menunjukkan bahwa sebagian besar Pokémon Legendary berhasil dideteksi oleh model, namun nilai precision yang lebih rendah menunjukkan bahwa masih terdapat beberapa Pokémon Non-Legendary yang salah diklasifikasikan sebagai Legendary. Faktor utama yang memengaruhi performa model adalah ketidakseimbangan jumlah data antara kategori Pokémon Legendary dan Non-Legendary. Selain itu, beberapa

Pokémon Non-Legendary memiliki nilai statistika basis yang mendekati Pokémon Legendary sehingga menyebabkan kesalahan klasifikasi. Meskipun demikian, secara keseluruhan algoritma Gaussian Naïve Bayes mampu memberikan performa yang cukup baik dalam proses klasifikasi Pokémon berbasis statistika dasar. Algoritma ini juga memiliki keunggulan berupa proses komputasi yang sederhana dan cepat. Namun, metode ini masih memiliki kelemahan karena tidak mempertimbangkan korelasi antar fitur dan cukup sensitif terhadap ketidakseimbangan data.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Gaussian Naive Bayes mampu digunakan untuk mengklasifikasikan performa Pokémon berdasarkan atribut statistika basis berupa HP, Attack, Defense, Special Attack, Special Defense, dan Speed. Model yang dibangun memperoleh akurasi sebesar 93% pada proses cross-validation dan 88% pada data uji, sehingga menunjukkan kemampuan klasifikasi yang cukup baik terhadap data Pokémon. Hasil penelitian juga menunjukkan bahwa ketidakseimbangan data antar kelas menjadi salah satu faktor utama yang memengaruhi hasil klasifikasi. Selain itu, adanya kesamaan nilai statistika basis antar Pokémon menyebabkan beberapa kesalahan prediksi pada kategori tertentu.

Untuk penelitian selanjutnya, disarankan untuk menerapkan teknik penanganan class imbalance seperti SMOTE, oversampling, atau undersampling agar distribusi data menjadi lebih seimbang. Selain itu, penelitian berikutnya dapat membandingkan performa algoritma Naive Bayes dengan metode lain seperti Decision Tree, Random Forest, atau Support Vector Machine untuk memperoleh hasil klasifikasi yang lebih optimal.

