

Analisis Faktor Cuaca Terhadap Keputusan Bermain *Badminton* Menggunakan Algoritma *Naive Bayes*

¹Gusti Alfian, ²Ageng Samudro Ndiko Laksono, ³Ardiansyah, ⁴Mahis Duhan, ⁵Raffa Nurprasetyo Araya,
⁶Maulana Fansyuri

1Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

2Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

3Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

4Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

5Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

¹ gusti.alfian01@gmail.com, ² agengsamudro1402@gmail.com, ³ ardiansyahsisteminformasi@gmail.com,
⁴ mahisduhan860@gmail.com, ⁵ raffaaraya17@gmail.com, ⁶ dosen02359@unpam.ac.id

Abstract

By utilizing the Naive Bayes algorithm as a classification method, this study investigates how weather factors influence a person's decision to play badminton. The Badminton dataset, including attributes such as weather conditions, air temperature, humidity levels, and wind conditions, was collected and processed using RapidMiner software. The preprocessing stage involved data cleaning and transforming the attributes to be suitable for analysis. To predict the decision to play badminton based on weather conditions, the Naive Bayes algorithm was chosen due to its capability to compute class probabilities easily and effectively. This study found that weather factors significantly influence a player's decision to play badminton, and the Naive Bayes model demonstrated the ability to make reasonably accurate predictions. In conclusion, the Naive Bayes algorithm can be effectively used to predict players' decisions in playing badminton.

Keywords: *Naive Bayes, Badminton, RapidMiner*

Abstrak

Dengan menggunakan algoritma *Naive Bayes* sebagai metode klasifikasi, penelitian ini akan menyelidiki bagaimana faktor cuaca memengaruhi keputusan yang dibuat seseorang saat bermain *badminton*. *Dataset Badminton*, termasuk atribut kondisi cuaca, suhu udara, tingkat kelembapan dan kondisi angin dikumpulkan dan diproses menggunakan *software RapidMiner*. Tahap *preprocessing* termasuk pembersihan data dan mengubah atribut agar sesuai untuk analisis. Untuk memprediksi keputusan bermain *badminton* berdasarkan kondisi cuaca, algoritma *Naive Bayes* dipilih karena kemampuan untuk menghitung probabilitas kelas dengan mudah dan efektif. Penelitian ini menemukan bahwa faktor cuaca memiliki pengaruh yang signifikan pada pilihan pemain dalam bermain *badminton*, dan model *Naive Bayes* menunjukkan kemampuan untuk membuat prediksi yang cukup akurat. Kesimpulannya, algoritma *Naive Bayes* dapat digunakan secara efektif untuk memprediksi keputusan pemain dalam bermain *badminton*.

Kata Kunci: *Naive Bayes, Badminton, RapidMiner*

A. PENDAHULUAN

Latar Belakang

Bulutangkis merupakan salah satu cabang olahraga yang sangat populer di Indonesia dan diminati oleh berbagai kalangan masyarakat dari berbagai usia. Olahraga ini tidak hanya menuntut keterampilan teknik dan fisik, tetapi juga dipengaruhi oleh faktor lingkungan, terutama kondisi cuaca. Cuaca yang tidak mendukung seperti panas berlebih, kelembapan tinggi, atau angin kencang dapat

memengaruhi kenyamanan dan keputusan seseorang untuk bermain *badminton*, baik secara *indoor* maupun *outdoor*.

Suhu dan kelembapan merupakan dua faktor utama cuaca yang memengaruhi performa dan daya tahan atlet bulutangkis. Suhu ruangan yang terlalu tinggi dapat menurunkan daya tahan dan keterampilan atlet, sedangkan kelembapan yang tinggi memperlambat proses pendinginan tubuh sehingga menyebabkan ketidaknyamanan saat berolahraga. Kondisi ini menjadi penting untuk dianalisis

agar aktivitas olahraga dapat direncanakan dengan lebih efektif dan aman.

