

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Badminton adalah salah satu olahraga yang sangat populer di Indonesia, dengan sejarah panjang dan prestasi cemerlang di tingkat internasional. Sejak tahun 1950-an, Indonesia telah menjadi kekuatan besar dalam badminton dunia, mencetak banyak pemain legendaris seperti Rudy Hartono, Liem Swie King, dan Christian Hadinata, yang telah mengukir banyak gelar juara dunia dan Olimpiade. Bahkan pada era 1990-an, tim badminton Indonesia mengukir prestasi luar biasa dengan memenangkan medali emas di setiap Olimpiade. Oleh karena itu, penting untuk mengeksplorasi cara-cara baru untuk menganalisis dan memprediksi hasil pertandingan badminton, yang dapat memberikan wawasan yang lebih mendalam dalam mendukung pengembangan strategi dan analisis pertandingan.

Dieu [1] telah melakukan penelitian mengenai data subyektif dan empiris dapat mengidentifikasi tingkat keahlian pemain badminton dalam sebuah reli menggunakan algoritma random forest, hasilnya di tunjukan di antara parameter yang dipertimbangkan. Sinadia [2] telah melakukan penelitian menggunakan berbagai macam model seperti Random Forest dan SVM yang membandingkan pemanfaatan metode regresi, klasifikasi, dan clustering dalam melakukan *profiling* atlet badminton. Zhou [3] telah melakukan penelitian menggunakan model *HMM* (*Hidden Markov model*) untuk mengatasi kurangnya perhatian pada pemanasan

dalam latihan badminton, dengan sehingga dapat meningkatkan keterampilan motorik, mencegah cedera, dan juga meningkatkan efektivitas pelatihan.

Penelitian ini bertujuan untuk memperluas pemahaman dan penerapan *machine learning* dalam konteks olahraga, dengan fokus pada cabang olahraga badminton. Penelitian akan menggunakan model *machine learning* untuk memprediksi kemenangan dan memunculkan berapa besar presentase kemenangan suatu pemain dalam sesi pertandingan tunggal putra dalam cabang olahraga badminton menggunakan metode *Supervised Learning*. Model yang akan digunakan dalam penelitian ini adalah *Logistic Regression*, *Naive Bayes*, dan *Random Forest*. Ketiga model akan dibandingkan tingkat akurasi dalam melakukan prediksi pemenang dari suatu pertandingan. Data yang digunakan berasal dari seluruh pertandingan Men Single di *BWF World Tour* dari periode April 2018 hingga 2021.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut:

- 1) Bagaimana proses analitik pertandingan tunggal putra dalam cabang olahraga badminton?
- 2) Bagaimana kinerja model algoritma *Logistic Regression*, *Naive Bayes*, dan *Random Forest*, untuk memprediksi pertandingan tunggal putra dalam cabang olahraga badminton?

- 3) Seberapa akurat dan efektif setiap model prediksi yang telah dibuat untuk memprediksi pemenang pertandingan badminton?

### 1.3 Batasan Masalah

Dalam penelitian ini, diperlukan beberapa batasan yang digunakan sebagai acuan dalam pelaksanaan penelitian untuk memberikan arah yang jelas dalam pengembangannya. Batasan-batasan yang terdapat dalam penelitian adalah sebagai berikut:

- 1) Penelitian ini hanya menggunakan data dalam *BWF World Tour* dari periode April 2018 hingga 2021.
- 2) Model yang akan dibuat hanya untuk memprediksi antara pemain tunggal pria yang terdaftar dalam *BWF World Tour*.
- 3) Model akan dibuat dan dibandingkan hanya menggunakan metode *supervised learning* dengan algoritma *Logistic Regression*, *Naive Bayes*, dan *Random Forest*.

### 1.4 Tujuan Penelitian

Penelitian ini akan mengembangkan model *machine learning* yang dapat memprediksi hasil pertandingan badminton dan menganalisis kinerja atlet badminton tunggal pria dari seluruh dunia berdasarkan data historis pertandingan.

Tujuan penelitian ini meliputi:

- 1) Membangun model *machine learning* yang dapat memprediksi hasil pertandingan badminton.

- 2) Membandingkan akurasi dari tiga model algoritma *Supervised Learning*, yaitu *Logistic Regression*, *Naive Bayes*, dan *Random Forest*, dalam memprediksi hasil pertandingan.

