

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Machine Learning**

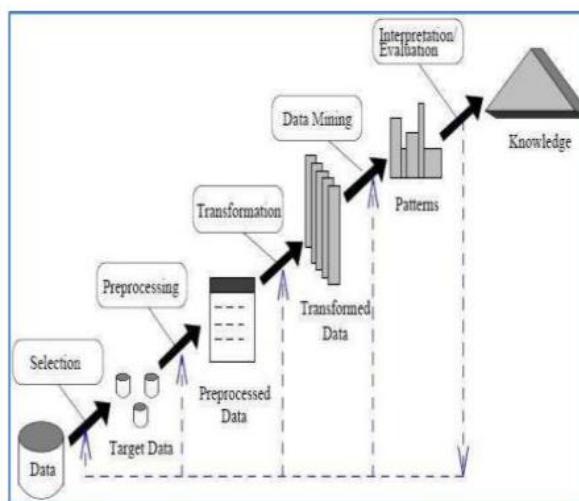
Machine learning adalah cabang dari kecerdasan buatan (AI) yang memungkinkan komputer untuk belajar dari data dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit [9] [10] [11]. Salah satu metode dalam machine learning yang sering digunakan adalah *Artificial Neural Network (ANN)*, yang meniru cara kerja otak manusia untuk mengenali pola dan hubungan dalam data yang kompleks [12] [13]. *ANN* sangat efektif dalam menangani masalah non-linear dan dapat digunakan untuk menganalisis berbagai faktor yang mempengaruhi hasil, seperti kualitas layanan, fasilitas, dan kebersihan dalam konteks penelitian kepuasan Masyarakat [13] [14]. Dalam machine learning, model dapat dilatih menggunakan data historis untuk memprediksi variabel tertentu, seperti tingkat kepuasan masyarakat terhadap program olahraga.

Dalam penelitian ini, machine learning, khususnya metode *ANN*, digunakan untuk menganalisis faktor-faktor yang mempengaruhi kepuasan masyarakat terhadap program olahraga yang diselenggarakan oleh Dinas Kepemudaan dan Olahraga. Dengan memanfaatkan *ANN*, penelitian ini dapat mengidentifikasi pola-pola yang mungkin tidak terlihat dengan analisis tradisional, seperti hubungan kompleks antara kualitas fasilitas, kebersihan, dan pemeliharaan dengan tingkat kepuasan peserta. Melalui pemodelan dan pelatihan data yang diperoleh, *ANN* dapat memberikan prediksi yang lebih akurat tentang bagaimana setiap atribut tersebut

berkontribusi terhadap kepuasan masyarakat, membantu dinas dalam merancang dan meningkatkan program olahraga yang lebih efektif dan sesuai dengan kebutuhan masyarakat.

## 2.2. *Knowledge Discovery in Database*

Knowledge Discovery in Databases (KDD) adalah proses kompleks untuk menemukan pola dan informasi yang berguna dari kumpulan data besar, yang melibatkan beberapa tahapan seperti seleksi data, pembersihan data, transformasi data, penambangan data, dan evaluasi pola yang ditemukan [15] [16]. Proses ini bertujuan untuk mengidentifikasi pola tersembunyi atau hubungan dalam data yang dapat digunakan untuk membuat keputusan yang lebih baik [17] [18] [19]. Dalam konteks penelitian ini, KDD digunakan untuk menganalisis data yang terkait dengan kualitas layanan olahraga, termasuk fasilitas, kebersihan, dan kepuasan masyarakat [20]. Melalui pendekatan KDD, peneliti dapat memperoleh wawasan yang lebih mendalam tentang faktor-faktor yang mempengaruhi kepuasan masyarakat terhadap program olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga [21].



***Gambar 2. 1. Knowledge Discovery in Database (KDD)***

Penelitian ini menggunakan metode KDD untuk memproses data terkait kualitas gedung, fasilitas olahraga, dan kebersihan yang dikumpulkan dari masyarakat yang berpartisipasi dalam program olahraga. Tahapan seleksi dan pembersihan data akan memastikan bahwa data yang digunakan akurat dan relevan, sementara tahap transformasi data akan mengubah data menjadi format yang dapat dianalisis lebih lanjut. Setelah itu, proses penambangan data menggunakan *Artificial Neural Network (ANN)* akan digunakan untuk menggali pola yang ada dalam data, sehingga dapat memprediksi tingkat kepuasan masyarakat berdasarkan berbagai atribut yang diteliti. Dengan demikian, KDD memungkinkan peneliti untuk mendapatkan pengetahuan baru dan terperinci mengenai faktor-faktor yang mempengaruhi kepuasan masyarakat terhadap layanan olahraga.

**2.3. Database dan Data Processing**

Database dalam konteks penelitian ini berperan sebagai tempat untuk menyimpan dan mengelola data yang terkait dengan kualitas fasilitas olahraga, kebersihan, serta tingkat kepuasan masyarakat terhadap layanan yang diberikan oleh Dinas Kepemudaan dan Olahraga. Data yang dikumpulkan, seperti informasi mengenai gedung, fasilitas yang tersedia, kebersihan, serta feedback dari peserta program olahraga, disimpan dalam sistem basis data yang terstruktur. Database ini menjadi sumber utama dalam melakukan analisis menggunakan metode *Artificial Neural Network (ANN)* untuk mengidentifikasi faktor-faktor yang paling berpengaruh terhadap kepuasan masyarakat. Selain itu, database memungkinkan pengelolaan data yang efisien dan akses yang mudah untuk dilakukan pengolahan lebih lanjut.

Data processing adalah tahap penting dalam penelitian ini yang mencakup pengolahan dan transformasi data mentah menjadi informasi yang berguna untuk analisis. Proses ini melibatkan pembersihan data untuk memastikan kualitas dan konsistensi informasi, seperti menghapus data yang hilang atau tidak valid, serta normalisasi data agar dapat digunakan dalam model *ANN*. Data processing juga mencakup seleksi fitur untuk menentukan atribut mana yang paling relevan dalam memprediksi kepuasan masyarakat, seperti kualitas gedung, fasilitas olahraga, dan kebersihan. Dengan pengolahan data yang tepat, model *ANN* dapat menghasilkan prediksi yang lebih akurat dan memberikan wawasan yang lebih mendalam bagi Dinas Kepemudaan dan Olahraga dalam meningkatkan pelayanan dan fasilitas yang ada.

### **2.3.1. Visualization**

Visualization dalam konteks penelitian ini merujuk pada penggunaan teknik grafis untuk menggambarkan data dan hasil analisis dalam bentuk yang mudah dipahami, seperti grafik, diagram, dan peta. Dengan menggunakan visualisasi, peneliti dapat menyajikan hubungan antara variabel-variabel seperti kualitas gedung, fasilitas, kebersihan, dan kepuasan masyarakat secara lebih jelas dan informatif. Visualisasi ini sangat penting untuk menggambarkan tren atau pola dalam data yang mungkin sulit diidentifikasi hanya dengan angka atau tabel. Misalnya, grafik yang menunjukkan hubungan antara kualitas fasilitas dan tingkat kepuasan peserta dapat memberikan wawasan yang lebih mendalam tentang aspek mana yang paling mempengaruhi pengalaman mereka.

Dalam penelitian ini, visualisasi juga dapat digunakan untuk mendemonstrasikan hasil dari analisis menggunakan metode *Artificial Neural Network (ANN)*. Melalui teknik ini, peneliti dapat memetakan data input, seperti kualitas gedung dan fasilitas, ke dalam output berupa tingkat kepuasan masyarakat. Grafik dan diagram yang dihasilkan dari model *ANN* dapat memberikan gambaran yang jelas tentang seberapa besar kontribusi masing-masing faktor terhadap kepuasan masyarakat. Dengan visualisasi yang efektif, hasil penelitian ini akan lebih mudah dipahami dan diterjemahkan ke dalam kebijakan yang dapat digunakan oleh Dinas Kepemudaan dan Olahraga untuk meningkatkan layanan dan fasilitas olahraga mereka.