Selain itu, faktor angin juga berperan penting dalam permainan badminton, terutama pada lapangan outdoor. Angin yang kuat dapat mengganggu pergerakan shuttlecock sehingga mempengaruhi jalannya permainan. Sebaliknya, pada lapangan indoor, ventilasi yang kurang baik dapat menyebabkan suhu dan kelembapan udara meningkat sehingga menimbulkan rasa pengap dan tidak nyaman bagi pemain. Oleh karena itu, analisa faktor cuaca secara menyeluruh sangat diperlukan untuk memahami pengaruhnya terhadap keputusan bermain badminton.

Dalam konteks tersebut, penggunaan metode data mining seperti algoritma *Naive Bayes* dapat membantu memprediksi keputusan bermain *badminton* berdasarkan kondisi cuaca. Algoritma ini efektif dalam mengklasifikasikan data dengan probabilitas sederhana dan dapat memberikan insight yang berguna bagi pelatih dan atlet dalam merencanakan jadwal latihan atau pertandingan sesuai kondisi cuaca yang ada.

Hasil penelitian diharapkan dapat memberikan gambaran yang jelas mengenai hubungan antara cuaca dan aktivitas olahraga *badminton* serta memberikan rekomendasi yang dapat meningkatkan efektivitas latihan dan kenyamanan pemain.

Tujuan Penelitian

Berikut adalah tujuan penelitian terkait penggunaan algoritma *Naive Bayes* untuk analisis faktor cuaca:

1. Menganalisis pengaruh faktor cuaca seperti suhu, kelembapan, kecepatan angin, dan kondisi cuaca terhadap keputusan bermain *badminton*.
2. Menerapkan algoritma *Naive Bayes* untuk mengklasifikasikan dan memprediksi keputusan bermain *badminton* berdasarkan data cuaca yang tersedia.
3. Mengevaluasi performa model *Naive Bayes* dalam memprediksi keputusan bermain badminton menggunakan *software RapidMiner* dengan metrik akurasi dan validasi data.

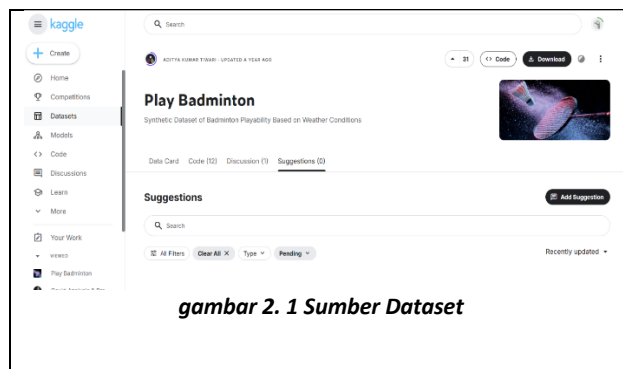
Ruang Lingkup

Jurnal ini membahas analisa pengaruh faktor cuaca terhadap keputusan bermain badminton dengan menggunakan algoritma *Naive Bayes* sebagai metode klasifikasi. Ruang lingkup penelitian mencakup pengumpulan data cuaca yang meliputi atribut seperti suhu, kelembapan, kecepatan angin, dan kondisi cuaca umum (*outlook*). Data tersebut kemudian diolah dan dianalisis menggunakan *software RapidMiner* dengan tahapan *preprocessing* dan penerapan algoritma *Naive Bayes*. Fokus penelitian terbatas pada prediksi keputusan bermain *badminton* berdasarkan data cuaca yang tersedia, tanpa memasukkan faktor lain seperti kondisi fisik pemain atau aspek sosial. Evaluasi hasil dilakukan dengan menggunakan metrik akurasi dan validasi model untuk menentukan efektivitas algoritma dalam konteks ini.