REFERENCES

- [1] C. Maitland and E. Agu, "Machine Learning Prediction of Pokemon Go Exergame Enjoyment," in *2025 IEEE Conference on Serious Games and Applications for Health (SeGAH)*, IEEE, Aug. 2025, pp. 1–8. doi: 10.1109/SeGAH65397.2025.11168299.
- [2] R. R. Saxena, E. Nieters, and I. Mamudu, "Pokémondium: A Machine Learning Approach to Detecting Images of Pokémon," Jun. 27, 2025. doi: 10.36227/techrxiv.175099879.95688297/v1.
- [3] E. J. Martínez-López, S. López-Serrano, M. De La Torre-Cruz, and A. Ruiz-Ariza, "Effects of the augmented reality game Pokémon GO on fitness and fatness in secondary school students," *Health Educ. J.*, vol. 81, no. 1, pp. 54–68, Feb. 2022, doi: 10.1177/00178969211047800.
- [4] M. Khamzina, K. V. Parab, R. An, T. Bullard, and D. S. Grigsby-Toussaint, "Impact of Pokémon Go on Physical Activity: A Systematic Review and Meta-Analysis," *Am. J. Prev. Med.*, vol. 58, no. 2, pp. 270–282, Feb. 2020, doi: 10.1016/j.amepre.2019.09.005.
- [5] A. Ruiz-Ariza, R. A. Casuso, S. Suarez-Manzano, and E. J. Martínez-López, "Effect of augmented reality game Pokémon GO on cognitive performance and emotional intelligence in adolescent young," *Comput. Educ.*, vol. 116, pp. 49–63, Jan. 2018, doi: 10.1016/j.compedu.2017.09.002.
- [6] N. Hidayanah, Y. Simarmata, I. W. Permadani, N. K. Destianto, and H. Sulistiani, "Klasifikasi Tipe Pokemon Berdasarkan Statistik Tempur Menggunakan Algoritma Random Forest," *Journal of Computing and Informatics Research*, vol. 4, no. 3, p. 367, 2025, doi: 10.47065/comforch.v4i3.2083.
- [7] E. Jessica, "Analisis User-Experience: Limitasi Pemain Pada Hybrid Reality Games (HRG) Pokemon Go," *JURNAL Dasarrupa: Desain dan Seni Rupa*, vol. 6, no. 3, pp. 32–41, Dec. 2024, doi: 10.52005/dasarrupa.v6i3.335.
- [8] Fauzi, "HUBUNGAN ANTARA PERSEPSI NILAI DAN NIAT BELI PADA KOMUNITAS TRADING CARD GAME POKEMON SOLO," *JURNAL TALENTA PSIKOLOGI*, vol. 13, no. 2, Apr. 2025, doi: 10.47942/talenta.v13i2.2006.
- [9] Y. Yuricha and I. K. Phan, "Penerapan Role Based Access Control dalam Sistem Supply Chain Management Berbasis Cloud," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 339–348, Nov. 2023, doi: 10.57152/malcom.v3i2.1259.
- [10] A. R. Dikananda, O. Nurdian, A. Faqih, A. Surip, and D. B. Saputra, "Estimasi Kemampuan Computational Thinking Pemain Game Dota Dua Menggunakan Teknik Machine Learning," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 5, p. 1599, Oct. 2022, doi: 10.30865/jurikom.v9i5.5025.
- [11] D. S. Sany, F. Zikri, B. E. Putro, M. M. Angkasa, and R. Moch. R. Assyidiq, "Penerapan Game Edukasi Berbasis Machine Learning untuk Meningkatkan Kesadaran Pengelolaan Sampah di SMKS Bina Bangsa Pertiwi Cianjur," *Jurnal Altifani Penelitian dan Pengabdian kepada Masyarakat*, vol. 6, no. 1, pp. 198–209, Dec. 2025, doi: 10.59395/altifani.v6i1.906.
- [12] T. R. Hakim *et al.*, "Getting Started with Machine Learning in Game Development," *Society : Jurnal Pengabdian Masyarakat*, vol. 4, no. 2, pp. 277–276, Mar. 2025, doi: 10.55824/jpm.v4i2.539.

-
- [13] A. T. Bisono and A. Zulherry, “Analisis Sentimen Game Genshin Impact untuk Mengetahui Reaksi dan Harapan Pemain Menggunakan Metode Naïve Bayes,” *sudo Jurnal Teknik Informatika*, vol. 4, no. 2, pp. 183–193, Aug. 2025, doi: 10.56211/sudo.v4i2.1131.
 - [14] V. Fazrian, T. Suprapti, and R. Narasati, “PENERAPAN ALGORITMA NAIVE BAYES TERHADAP ANALISIS SENTIMEN APLIKASI GAME MULTIPLAYER ONLINE BATTLE ARENA,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 1005–1012, Mar. 2024, doi: 10.36040/jati.v8i1.8432.
 - [15] M. Haikal, M. Martanto, and U. Hayati, “ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI GAME ONLINE PUBG MOBILE MENGGUNAKAN ALGORITMA NAIVE BAYES,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 6, pp. 3275–3281, Jan. 2024, doi: 10.36040/jati.v7i6.8174.
 - [16] D. Sebagai *et al.*, “TUGAS AKHIR.”
 - [17] A. Kusuma Wardana *et al.*, “RISET : Jurnal Ilmiah Multidisiplin Ilmu Analisis Sentimen Ulasan Produk E-Commerce Menggunakan Naive Bayes Dan TF-IDF,” 2025.
 - [18] A. Padilah, H. Perdana, and S. A. Intisari, “IMPLEMENTASI METODE NAÏVE BAYES DALAM PREDIKSI TINGKAT KEMENANGAN PADA GAME MOBILE LEGENDS,” 2024.
 - [19] R. Afhalla Hendrawan, A. Fasha, and J. Jestian, “Implementasi Sistem Manajemen Inventaris, Penjualan, dan Penyewaan Sneakers Berbasis Bahasa Pemrograman C++ dengan File I/O,” 2025.
 - [20] N. Nur Azzahrani and D. Rika Susanti, “Perancangan dan Implementasi Program Penjualan Produk Matcha Menggunakan Bahasa Pemrograman C++ pada Matchuchu Griya,” 2025.