### **2.3.2. Statistik**

Statistik adalah ilmu yang digunakan untuk mengumpulkan, menganalisis, dan menginterpretasikan data untuk membuat keputusan yang lebih tepat. Dalam konteks penelitian ini, statistik digunakan untuk menganalisis data yang berkaitan dengan kepuasan masyarakat terhadap program olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga. Dengan menggunakan statistik, peneliti dapat mengidentifikasi pola atau hubungan antara faktor-faktor seperti kualitas gedung, fasilitas olahraga, kebersihan, dan pemeliharaan fasilitas dengan tingkat kepuasan masyarakat. Analisis statistik memungkinkan untuk menggambarkan sejauh mana faktor-faktor tersebut memengaruhi persepsi masyarakat terhadap layanan yang diberikan dan memberikan gambaran yang jelas tentang kualitas program olahraga yang ada.

Dalam penelitian ini, metode statistik yang digunakan termasuk teknik *Artificial Neural Network (ANN)* yang dapat mengolah data non-linear dan kompleks untuk memprediksi kepuasan masyarakat berdasarkan atribut-atribut tersebut. Statistik memungkinkan peneliti untuk menyusun model yang dapat memberikan wawasan lebih mendalam tentang faktor-faktor yang memengaruhi kepuasan, serta untuk menguji hipotesis yang telah ditetapkan. Dengan demikian, statistik tidak hanya berfungsi sebagai alat analisis, tetapi juga sebagai dasar untuk mengambil keputusan yang berbasis data guna meningkatkan kualitas layanan dan fasilitas yang disediakan oleh Dinas Kepemudaan dan Olahraga.

### **2.3.3. Pattern Recognition**

Pattern recognition atau pengenalan pola adalah teknik dalam kecerdasan buatan yang bertujuan untuk mengidentifikasi pola atau struktur yang tersembunyi dalam data, baik itu data numerik, citra, atau informasi lainnya. Dalam konteks penelitian ini, pattern recognition digunakan untuk menganalisis data yang diperoleh dari tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga. Melalui pengenalan pola, penelitian ini dapat menemukan hubungan atau pola antara berbagai faktor yang mempengaruhi kepuasan masyarakat, seperti kualitas gedung, fasilitas, kebersihan, dan biaya partisipasi. Pola-pola ini kemudian bisa digunakan untuk memprediksi dan mengevaluasi faktor-faktor yang perlu diperbaiki agar meningkatkan kepuasan masyarakat.

Hubungan antara pattern recognition dan penelitian ini terletak pada penggunaan metode *Artificial Neural Network (ANN)* untuk mengidentifikasi pola

yang ada dalam data. Dengan *ANN*, penelitian ini dapat menganalisis hubungan kompleks antara atribut-atribut yang mempengaruhi kepuasan, seperti kualitas pengajar, fasilitas, dan kebersihan. Pengenalan pola ini memungkinkan untuk mengetahui faktor mana yang paling berpengaruh dalam meningkatkan kepuasan masyarakat terhadap layanan olahraga. Dengan demikian, pattern recognition memainkan peran penting dalam menghasilkan wawasan yang berguna bagi Dinas Kepemudaan dan Olahraga untuk merancang perbaikan program yang lebih efektif dan sesuai dengan harapan masyarakat.

#### **2.4. Metode Artifical Neural Network (*ANN*)**

Metode *Artificial Neural Network (ANN)* adalah salah satu teknik pembelajaran mesin yang meniru cara kerja otak manusia dalam memproses informasi [22] [23] [24]. *ANN* terdiri dari jaringan neuron yang saling terhubung dan dapat dilatih untuk mengenali pola dalam data. Teknik ini sangat berguna dalam menangani masalah dengan data yang kompleks dan non-linear, sehingga dapat menghasilkan prediksi atau klasifikasi yang lebih akurat [25] [26]. *ANN* digunakan untuk menganalisis hubungan antara berbagai faktor yang mempengaruhi suatu hasil, dengan cara mengolah data input dan memberikan output yang diinginkan setelah melalui proses pelatihan [27] [28].

Dalam konteks penelitian ini, *Artificial Neural Network (ANN)* digunakan untuk menganalisis dan memprediksi tingkat kepuasan masyarakat terhadap kualitas layanan olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga. Dengan menggunakan atribut seperti kualitas gedung, fasilitas, kebersihan, dan pemeliharaan, *ANN* mampu mengidentifikasi hubungan yang kompleks antara

faktor-faktor tersebut dan kepuasan masyarakat. Melalui model *ANN*, penelitian ini dapat memberikan gambaran yang lebih akurat tentang bagaimana setiap faktor berkontribusi terhadap kepuasan peserta, sehingga memberikan dasar yang kuat bagi pihak dinas untuk melakukan perbaikan dan peningkatan kualitas layanan olahraga yang lebih efektif.

### **2.7.1. Informasi System**

Sistem informasi yang menggunakan metode *Artificial Neural Network* (*ANN*) memiliki kemampuan untuk memproses data secara kompleks dan menghasilkan prediksi yang akurat berdasarkan pola yang ditemukan dalam dataset. *ANN* bekerja dengan meniru jaringan saraf biologis untuk mengidentifikasi hubungan non-linear antara variabel input dan output, sehingga cocok digunakan untuk analisis tingkat kepuasan masyarakat. Dalam konteks penelitian ini, *ANN* digunakan untuk mengevaluasi data kuesioner yang mencakup atribut seperti kualitas gedung, fasilitas olahraga, kebersihan, dan aksesibilitas program. Dengan menerapkan *ANN*, penelitian ini dapat mengolah data secara efisien untuk memprediksi kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga, serta memberikan wawasan yang dapat digunakan untuk meningkatkan kualitas layanan tersebut.

### **2.7.2. Decision System**

Sistem pengambilan keputusan (Decision System) berbasis *Artificial Neural Network* (*ANN*) merupakan metode yang digunakan untuk menganalisis pola dan hubungan kompleks dalam data. *ANN* bekerja dengan meniru cara kerja otak manusia, menggunakan lapisan neuron untuk memproses informasi dan

menghasilkan prediksi atau keputusan berdasarkan input data. Dalam konteks penelitian ini, *ANN* digunakan untuk mengevaluasi tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga, dengan atribut seperti kualitas gedung, fasilitas olahraga, kebersihan, aksesibilitas program, dan biaya partisipasi. Metode ini memungkinkan penelitian untuk memprediksi kepuasan masyarakat secara lebih akurat, karena *ANN* mampu mengolah data dalam jumlah besar dengan mempertimbangkan interaksi antar variabel yang kompleks. Dengan demikian, hasil analisis dapat memberikan wawasan yang lebih mendalam bagi dinas untuk meningkatkan kualitas layanan dan merancang strategi perbaikan yang efektif.

### **2.7.3. Algoritma System**

Algoritma sistem dalam metode *Artificial Neural Network (ANN)* dirancang untuk meniru cara kerja otak manusia dalam memproses data dan menghasilkan output yang optimal berdasarkan pola-pola tertentu. *ANN* menggunakan struktur jaringan yang terdiri dari lapisan input, lapisan tersembunyi, dan lapisan output, di mana setiap lapisan memiliki neuron yang terhubung dengan bobot tertentu. Proses analisis melibatkan perhitungan feedforward untuk menghasilkan output awal, serta backpropagation untuk memperbaiki bobot berdasarkan kesalahan prediksi. Dalam konteks penelitian ini, *ANN* digunakan untuk menganalisis dan memprediksi tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga. Dengan atribut seperti kualitas gedung, fasilitas, kebersihan, dan aksesibilitas, *ANN* mampu mengidentifikasi pola hubungan antara faktor-faktor

tersebut dan kepuasan masyarakat, sehingga menghasilkan evaluasi yang akurat untuk meningkatkan kualitas layanan di masa depan.

#### 2.7.4. Rumus Metode

Berikut adalah rumus untuk melakukan ataupun menghitung dengan Rumus *Artificial Neural Network (ANN)* dengan menggunakan rumus Feedforward yaitu sebagai berikut.