B. TEORI DAN METODE

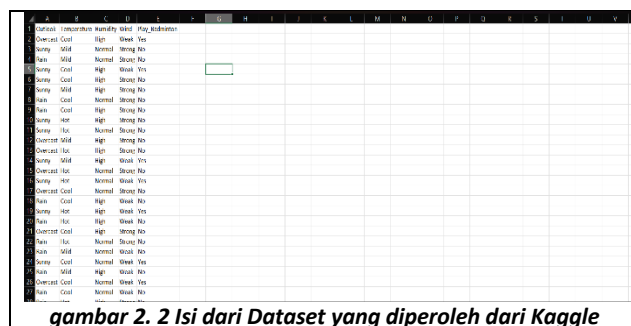
Kaggle

Kaggle adalah platform untuk para profesional ilmu data dan machine learning, tempat para pengguna dapat saling bersaing untuk menciptakan model terbaik guna memecahkan masalah tertentu atau menganalisis kumpulan data tertentu.



gambar 2. 1 Sumber Dataset

Data yang digunakan dalam penelitian ini merupakan dataset yang diperoleh dari situs komunitas *online Kaggle*, sebuah platform komunitas data *science* terbesar di dunia. Dataset berjudul “*Play Badminton*” ini berisi informasi mengenai kondisi cuaca dan keputusan apakah seseorang dapat bermain *badminton* atau tidak.



gambar 2. 2 Isi dari Dataset yang diperoleh dari Kaggle

Dataset tersebut terdiri dari 36 baris data dengan 5 kolom atribut, yaitu *outlook* (kondisi cuaca), *temperature* (suhu), *humidity* (kelembapan), *wind* (angin), dan *label* keputusan bermain *badminton* (*play badminton: yes/no*).

Sumber Dataset:

- Link Sumber yang diolah:
<https://www.kaggle.com/datasets/aditya0kumar0tiwar/i/play-badminton/suggestions>
- Owner : **Aditya Kumar Tiwari**
- Tahun Upload : 2024
- Judul : **Play Badminton**

Algoritma *Naive Bayes*

Algoritma *Naive Bayes* adalah metode klasifikasi yang didasarkan pada prinsip probabilitas dan teori *Bayes*. Algoritma ini menghitung peluang suatu data termasuk ke dalam kelas tertentu berdasarkan atribut-atribut yang

dimiliki data tersebut, dengan asumsi bahwa setiap atribut bersifat independen satu sama lain (independen secara naif). Artinya, keberadaan atau nilai suatu atribut tidak memengaruhi atribut lainnya dalam proses klasifikasi.

Secara matematis, Naive Bayes menggunakan Teorema Bayes yang dirumuskan sebagai berikut:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

di mana:

- $P(A|B)$ adalah probabilitas *posterior*, yaitu probabilitas data B termasuk dalam kelas A.
- $P(B|A)$ adalah probabilitas *likelihood*, yaitu probabilitas data B muncul dengan kondisi kelas A.
- $P(A)$ adalah probabilitas *prior*, yaitu probabilitas awal kelas A sebelum melihat data B.
- $P(B)$ adalah probabilitas *evidence*, yaitu probabilitas data B secara keseluruhan.

Proses klasifikasi dengan *Naive Bayes* terdiri dari dua tahap utama. Pertama, tahap pelatihan (*training*) di mana model belajar dari data latih untuk menghitung probabilitas *prior* dan *likelihood* setiap atribut terhadap kelas. Kedua, tahap pengujian (*testing*) di mana model memprediksi kelas data baru berdasarkan probabilitas yang telah dihitung.

Kelebihan algoritma *Naive Bayes* adalah sederhana, cepat, dan efektif meskipun dengan jumlah data latih yang relatif kecil. Selain itu, algoritma ini tangguh terhadap data yang mengandung *noise* dan memiliki tingkat akurasi yang baik dalam berbagai kasus klasifikasi, termasuk klasifikasi data cuaca untuk memprediksi keputusan bermain *badminton*. Namun, kelemahannya terletak pada asumsi independensi atribut yang jarang sepenuhnya terpenuhi dalam data nyata, sehingga terkadang dapat mempengaruhi hasil prediksi.