## 1.5 Metodologi

Untuk menyelesaikan masalah pada penelitian ini, metode-metode yang digunakan adalah:

- 1) Studi Literatur

Melakukan pencarian literatur terkait pertandingan badminton, analitik score, dan metode machine learning yang dapat digunakan untuk menganalisis hasil pertandingan dan kinerja atlet.

- 2) Pengumpulan Data

Mengumpulkan data historis pertandingan badminton yang melibatkan atlet Indonesia dari sumber-sumber terpercaya.

- 3) *Cleansing Data*

Membersihkan dan mempersiapkan data untuk analisis, termasuk menghapus nilai yang hilang, mengatasi duplikasi, dan menangani nilai yang tidak valid dan melakukan transformasi data jika diperlukan untuk memenuhi kebutuhan analisis.

- 4) EDA (*Exploratory Data Analysis*)

Menganalisis data secara eksploratif untuk memperoleh pemahaman yang lebih baik tentang karakteristik data, termasuk statistik deskriptif, distribusi, dan korelasi antara variabel. Melakukan visualisasi data menggunakan grafik dan diagram untuk membantu memahami pola

dan tren dalam data.

#### 1) Metode Analitik

- Menggunakan beberapa metode machine learning, seperti regresi linier, regresi logistik, dan decision tree, untuk membangun model prediksi hasil pertandingan badminton dan kinerja atlet.
- Memilih metode machine learning yang paling cocok untuk tujuan penelitian berdasarkan hasil analisis EDA dan literatur yang telah dipelajari.
- Menguji dan mengevaluasi model menggunakan teknik validasi silang untuk memastikan keakuratan dan keterandalan model.

#### 2) Interpretasi Hasil

- Menganalisis hasil prediksi dan mengevaluasi kinerja model untuk menentukan tingkat keakuratan dan relevansi.
- Menginterpretasikan hasil analisis untuk menarik kesimpulan dan membuat rekomendasi model *machine learning* yang telah dibuat.

### 1.6 Sistematika Penulisan

Laporan tugas akhir ini disusun dengan menggunakan sistematika penulisan yang dapat dijelaskan sebagai berikut:

## **BAB I            PENDAHULUAN**

Bab ini dimulai dengan penjelasan mengenai latar belakang penelitian yang berjudul “Perancangan Model *Machine Learning* Menggunakan Algoritma *Logistic Regression*, *Naive Bayes*, dan *Random Forest* dalam Memprediksi Hasil Pertandingan Badminton”. Kemudian dalam bab ini juga dibahas penentuan rumusan dan batasan masalah serta penjelasan tujuan penelitian dan metodologi yang digunakan pada penelitian ini. Pada akhir bab ini dijelaskan mengenai sistematika penulisan yang digunakan.

## **BAB II            LANDASAN TEORI**

Bab ini berisi tentang teori-teori yang digunakan sebagai acuan dalam merancang dan mengembangkan tugas akhir meliputi penjelasan mengenai *Exploratory Data Analysis*, teknik evaluasi dan visualisasi model *machine learning* yang akan dibuat, serta penjelasan mengenai model algoritma *machine learning* yang akan digunakan yaitu *naïve bayes*, *random forest*, dan *logistic regression*,

## **BAB III            METODOLOGI**

Bab ketiga menjelaskan mengenai gambaran perancangan penelitian. Perancangan penelitian akan membahas langkah – langkah metodologi yang diperlukan dalam mengelola data yang digunakan agar dapat digunakan dengan model *machine learning* yang akan dirancang.

## **BAB IV HASIL DAN ANALISIS**

Bab ini berisikan penjelasan mengenai implementasi aplikasi yang dirancang, serta penjelasan singkat mengenai cara penggunaan *graphical user interface* aplikasi. Hasil dari pengujian sistem juga dijelaskan pada bab ini. Dalam pengujian program akan memaparkan analisis yang dilakukan berdasarkan hasil pengujian yang akan digunakan untuk menarik kesimpulan dan saran.

## **BAB V PENUTUP**

Bab lima merupakan bab terakhir yang berisikan kesimpulan yang diperoleh dari penelitian yang telah dilakukan serta saran yang dianjurkan untuk penelitian lebih lanjut.