##### 1. Perhitungan Input Net untuk Neuron

$$z = \sum_{i=1}^n (w_1 \cdot w_2) + b \quad \dots \dots \dots \quad (1)$$

- $z$  : Input net pada neuron.
- $w_1$  : Bobot antara neuron tersembunyi atau output.
- $w_1$  : Input dari neuron sebelumnya.
- $b$  : Bias

##### 2. Fungsi Aktivasi Sigmoid

Fungsi ini berfungsi ataupun akan digunakan jika nilai yang dihasilkan oleh  $X_1, X_2$  dan  $X_3$  lebih dari 1. Maka akan dilakukan fungsi Aktivasi Sigmoid.

$$a = f(z) = \frac{1}{1+e^{-z}} \quad \dots \dots \dots \quad (2)$$

$a$  : Output neuron setelah fungsi aktivasi diterapkan.

Untuk tahapan yang akan dilakukan yaitu sebagai berikut.

##### 1. Menghitung Nilai Bobot (W) dan Bias (B)

- Tetapkan bobot ( $w_1, w_2, \dots$ ) dan Bias ( $b_1, b_2, \dots$ ) dengan nilai tertentu (misalnya 0,5 atau 0,1)

##### 2. Menghitung Nilai Input ( $x_1, x_2$ , atau $z$ )

- Gunakan rumus perhitungan input net:

$$z = (w_1 \cdot x_1) + (w_2 \cdot x_2) + b \quad \dots \dots \dots \quad (1)$$

- Masukkan nilai  $x_1$  dan  $x_2$  dari data.

3. Fungsi Aktivasi untuk Nilai  $x_1$  dan  $x_2$

- Jika nilai  $z > 1$ , gunakan fungsi sigmoid:

4. Menghitung Output untuk Neuron Berikutnya ( $x_3$ )

- Output dari neuron sebelumnya menjadi input untuk neuron berikutnya.
- Hitung nilai  $x_3$  dengan:

$$z_3 = f(z_3) = \frac{1}{1+e^{-z}} \quad \dots \dots \dots \quad (2)$$

## 2.5. Metode *Decision Tree*

Metode *Decision Tree* adalah algoritma pembelajaran mesin yang digunakan untuk menghasilkan pohon keputusan (*Decision Tree*) dengan cara membagi dataset berdasarkan atribut yang memberikan informasi terbaik atau paling signifikan [29] [30]. *Decision Tree* bekerja dengan memilih atribut yang memiliki nilai informasi tertinggi untuk membagi data pada setiap langkah, sehingga pohon keputusan yang dihasilkan dapat digunakan untuk melakukan klasifikasi data [31]. Atribut yang dipilih untuk pemisahan didasarkan pada pengukuran entropi, yang mengukur ketidakpastian dalam data, dan informasi gain yang menggambarkan pengurangan ketidakpastian tersebut setelah pemisahan [32] [33]. Metode ini sangat efektif dalam analisis data dengan banyak atribut dan kategori, yang membuatnya relevan dalam penelitian yang berfokus pada kepuasan masyarakat terhadap layanan olahraga [34] [35].

### **2.8.1. Informasi System**

Sistem informasi berbasis metode *Decision Tree* merupakan salah satu pendekatan yang digunakan untuk menganalisis data secara terstruktur, dengan tujuan memetakan hubungan antara berbagai variabel untuk menghasilkan keputusan yang optimal. Metode ini bekerja dengan membangun model berbentuk pohon keputusan, di mana setiap cabang merepresentasikan kondisi atau atribut tertentu, dan setiap daun menunjukkan hasil atau klasifikasi akhir. Dalam konteks penelitian ini, *Decision Tree* dapat dimanfaatkan untuk mengevaluasi pengaruh faktor-faktor seperti kualitas gedung, fasilitas, dan kebersihan terhadap tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga. Dengan menganalisis data kuesioner yang dikumpulkan, *Decision Tree* mampu memberikan pemahaman yang lebih mendalam mengenai atribut mana yang memiliki dampak signifikan terhadap kepuasan masyarakat, sehingga dapat membantu dinas dalam merancang strategi peningkatan layanan yang lebih efektif.

### **2.8.2. Decision System**

Decision Support System (DSS) dengan metode *Decision Tree* adalah alat yang berguna untuk membantu pengambilan keputusan berdasarkan analisis data yang terstruktur. Metode ini bekerja dengan memecah data ke dalam cabang-cabang berdasarkan atribut tertentu, seperti kualitas layanan, fasilitas, kebersihan, atau atribut lainnya yang relevan dengan penelitian. Dalam konteks penelitian ini, metode *Decision Tree* dapat digunakan untuk mengidentifikasi faktor utama yang memengaruhi tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga. Dengan membangun model berbasis

*Decision Tree*, data dari kuesioner dapat diolah untuk memberikan wawasan tentang pengaruh kualitas gedung, kebersihan, atau fasilitas terhadap kepuasan masyarakat, sehingga menghasilkan rekomendasi yang berbasis data untuk meningkatkan kualitas layanan yang diberikan.

### **2.8.3. Algoritma System**

Algoritma *Decision Tree* adalah salah satu metode dalam data mining yang digunakan untuk membuat model prediksi atau klasifikasi berdasarkan serangkaian aturan keputusan dari data input. Algoritma ini membangun pohon keputusan dengan memecah dataset menjadi subset berdasarkan nilai atribut tertentu, menggunakan konsep Entropy dan Gain untuk menentukan atribut terbaik yang menjadi akar pohon. Dalam konteks penelitian ini, *Decision Tree* dapat diterapkan untuk menganalisis faktor-faktor yang memengaruhi kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga. Dengan atribut seperti kualitas gedung, fasilitas olahraga, kebersihan, dan biaya partisipasi, *Decision Tree* dapat membantu dalam mengidentifikasi atribut mana yang paling signifikan dalam memengaruhi tingkat kepuasan. Hasil analisis ini memberikan wawasan yang lebih jelas bagi pengambil keputusan dalam meningkatkan kualitas layanan olahraga yang sesuai dengan kebutuhan masyarakat.

### **2.8.4. Rumus Metode**

Berikut adalah rumus dan tahapan metode *Decision Tree*. Untuk rumus nya yaitu ada Rumus Entropy dan Rumus Gain. Untuk rumus nya yaitu sebagai berikut.

#### **1. Rumus Entropy**

Entropy digunakan untuk mengukur ketidakpastian data.

$$E(S) = - \sum_{i=1}^n p_i \cdot \log_2(p_i) \quad \dots \quad (1)$$

- $E(S)$  : Entropy dari dataset S.
- $p_i$  : Proporsi jumlah data kategori ke-I terhadap total data.
- $n$  : Jumlah kategori dalam dataset.

Tahapan untuk Perhitungan Entropy yaitu sebagai berikut.

1) Hitung Proporsi Tiap Kategori

- Misalkan dataset memiliki kategori positif dan negatif.
- Hitung proporsi ppositif dan pnegatif:

$$p_{positif} = \frac{\text{Jumlah Positif}}{\text{Total Data}}, p_{negatif} = \frac{\text{Jumlah Negatif}}{\text{Total Data}} \quad \dots \dots \dots \quad (2)$$

2) Masukkan ke Rumus Entropy

$$E(S) = - (p_{positif} \cdot \log_2(p_{positif})) - (p_{negatif} \cdot \log_2(p_{negatif})) \quad \dots \dots \dots \quad (3)$$

## 2. Rumus Gain

Gain digunakan untuk mengukur pengurangan ketidakpastian setelah data dibagi berdasarkan suatu atribut.

$$Gain(S, A) = E(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} \cdot E(S_i) \quad \dots \dots \dots \quad (4)$$

- $Gain(S, A)$  : Informasi yang diperoleh dari pemisahan dataset S berdasarkan atribut A.
- $E(S)$  : Entropy dari dataset awal.
- $S_i$  : Subset data untuk nilai ke-I dari atribut A.
- $|S_i|$  : Jumlah data pada subset  $S_i$ .
- $|S|$  : Jumlah total data dalam dataset.
- $E(S_i)$  : Entropy subset  $S_i$ .
- $k$  : Jumlah nilai unik pada atribut A.

Tahapan untuk perhitungan Gain.