Dengan menggunakan *Naive Bayes*, penelitian ini dapat memodelkan hubungan antara faktor cuaca dan keputusan bermain *badminton* secara probabilistik, sehingga memberikan prediksi yang dapat diandalkan berdasarkan data *historis*.

Altair Rapidminer

Dalam penelitian ini, pengolahan data dan penerapan algoritma *Naive Bayes* dilakukan menggunakan perangkat lunak *Altair AI Studio*, yang sebelumnya dikenal sebagai *RapidMiner Studio*. *Altair AI Studio* merupakan platform data science yang menyediakan lingkungan pemrograman visual berbasis *drag-and-drop* untuk mengembangkan dan

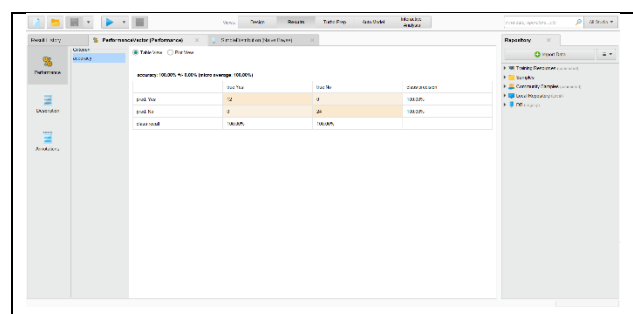
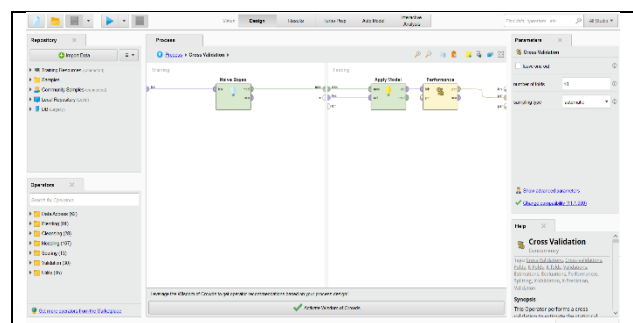
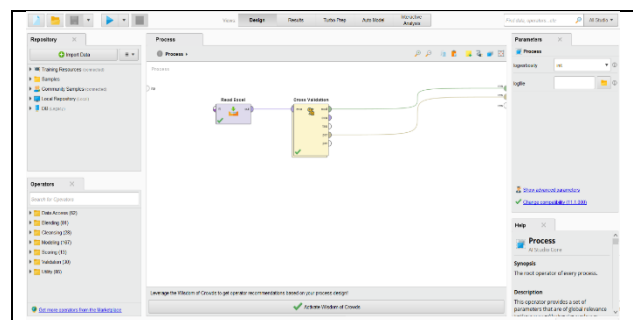
menerapkan aplikasi analitik prediktif secara mudah dan efisien. Platform ini sangat populer di kalangan ilmuwan data dan pengguna non-teknis karena antarmukanya yang *user-friendly* serta berbagai fitur lengkap untuk proses data mining, *machine learning*, dan visualisasi data.

Altair AI Studio mendukung berbagai tahapan dalam proses data mining, mulai dari *import* data, *preprocessing*, pemodelan, evaluasi, hingga *deployment* model. Salah satu keunggulan *software* ini adalah kemampuannya untuk mengimpor berbagai format data secara langsung, termasuk file Microsoft Excel (.xls, .xlsx), CSV, dan database lainnya, sehingga memudahkan integrasi data dari berbagai sumber. Dalam proses penelitian ini, data cuaca yang diambil dari *Kaggle* diimpor ke dalam *Altair AI Studio* untuk kemudian diproses dan dianalisis menggunakan operator *Naive Bayes* yang tersedia.