1) Hitung Entropy Awal Dataset

- Gunakan rumus entropy untuk dataset awal S.

2) Pisahkan Dataset Berdasarkan Atribut

- Bagi dataset kedalam subset S1, S2, ..., Sk berdasarkan nilai atribut A.

3) Hitung Entropy Setiap Subset

- Untuk setiap subset Si, hitung entropy:

$$E(S) = - \sum_{j=1}^n p_{ij} \cdot \log_2(p_{ij}) \quad \dots \dots \dots \quad (1)$$

- Di mana  $p_{ij}$  adalah proporsi kategori ke-j dalam subset Si.

4) Hitung Gain

- Masukkan nilai entropy awal dan subset ke rumus gain:

$$Gain(S, A) = E(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} \cdot E(S_i) \quad \dots \dots \dots \quad (2)$$

Dalam penelitian ini, metode *Decision Tree* dapat digunakan untuk mengklasifikasikan dan menganalisis faktor-faktor yang mempengaruhi kepuasan masyarakat terhadap layanan olahraga, seperti kualitas gedung, fasilitas, kebersihan, dan pemeliharaan. Dengan menggunakan *Decision Tree*, data yang dikumpulkan dari responden dapat diolah untuk menghasilkan pohon keputusan yang menggambarkan hubungan antara atribut-atribut tersebut dan tingkat kepuasan masyarakat. Hasilnya dapat memberikan wawasan yang lebih jelas tentang faktor mana yang paling berpengaruh terhadap kepuasan masyarakat dan memberikan rekomendasi berbasis data yang dapat digunakan oleh Dinas

Kepemudaan dan Olahraga untuk meningkatkan kualitas layanan dan fasilitas olahraga yang disediakan.

## 2.6. Model Klasifikasi

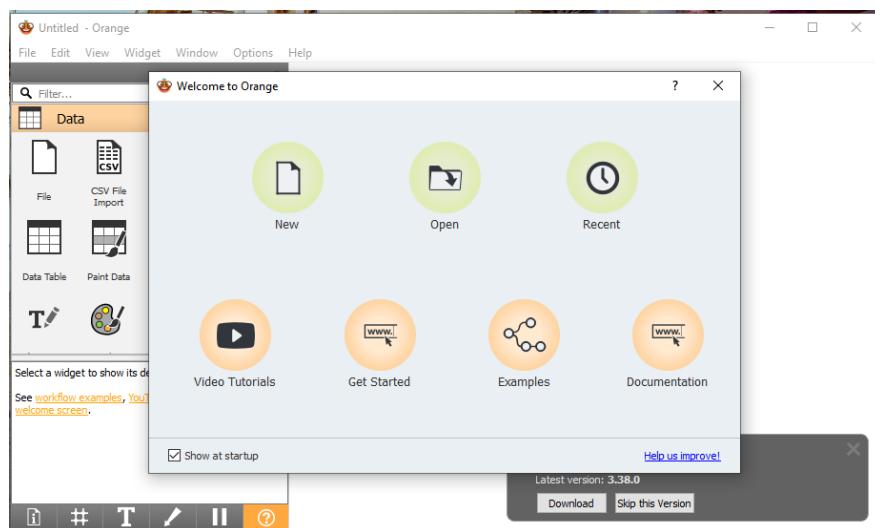
Model klasifikasi adalah teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam kategori-kategori tertentu berdasarkan fitur atau atribut yang ada. Dalam penelitian ini, model klasifikasi dapat digunakan untuk menganalisis tingkat kepuasan masyarakat terhadap layanan olahraga berdasarkan berbagai atribut, seperti kualitas gedung, fasilitas, kebersihan, dan pemeliharaan.

Salah satu metode klasifikasi yang dapat diterapkan adalah *Artificial Neural Network (ANN)*, yang mampu menangani data yang kompleks dan non-linear dengan akurasi tinggi. Dengan menggunakan model klasifikasi, penelitian ini dapat mengidentifikasi faktor-faktor utama yang mempengaruhi kepuasan masyarakat, serta mengelompokkan masyarakat berdasarkan tingkat kepuasan mereka terhadap kualitas layanan yang diberikan.

Pada penelitian ini, model klasifikasi seperti *ANN* memungkinkan peneliti untuk mengolah dan menganalisis data kepuasan masyarakat yang melibatkan variabel-variabel seperti kualitas fasilitas olahraga, kebersihan, dan pemeliharaan. Hasil dari klasifikasi ini dapat memberikan pemahaman yang lebih baik mengenai hubungan antara kualitas fasilitas dan kepuasan masyarakat, serta memberikan rekomendasi yang lebih terarah untuk peningkatan layanan olahraga oleh Dinas Kepemudaan dan Olahraga. Dengan memanfaatkan model klasifikasi, dinas dapat merancang strategi perbaikan yang lebih efektif berdasarkan segmentasi kepuasan masyarakat yang lebih terperinci dan berbasis data.

## 2.7. Tools Pendukung Aplikasi Orange

Aplikasi Orange adalah perangkat lunak open-source yang digunakan untuk analisis data dan pemodelan prediktif, yang sangat berguna dalam penelitian yang melibatkan data kompleks. Aplikasi ini menyediakan berbagai alat visualisasi data dan algoritma pembelajaran mesin, termasuk *Artificial Neural Network (ANN)*, yang dapat digunakan untuk menganalisis faktor-faktor yang mempengaruhi kepuasan masyarakat terhadap layanan olahraga. Dengan Orange, peneliti dapat memproses dan mengolah data mengenai kualitas gedung, fasilitas, kebersihan, dan pemeliharaan fasilitas secara lebih efisien, serta membangun model prediktif yang akurat untuk memprediksi tingkat kepuasan masyarakat berdasarkan atribut-atribut tersebut.



*Gambar 2. 2. Tampilan Awal Aplikasi Orange*

Dalam konteks penelitian ini, Aplikasi Orange dapat menjadi alat yang sangat efektif untuk menganalisis data yang dikumpulkan dari peserta program olahraga dan untuk mengidentifikasi pola-pola yang berkaitan dengan kualitas layanan yang diberikan oleh Dinas Kepemudaan dan Olahraga. Dengan menggunakan metode

*Artificial Neural Network* yang tersedia di Orange, penelitian ini dapat mengembangkan model prediktif yang akan membantu dinas dalam meningkatkan fasilitas olahraga dan layanan yang diberikan. Hal ini akan memungkinkan Dinas Kepemudaan dan Olahraga untuk lebih memahami faktor-faktor kunci yang berpengaruh terhadap kepuasan masyarakat dan membuat keputusan berbasis data untuk peningkatan program olahraga di masa depan.

## **2.8. Metodologi Penelitian**

Untuk metodologi Penelitian yang dilakukan yaitu menggunakan metode *Artificial Neural Network (ANN)* dan metode *Decision Tree*. Untuk metodologi Penelitian yang akan dipaparkan yaitu Penelitian terdahulu dan kerangka kerja penelitian yang dapat dilihat pada penjelasan dibawah ini.

### **2.8.1. Penelitian Terdahulu**

Penelitian terdahulu yang akan dipaparkan pada penelitian ini yaitu akan membahas tentang peneliti-peneliti terdahulu yang berhubungan dengan penelitian ataupun metode yang digunakan pada penelitian ini.

***Tabel 2. 1. Penelitian Terdahulu***

Referensi Penelitian	1
Judul	Analysis of the Neural Network Method to Determine Interest in Buying Pertamax Fuel
Nama	Mayang Sari1)*, Gomal Juni Yanris2), Mila Nirmala Sari Hasibuan3)
Tahun	2023

Hasil	Penelitian ini menggunakan metode <i>Artificial Neural Network (ANN)</i> untuk mengklasifikasikan data masyarakat terkait minat terhadap bahan bakar Pertamax. Hasil klasifikasi menunjukkan bahwa sebanyak 23 orang berminat menggunakan Pertamax, sementara 18 orang tidak berminat, yang menegaskan keandalan metode <i>ANN</i> dalam menganalisis minat masyarakat terhadap bahan bakar [36].
Referensi Penelitian	2
Judul	Comparison of Convolutional Neural Network and <i>Artificial Neural Network</i> for Rice Detection
Nama	Endang Suherman1)* , Djarot Hindarto2) , Amelia Makmur3) , Handri Santoso4)
Tahun	2023
Hasil	Penelitian ini menggunakan metode <i>Artificial Neural Network (ANN)</i> untuk mengklasifikasi dan mendeteksi citra padi pada dataset publik, dengan hasil akurasi pelatihan sebesar 98,2% dan loss sebesar 0,2351 dalam waktu pelatihan sekitar 10 menit. Pada tahap pengujian, <i>ANN</i> menunjukkan akurasi sebesar 96% dengan loss sebesar 0,6641, sehingga membuktikan efektivitas metode ini dalam mendeteksi berbagai jenis padi [37].
Referensi Penelitian	3

Judul	<i>Artificial Neural Network</i> Backpropagation Method to Predict Tuberculosis Cases
Nama	Valencya Lestari1), Herman Mawengkang2) *, Zakarias Situmorang3)
Tahun	2023
Hasil	Penelitian ini menggunakan metode <i>Artificial Neural Network (ANN)</i> dengan algoritma Backpropagation Nguyen-Widrow untuk memprediksi kasus Tuberkulosis berdasarkan data dari Dinas Kesehatan Provinsi Sumatera Utara tahun 2019 hingga 2021. Hasil pengujian menunjukkan bahwa dengan tingkat pembelajaran 0,5 dan momentum 0-1, proses pelatihan berhenti pada iterasi ke-172 dengan nilai gradien 0,0001598, menghasilkan nilai R sebesar 0,99841 yang menunjukkan tingkat akurasi sangat baik sebesar 81,82% [38].
Referensi Penelitian	4
Judul	Improved Accuracy In Data Mining <i>Decision Tree</i> Classification Using Adaptive Boosting (Adaboost)
Nama	Muhammad Riansyah1)*, Saib Suwilo 2 3)
Tahun	2023
Hasil	Penelitian ini membahas penggunaan metode <i>Decision Tree</i> C5.0 untuk klasifikasi data, yang memiliki tingkat

	akurasi 80,58% dengan tingkat kesalahan klasifikasi 19,15% saat diterapkan pada dataset Hepatitis. Dengan penerapan algoritma Adaptive Boosting (Adaboost), akurasi metode <i>Decision Tree</i> meningkat menjadi 82,98%, karena Adaboost mampu mengubah classifier lemah menjadi kuat dengan memberikan bobot lebih besar pada data yang salah klasifikasi, sehingga mengurangi tingkat kesalahan dan meningkatkan kinerja model [39].
Referensi Penelitian	5
Judul	Analysis of the <i>Decision Tree</i> Method for Determining Interest in Prospective Student College
Nama	Safrina Maizura1)*, Volvo Sihombing2), Muhammad Halmi Dar3)
Tahun	2023
Hasil	Penelitian ini menggunakan metode <i>Decision Tree</i> untuk mengklasifikasikan minat calon mahasiswa terhadap pendidikan perguruan tinggi berdasarkan data 65 calon mahasiswa. Hasil klasifikasi menunjukkan bahwa 46 calon mahasiswa tertarik untuk melanjutkan pendidikan, mengindikasikan bahwa metode <i>Decision Tree</i> efektif dalam mengelompokkan data dan memberikan wawasan

	tentang minat calon mahasiswa terhadap perguruan tinggi [40].
--	---

### 2.8.2. Kerangka Kerja Penelitian

Untuk kerangka kerja pada penelitian ini, penulis akan paparkan dalam bentuk tabel seperti dibawah ini.

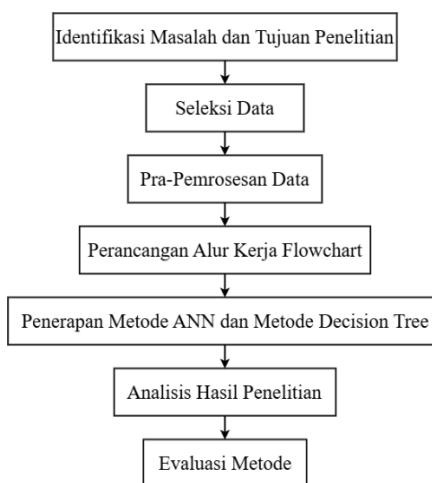
No	Kegiatan	Bulan				
		November 2024	Desember 2024	Januari 2025	Februari 2025	Maret 2025
1	Penentuan Topik Penelitian	✓				
2	Pendefenisian Masalah	✓				
3	Menganalisa Masalah		✓			
4	Menentukan Tujuan		✓			
5	Pengumpulan Data		✓			
6	Preprocessing Data		✓			
7	Konversi Data		✓			
8	Perancangan Algoritma			✓		
9	Pengujian Model				✓	
10	Evaluasi Metode			✓		
11	Pengajuan Seminar Proposal			✓		
12	Seminar Proposal			✓		
14	Pengajuan Sidang Akhir				✓	
14	Sidang Akhir					✓

## **BAB III**

### **ANALISIS DAN PERANCANGAN**

#### **3.1. Arsitektur Sistem**

Arsitektur sistem dalam penelitian ini dirancang untuk mendukung analisis kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga menggunakan metode *Decision Tree*. Sistem ini terdiri dari beberapa komponen utama, yaitu pengumpulan data, pra-pemrosesan data, pemodelan menggunakan metode *Decision Tree*, evaluasi model, dan visualisasi hasil. Data awal dikumpulkan melalui kuesioner yang dibagikan kepada masyarakat yang telah berpartisipasi dalam kegiatan olahraga. Selanjutnya, data yang diperoleh diproses melalui tahap pra-pemrosesan untuk memastikan kualitas data, termasuk penanganan data yang hilang, normalisasi, dan transformasi data agar sesuai dengan format yang dibutuhkan oleh algoritma *Decision Tree*.



Sistem ini juga dilengkapi dengan komponen evaluasi untuk memastikan akurasi dan keandalan model yang dihasilkan. Widget seperti Confusion Matrix di Aplikasi Orange digunakan untuk mengevaluasi kinerja model berdasarkan metrik

akurasi, precision, recall, dan F1-score. Visualisasi hasil analisis ditampilkan dalam bentuk grafik pohon keputusan yang mudah dipahami oleh pengguna, sehingga faktor-faktor yang paling berpengaruh terhadap tingkat kepuasan dapat diidentifikasi dengan jelas. Dengan arsitektur ini, sistem dapat menghasilkan informasi yang akurat dan bermanfaat untuk membantu Dinas Kepemudaan dan Olahraga dalam meningkatkan kualitas layanan serta mengidentifikasi area yang perlu diperbaiki demi meningkatkan kepuasan masyarakat secara keseluruhan.

### 3.1.1. Perbandingan Metode *ANN* dengan Metode *Decision Tree*

Pada tahapan perbandingan merupakan tahapan yang dilakukan untuk membandingkan antara metode *ANN* dengan metode *Decision Tree*. Untuk hasil perbandingannya dapat dilihat pada table dibawah ini yaitu sebagai berikut.

*Tabel 3. 1. Perbandingan antara Metode ANN dan Metode Decision Tree*

No	Perbandingan	Metode <i>Artificial Neural Network</i>	Metode <i>Decision Tree</i>
1	Struktur dan Pendekatan Pemodelan	Sangat baik dalam menangani data yang memiliki pola nonlinier yang kompleks. Cocok untuk masalah prediksi yang sulit dan memerlukan generalisasi tinggi.	Menggunakan struktur seperti pohon untuk membagi data berdasarkan aturan atau keputusan sederhana. Setiap cabang mewakili keputusan, dan setiap daun mewakili hasil.
2	Pendekatan dalam Menangani Data Nonlinear	Sangat baik dalam menangani data yang memiliki pola nonlinier yang kompleks. Cocok untuk masalah prediksi yang sulit dan memerlukan generalisasi tinggi.	Kurang efektif dalam menangani pola nonlinier yang kompleks tanpa teknik tambahan seperti <i>ensemble methods</i> (contohnya <i>Random Forest</i> ).