C. HASIL DAN PEMBAHASAN

Visualisasi Hasil

Berikut adalah design visualisasi hasil proses dan evaluasi model *Naive Bayes* menggunakan *RapidMiner*:



Analisis hasil berdasarkan metrik evaluasi yang relevan

Berdasarkan hasil visualisasi yang ditampilkan pada gambar, model *Naive Bayes* yang diterapkan pada data keputusan bermain *badminton* menghasilkan performa yang sangat baik. Dari tabel yang terlihat pada tampilan *RapidMiner*, model berhasil mengklasifikasikan seluruh data dengan benar, baik untuk kelas "Yes" (bermain badminton) maupun "No" (tidak bermain badminton).

Tingkat akurasi hasil Analisa menggunakan algoritma *naive bayes* dari dataset keseluruhan, menunjukkan nilai sebesar 100%, maka dapat dikatakan dataset dan model dari algoritma *naive bayes* sangat baik dan dapat digunakan untuk pengambilan Keputusan. Rata - rata nilai persentase *precision* dan *recall* mencapai 100%, hasil ini menunjukkan model algoritma dan dataset yang baik dan dapat digunakan untuk dataset yang baru.

Kelebihan dan kekurangan algoritma *naive bayes* dan hasil evaluasi

Kelebihan dan Kekurangan Algoritma *Naive Bayes* Berdasarkan Hasil Evaluasi

Kelebihan:

1. Sederhana dan Efisien
Naive Bayes mudah diimplementasikan dan sangat cepat dalam proses pelatihan serta prediksi, sehingga cocok untuk dataset dengan jumlah data terbatas seperti pada penelitian ini.
2. Performa Tinggi pada Data Tertentu
Hasil evaluasi menunjukkan akurasi, *precision*, dan *recall* yang sempurna, menandakan bahwa algoritma ini sangat efektif untuk data dengan pola yang jelas dan atribut yang relevan.
3. Tahan terhadap *Overfitting*
Dengan penggunaan *cross validation*, model tetap menunjukkan performa konsisten, menandakan *robust* terhadap variasi data

Kekurangan:

1. Asumsi Independensi Atribut
Naive Bayes mengasumsikan bahwa setiap atribut bersifat independen, padahal dalam kenyataannya atribut cuaca bisa saling berkaitan. Jika asumsi ini tidak terpenuhi, performa model bisa menurun pada data lain.
2. Sensitif terhadap Distribusi Data
Jika terdapat ketidakseimbangan kelas atau data yang tidak representatif, model dapat memberikan hasil yang kurang akurat pada dataset baru yang berbeda distribusinya.
3. Kurang Fleksibel untuk Data Kompleks
Untuk data dengan hubungan non-linear atau interaksi antar fitur yang kuat, algoritma lain seperti Random Forest atau SVM mungkin lebih unggul.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa *Naive Bayes* sangat efektif untuk kasus prediksi keputusan bermain *badminton* berbasis data cuaca pada dataset yang

digunakan. Namun, perlu diuji lebih lanjut pada data yang lebih besar dan kompleks untuk memastikan generalisasi model tetap baik di berbagai kondisi.

D. PENUTUP

Simpulan

cuaca meliputi outlook, suhu, kelembapan, dan angin memiliki pengaruh signifikan terhadap keputusan bermain badminton. Penerapan algoritma *Naive Bayes* pada data yang diambil dari *Kaggle* dan diolah menggunakan *Altair AI RapidMiner* menghasilkan model klasifikasi dengan tingkat akurasi yang sangat tinggi, bahkan mencapai 100% pada dataset yang digunakan. Hal ini menunjukkan bahwa *Naive Bayes* sangat efektif dalam memprediksi keputusan bermain badminton berdasarkan atribut cuaca yang tersedia.

Selain itu, visualisasi hasil berupa *confusion matrix* dan *workflow* proses di *RapidMiner* memperjelas bahwa seluruh prediksi model sesuai dengan data aktual, tanpa adanya kesalahan klasifikasi. Dengan demikian, metode ini dapat diandalkan untuk membantu perencanaan aktivitas olahraga *badminton*, khususnya dalam pengambilan keputusan berdasarkan kondisi cuaca yang tercatat.