3	Kebutuhan terhadap Data dan Sumber Daya	Membutuhkan jumlah data yang besar dan sumber daya komputasi tinggi, terutama untuk pelatihan model yang melibatkan banyak paramet	Relatif lebih ringan dan dapat bekerja dengan jumlah data yang lebih kecil tanpa memerlukan sumber daya komputasi yang besar.
4	Interpretabilitas Model	Sulit diinterpretasikan karena berfungsi sebagai kotak hitam ( <i>black box</i> ). Hubungan antar data sulit dilacak oleh manusia.	Mudah diinterpretasikan karena menghasilkan aturan-aturan yang jelas dan dapat dipahami secara visual.
5	Kecepatan Pemrosesan	Proses pelatihan membutuhkan waktu lebih lama, tetapi inferensi (prediksi) biasanya lebih cepat setelah model selesai dilatih.	Pelatihan dan inferensi relatif lebih cepat dibandingkan <i>ANN</i> .
6	Ketahanan terhadap Noise	Relatif lebih tahan terhadap noise karena dapat mengenali pola meskipun ada gangguan pada data.	Rentan terhadap overfitting jika terdapat noise atau data dengan variasi yang terlalu kompleks.
7	Aplikasi	Digunakan untuk masalah yang kompleks seperti pengenalan pola, analisis gambar, pengolahan suara, dan prediksi berbasis data besar.	Sering digunakan untuk pengambilan keputusan sederhana, klasifikasi, dan analisis data tabular.

### 3.1.2. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di Dinas Pemuda dan Olahraga, yang berlokasi di Jl. WR. Supratman No.1, Kompleks GOR Rantauprapat. Lokasi ini dipilih karena merupakan pusat pengelolaan berbagai program olahraga yang bertujuan untuk

meningkatkan kebugaran dan partisipasi masyarakat. Dinas ini memiliki fasilitas olahraga yang cukup lengkap, seperti gedung olahraga tertutup, lapangan terbuka, dan sarana pendukung lainnya, yang menjadi fokus dalam analisis kualitas layanan olahraga. Dengan berbagai program yang dikelola, lokasi ini dianggap representatif untuk mengumpulkan data terkait tingkat kepuasan masyarakat terhadap layanan olahraga.

Penelitian ini dilakukan pada bulan Desember 2024, yang merupakan waktu yang sesuai untuk mengobservasi dan mengumpulkan data mengenai layanan olahraga yang disediakan. Pemilihan waktu ini juga mempertimbangkan kondisi aktivitas masyarakat yang umumnya lebih tinggi pada akhir tahun, baik dalam partisipasi olahraga rutin maupun kegiatan olahraga yang diadakan oleh dinas. Hal ini memungkinkan penelitian mendapatkan data yang lebih komprehensif dan relevan untuk dianalisis menggunakan metode *Artificial Neural Network (ANN)*.

### **3.1.3. Populasi dan Sampel Penelitian**

Populasi dalam penelitian ini adalah seluruh masyarakat yang berdomisili di Kabupaten Labuhanbatu, yang memiliki akses dan kesempatan untuk berpartisipasi dalam kegiatan yang diselenggarakan oleh Dinas Pemuda dan Olahraga. Kabupaten Labuhanbatu dipilih karena memiliki komunitas masyarakat yang beragam, dengan minat yang bervariasi terhadap kegiatan olahraga. Dengan populasi yang mencakup seluruh masyarakat, penelitian ini diharapkan dapat memberikan gambaran umum tentang persepsi dan kepuasan masyarakat terhadap layanan olahraga yang tersedia.

Sampel dalam penelitian ini adalah masyarakat Kabupaten Labuhanbatu yang telah berpartisipasi dalam kegiatan olahraga yang diselenggarakan oleh Dinas

Pemuda dan Olahraga. Pemilihan sampel ini dilakukan untuk memastikan bahwa data yang dikumpulkan berasal dari individu yang memiliki pengalaman langsung terhadap program dan layanan olahraga yang disediakan. Dengan demikian, informasi yang diperoleh akan lebih relevan dan akurat dalam mengevaluasi kualitas layanan, fasilitas, dan kebersihan, serta mengukur tingkat kepuasan masyarakat menggunakan metode *Artificial Neural Network (ANN)*.

#### **3.1.4. Teknik Pengumpulan Data**

Teknik pengumpulan data dalam penelitian ini dilakukan dengan cara membagikan kuesioner kepada masyarakat yang telah berpartisipasi dalam kegiatan yang diselenggarakan oleh Dinas Pemuda dan Olahraga. Kuesioner dirancang untuk menggali informasi mengenai tingkat kepuasan masyarakat terhadap layanan yang diterima, dengan fokus pada kualitas gedung, fasilitas olahraga, kebersihan, serta aksesibilitas program yang disediakan. Pertanyaan dalam kuesioner disusun secara terstruktur menggunakan skala Likert untuk memudahkan analisis data, sehingga hasilnya dapat digunakan untuk mendukung proses evaluasi dan prediksi kepuasan masyarakat.

### **3.2. Desain Aktivitas Sistem**

Pada tahapan pengolahan data, proses pengolahan data akan dilakukan dengan menggunakan 2 metode yaitu Metode *Artificial Neural Network (ANN)* dan Metode *Decision Tree*. Untuk proses pengolahANNya yaitu sebagai berikut.

#### **3.2.1. Langkah-Langkah Metode ANN dan Decision Tree**

Untuk melakukan pengolahan pada metode *ANN* dan Metode *Decision Tree*, untuk tahapnya ada beberapa tahapan yang harus dilakukan yaitu sebagai berikut.

## 1. Identifikasi Masalah dan Tujuan Penelitian

Tahap identifikasi masalah dan tujuan penelitian bertujuan untuk memahami faktor-faktor yang mempengaruhi tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Pemuda dan Olahraga. Fokus utama penelitian ini adalah pada kualitas gedung, fasilitas, kebersihan, serta pemeliharaan sarana olahraga. Selain itu, penelitian ini bertujuan untuk mengevaluasi sejauh mana faktor-faktor tersebut berkontribusi terhadap kepuasan masyarakat dengan menggunakan metode *Artificial Neural Network (ANN)* dan *Decision Tree*. Dengan identifikasi yang tepat, diharapkan penelitian ini dapat memberikan gambaran yang jelas mengenai aspek-aspek yang perlu ditingkatkan untuk meningkatkan kualitas layanan olahraga.

## 2. Seleksi Data

Pada tahapan seleksi data, terdapat jenis data yang digunakan pada metode yang berbeda pula yaitu sebagai berikut.

### 1) Metode Artificial Neural Network

Untuk data yang digunakan yaitu data testing.

*Tabel 3. 2. Data Testing (Sampel) Metode ANN*

Nama	Kualitas Pengajar	Kualitas layanan	Ketersediaan Fasilitas	Aksesibilitas Program	Biaya Partisipasi
Agus Santoso	Bagus	Bagus	Lengkap	Mudah	Murah
Ahmad Nasution	Kurang Bagus	Kurang Bagus	Lengkap	Sulit	Mahal
Andi Nasution	Bagus	Bagus	Kurang Lengkap	Sulit	Murah
Arif Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah
Bayu Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah
Budi Siregar	Kurang Bagus	Kurang Bagus	Lengkap	Sulit	Murah
Citra Dewi	Bagus	Kurang Bagus	Kurang Lengkap	Mudah	Murah
Dedi Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah
Dian Nasution	Bagus	Bagus	Lengkap	Mudah	Murah
Dian Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal
Dwi Siregar	Bagus	Bagus	Lengkap	Mudah	Murah
Eko Pranoto	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal
Heri Siregar	Bagus	Bagus	Lengkap	Sulit	Murah
Intan Maharani	Bagus	Bagus	Lengkap	Mudah	Murah
Joko Prasetyo	Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal
Lestari Wulandari	Kurang Bagus	Bagus	Lengkap	Mudah	Mahal
Lia Kartika	Bagus	Bagus	Lengkap	Mudah	Murah

Mira Nasution	Bagus	Kurang Bagus	Lengkap	Mudah	Murah
Putri Rahmadani	Bagus	Bagus	Lengkap	Mudah	Murah
Ratna Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal
Rina Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal
Rudi Nasution	Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal
Santi Kusuma	Bagus	Bagus	Lengkap	Mudah	Murah
Siti Nurhayati	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal
Sri Rahayu	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Mudah	Mahal
Tommy Siregar	Bagus	Bagus	Lengkap	Mudah	Murah
Tono Sukardi	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal
Tri Wahyuni	Bagus	Kurang Bagus	Lengkap	Sulit	Mahal
Wawan Purnomo	Bagus	Bagus	Lengkap	Sulit	Mahal
Yuni Andriyani	Bagus	Bagus	Kurang Lengkap	Mudah	Murah

Pada table diatas merupakan data sampel yang akan digunakan untuk dihitung dan diolah dengan menggunakan metode *ANN*. Tetapi sebelum data diatas digunakan, data terlebih dahulu dikonvert kedalam format binery. Untuk nilai binery yang akan digunakan yaitu 1 dan 0. Nilai 1 untuk kategori positif dan nilai 0 untuk kategori negatif. Untuk tabel yang sudah dikonvert yaitu sebagai berikut.