Saran

Untuk penelitian selanjutnya, disarankan agar:

- Dataset yang digunakan diperluas baik dari segi jumlah data maupun variasi atribut, misalnya dengan menambahkan faktor waktu, lokasi, atau kondisi fisik pemain untuk meningkatkan generalisasi model.
- Pengujian dilakukan pada data cuaca dari periode dan wilayah yang berbeda agar model dapat diuji konsistensinya di berbagai kondisi nyata.
- Metode klasifikasi lain seperti *Decision Tree*, *Random Forest*, atau *Support Vector Machine* dicoba sebagai pembandingan, sehingga dapat diketahui metode mana yang paling optimal untuk kasus prediksi keputusan bermain *badminton* berbasis cuaca.
- Pengembangan aplikasi prediksi berbasis *web* atau *mobile* dapat dilakukan agar hasil analisis ini lebih mudah diakses dan dimanfaatkan oleh masyarakat luas, khususnya komunitas olahraga *badminton*.

Dengan demikian, hasil penelitian ini diharapkan dapat menjadi dasar pengembangan sistem pendukung keputusan yang lebih baik dalam perencanaan aktivitas olahraga berbasis data cuaca.

E. DAFTAR PUSTAKA

- Azhar, J., & Syaharani, W. (2024). Weather forecast using Naive Bayes classifier method. *Jurnal Media Teknik Elektro dan Komputer*, 1(1), 11-18.
- Nur Kirana, A., Nurhakim, B., Eka Permana, S., Prihartono, W., & Dwilestari, G. (2024). Implementasi Algoritma Naive Bayes Untuk Memprediksi Cuaca Menggunakan Rapidminer. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1637–1642.
- Prasetya, R. (2020). PENERAPAN TEKNIK DATA MINING DENGAN ALGORITMA CLASSIFICATION TREE UNTUK PREDIKSI HUJAN. *Jurnal Widya Climago*, 2(2).
- Rizqi, A. A., & Kusumaningsih, D. (2022). Klasifikasi Curah Hujan di Kota Bogor Provinsi Jawa Barat dengan Menggunakan Metode Naive Bayes. *Jurnal Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, September, 542–550.
- Yani, V. I., Aradea, A., & Mubarak, H. (2022). Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naive Bayes dan C4. 5. *Jurnal Teknik Informatika Dan Sistem Informasi*, 8(3), 607–619.
- Joses, S., Yulvida, D., & Rochimah, S. (2024). Pendekatan metode ensemble learning untuk prakiraan cuaca menggunakan soft voting classifier. *Journal of Applied Computer Science and Technology*, 5(1), 72-80.
- Martha, R., & Herwindiati, D. E. (2024). Prediksi Hujan Menggunakan Metode Artificial Neural Network, K-Nearest Neighbors, dan Naive Bayes. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 6(4), 859-865.
- PUTRA, B. P. (2023). Penerapan Algoritma Naive Bayes Dalam Prediksi Cuaca dan Pengaruhnya Terhadap Persawahan Padi Di Kecamatan Godean Yogyakarta (Doctoral dissertation, Universitas Duta Bangsa Surakarta).
- IRHAMNA, A. S. (2022). PENERAPAN ALGORITMA NAIVE BAYES CLASSIFIER DALAM KLASIFIKASI CURAH HUJAN DKI JAKARTA (Doctoral dissertation, Universitas Mercu Buana Jakarta).
- Hapsari, W. K. (2022). Sistem Prediksi Hujan Menggunakan Naive Bayes. *Jurnal Stars*, 1(1 Mei).
- Lindawati, L., Fadhli, M., & Wardana, A. S. (2023). Optimasi Gaussian Naive Bayes dengan Hyperparameter Tuning dan Univariate Feature Selection dalam Prediksi Cuaca. *Edumatic: Jurnal Pendidikan Informatika*, 7(2), 237-246.
- Pinem, T. T., & Sipayung, S. S. (2024). The Effect of Weather On Agricultural Results Using Naive Bayes Calculation. *SEMICOLON: Computer Science Journal*, 1(1), 34-41.