**Tabel 3. 3. Data Testing setelah di Convert**

Nama	Kualitas Pengajar	Kualitas layanan	Ketersediaan Fasilitas	Aksesibilitas Program	Biaya Partisipasi
Agus Santoso	1	1	1	1	1
Ahmad Nasution	0	0	1	0	0
Andi Nasution	1	1	0	0	1
Arif Siregar	0	1	1	1	1
Bayu Siregar	0	1	1	1	1
Budi Siregar	0	0	1	0	1
Citra Dewi	1	0	0	1	1
Dedi Siregar	0	1	1	1	1
Dian Nasution	1	1	1	1	1
Dian Siregar	0	1	0	0	0
Dwi Siregar	1	1	1	1	1
Eko Pranoto	0	0	0	0	0
Heri Siregar	1	1	1	0	1
Intan Maharani	1	1	1	1	1
Joko Prasetyo	1	0	0	0	0
Lestari Wulandari	0	1	1	1	0
Lia Kartika	1	1	1	1	1
Mira Nasution	1	0	1	1	1
Putri Rahmadani	1	1	1	1	1
Ratna Siregar	0	1	0	0	0
Rina Siregar	0	1	0	0	0
Rudi Nasution	1	0	0	0	0
Santi Kusuma	1	1	1	1	1
Siti Nurhayati	0	0	0	0	0
Sri Rahayu	0	0	0	1	0
Tommy Siregar	1	1	1	1	1
Tono Sukardi	0	1	0	0	0
Tri Wahyuni	1	0	1	0	0
Wawan Purnomo	1	1	1	0	0

Yuni Andriyani	1	1	0	1	1
----------------	---	---	---	---	---

Pada tabel diatas merupakan data sampel yang sudah diconvert, maka table data sampel diatas sudah dapat digunakan untuk dihitung ataupun diolah dengan menggunakan metode *ANN*.

## 2) Metode *Decision Tree*

Pada tahapan seleksi data merupakan tahapan yang dilakukan untuk menentukan data yang akan digunakan pada penelitian ini. Untuk data yang digunakan yaitu sebagai berikut.

*Tabel 3. 4. Data Testing (Data Sampel) Metode Decision Tree*

Nama	Kualitas Pengajar	Kualitas layanan	Ketersediaan Fasilitas	Aksebilitas Program	Biaya Partisipasi	Kategori
Agus Santoso	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Ahmad Nasution	Kurang Bagus	Kurang Bagus	Lengkap	Sulit	Mahal	Tidak Puas
Andi Nasution	Bagus	Bagus	Kurang Lengkap	Sulit	Murah	Puas
Arif Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Bayu Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Budi Siregar	Kurang Bagus	Kurang Bagus	Lengkap	Sulit	Murah	Tidak Puas
Citra Dewi	Bagus	Kurang Bagus	Kurang Lengkap	Mudah	Murah	Puas
Dedi Siregar	Kurang Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Dian Nasution	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Dian Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Dwi Siregar	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Eko Pranoto	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Heri Siregar	Bagus	Bagus	Lengkap	Sulit	Murah	Puas
Intan Maharani	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Joko Prasetyo	Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Lestari Wulandari	Kurang Bagus	Bagus	Lengkap	Mudah	Mahal	Puas
Lia Kartika	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Mira Nasution	Bagus	Kurang Bagus	Lengkap	Mudah	Murah	Puas
Putri Rahmadani	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Ratna Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Rina Siregar	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Rudi Nasution	Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Santi Kusuma	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Siti Nurhayati	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Sri Rahayu	Kurang Bagus	Kurang Bagus	Kurang Lengkap	Mudah	Mahal	Tidak Puas
Tommy Siregar	Bagus	Bagus	Lengkap	Mudah	Murah	Puas
Tono Sukardi	Kurang Bagus	Bagus	Kurang Lengkap	Sulit	Mahal	Tidak Puas
Tri Wahyuni	Bagus	Kurang Bagus	Lengkap	Sulit	Mahal	Tidak Puas
Wawan Purnomo	Bagus	Bagus	Lengkap	Mudah	Mahal	Puas
Yuni Andriyani	Bagus	Bagus	Kurang Lengkap	Mudah	Murah	Puas

Pada tabel diatas merupakan tabel data sampel yang akan digunakan pada pengolahan data dengan menggunakan metode *Decision Tree*. Data yang digunakan sebanyak 30 data sampel.

### 3. Pra-Pemrosesan Data

Tahapan preprocessing data merupakan tahapan yang digunakan untuk membersihkan data yang akan digunakan pada Penelitian ini. Jadi sebelum data digunakan, data akan diseleksi menyesuaikan dengan kebutuhan dari metode yang digunakan. Untuk preprocessing data pada metode ini yaitu sebagai berikut.

*Tabel 3. 5. Atribut Kualitas Pengajar*

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Kualitas Pengajar	Bagus	17	3	14
	Kurang Bagus	13	9	4

*Tabel 3. 6. Atribut Kualitas layanan*

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Kualitas Layanan	Bagus	20	4	16
	Kurang Bagus	10	8	2

*Tabel 3. 7. Atribut Ketersediaan Fasilitas*

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Ketersediaan Fasilitas	Lengkap	18	3	15
	Kurang Lengkap	12	9	3

*Tabel 3. 8. Atribut Aksebilitas Program*

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Aksebilitas Program	Mudah	17	1	16
	Sulit	13	11	2

*Tabel 3. 9. Biaya Partisipasi*

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Biaya Partisipasi	Murah	17	1	16
	Mahal	13	11	2

*Tabel 3. 10. Atribut Kategori*

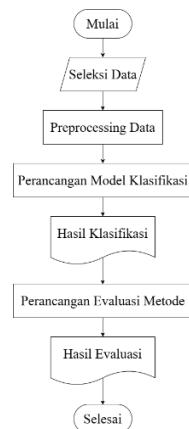
Atribut	Partisi	Kasus (S)
Kategori	Minat	18
	Tidak Minat	12

Tahapan preprocessing data pada metode *Decision Tree* yaitu memisahkan setiap atribut menjadi masing-masing 1 tabel disetiap atribut. Seperti pada table diatas, jadi setiap atribut memiliki table nya masing-masing.

#### 4. Perancangan Alur Kerja Flowchart

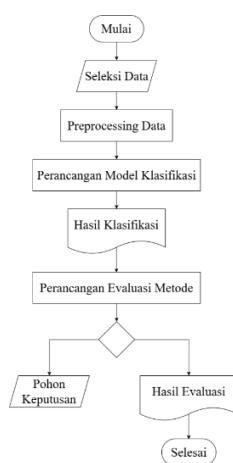
Pada tahapan perancangan alur, penulis menggunakan perancangan alur dengan menggunakan Flowchart agar dapat lebih mudah difahami.

##### 1. Metode Artificial Neural Network (ANN)



**Gambar 2. 3. Flowchart Metode Artificial Neural Network (ANN)**

##### 2. Metode Decision Tree



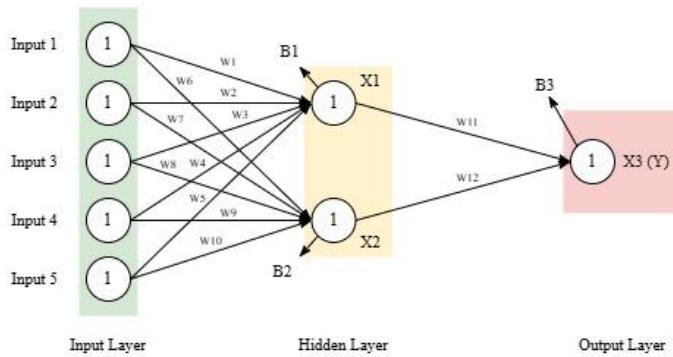
**Gambar 2. 4. Flowchart Metode Decision Tree**

## 5. Penerapan Metode *ANN* dan Metode *Decision Tree*

Pada tahapan penerapan metode yaitu hanya menggunakan rumus manual, hal ini dilakukan untuk dapat lebih memahami konsep metode *ANN* dan metode *Decision Tree* dengan perhitungan matematika.

### 1) Metode *Artificial Neural Network*

Pada proses perhitungan dengan menggunakan metode *ANN*, terlebih dahulu dibuat gambar neuron pada penelitian ini. Untuk gambar neuron nya yaitu sebagai berikut.



*Gambar 3. 1. Neural Model Backpropagation*

Pada gambar diatas merupakan gambar neuron pada perhitungan metode *ANN*, dengan adanya gambar diatas maka perhitungan pada metode *ANN* dapat dilakukan. Untuk gambar diatas terdapat tiga layer yang ditampilkan yaitu input layer yang berisi data yang digunakan, hidden layer berisi tentang data untuk nilai X1 dan X2. Untuk output layer berisi hasil X3. Untuk W merupakan nilai bobot dan untuk B merupakan nilai Bias. Untuk nilai W dan nilai B akan disajikan pada tabel dibawah ini yaitu sebagai berikut.

*Tabel 3. 11. Nilai Bobot dan Nilai Bias*

Bobot	Nilai	Bobot	Nilai	Bias	Nilai
W1	0,2	W7	0,12	B1	1
W2	0,4	W8	0,9	B2	2
W3	0,3	W9	0,8	B3	3

W4	0,6	W10	0,15
W5	0,7	W11	0,10
W6	0,1	W12	0,13

Untuk proses perhitungan ANNya dimulai dengan mencari nilai X1 dan X2, jika X1 dan X2 memiliki hasil lebih dari 1, maka akan dilakukan dengan fungsi aktivasi sigmoid, kemudian dilanjutkan dengan menghitung X3. tetapi jika nilai X1 dan X2 kurang dari 1, maka langsung dilakukan menghitung nilai X3.

## 2) Metode *Decision Tree*

Proses perhitungan metode *Decision Tree* menggunakan Entropy dan Gain diawali dengan mengukur tingkat ketidakteraturan (impurity) dalam dataset melalui rumus Entropy. Entropy digunakan untuk menentukan seberapa homogen atau heterogen suatu kumpulan data berdasarkan atribut target yang diinginkan. Nilai Entropy berkisar antara 0 hingga 1, di mana nilai 0 menunjukkan kondisi data yang sepenuhnya homogen (satu kelas dominan), sedangkan nilai 1 menunjukkan data yang sepenuhnya heterogen (kelas terbagi rata). Setelah Entropy dihitung untuk setiap atribut, langkah selanjutnya adalah menghitung Information Gain untuk setiap atribut. Gain mengukur seberapa besar pengurangan Entropy yang dihasilkan jika data dipisahkan berdasarkan atribut tertentu. Atribut dengan nilai Gain tertinggi akan dipilih sebagai node akar (root node) dalam pohon keputusan, karena atribut tersebut dianggap paling signifikan dalam memisahkan data ke dalam kelas yang berbeda.

Tahapan selanjutnya melibatkan pemisahan dataset berdasarkan atribut dengan Gain tertinggi, dan proses yang sama diulangi untuk setiap subset data hingga tercapai kondisi di mana semua data dalam subset memiliki kelas yang sama

atau memenuhi kriteria penghentian (stopping criteria). Proses ini menghasilkan struktur pohon yang terdiri dari node akar, node cabang, dan node daun (leaf nodes), di mana setiap node daun merepresentasikan hasil akhir dari klasifikasi. Dengan menggunakan metode ini, pola dalam data dapat diidentifikasi dengan jelas, dan setiap cabang pohon merepresentasikan jalur pengambilan keputusan berdasarkan atribut yang dipilih. Melalui penerapan *Decision Tree*, penelitian ini bertujuan untuk menganalisis dan mengklasifikasikan tingkat kepuasan masyarakat terhadap layanan olahraga berdasarkan atribut kualitas gedung, fasilitas, kebersihan, dan pemeliharaan.

## 6. Analisis Hasil Penelitian

Analisis hasil penelitian ini dilakukan menggunakan Aplikasi Orange, sebuah perangkat lunak data mining yang memungkinkan visualisasi data dan penerapan algoritma klasifikasi dengan mudah, termasuk metode *Decision Tree*. Data yang telah dikumpulkan melalui kuesioner diolah dan dianalisis dalam aplikasi ini untuk menghasilkan pola dan hubungan antara variabel penelitian, seperti kualitas gedung, fasilitas, kebersihan, dan pemeliharaan. Proses analisis dimulai dengan tahap data preprocessing untuk memastikan data bersih dan siap digunakan. Selanjutnya, data diolah menggunakan algoritma *Decision Tree* untuk membangun model klasifikasi yang dapat memprediksi tingkat kepuasan masyarakat terhadap layanan olahraga yang disediakan oleh Dinas Kepemudaan dan Olahraga.

Hasil analisis dari aplikasi Orange disajikan dalam bentuk visualisasi pohon keputusan yang mempermudah pemahaman pola dan hubungan antar variabel. Pohon keputusan menunjukkan atribut mana yang paling berpengaruh dalam

menentukan kepuasan masyarakat dan bagaimana atribut tersebut berkontribusi terhadap hasil akhir klasifikasi. Selain itu, metrik evaluasi seperti akurasi, precision, dan recall juga digunakan untuk menilai sejauh mana model yang dihasilkan dapat memprediksi dengan baik. Dengan pendekatan ini, penelitian dapat memberikan gambaran yang lebih jelas mengenai faktor-faktor yang paling memengaruhi kepuasan masyarakat terhadap fasilitas dan layanan olahraga.

## 7. Evaluasi Metode

Evaluasi metode *Decision Tree* dalam penelitian ini dilakukan menggunakan Aplikasi Orange, dengan tambahan widget Confusion Matrix untuk mengukur kinerja model klasifikasi yang dihasilkan. Confusion Matrix berfungsi untuk menunjukkan perbandingan antara prediksi yang dihasilkan oleh model dan nilai aktual dari data yang diuji. Melalui matriks ini, dapat dilihat jumlah prediksi yang benar (True Positive dan True Negative) serta prediksi yang salah (False Positive dan False Negative). Selain itu, metrik evaluasi seperti akurasi, precision, recall, dan F1-score dapat dihitung untuk memberikan gambaran yang lebih komprehensif tentang kinerja algoritma *Decision Tree* dalam memprediksi tingkat kepuasan masyarakat terhadap layanan olahraga. Dengan menggunakan widget ini, hasil evaluasi dapat ditampilkan dalam bentuk visualisasi yang mudah dipahami.

Akurasi yang dihasilkan dari model *Decision Tree* dalam aplikasi Orange mencerminkan sejauh mana model mampu mengklasifikasikan data dengan benar. Nilai akurasi yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam memprediksi tingkat kepuasan berdasarkan atribut yang telah ditentukan, seperti kualitas gedung, fasilitas olahraga, kebersihan, dan

pemeliharaan. Selain akurasi, nilai-nilai lain seperti precision dan recall juga memberikan wawasan tambahan tentang sejauh mana model dapat meminimalkan kesalahan dalam prediksi. Hasil evaluasi ini menjadi dasar untuk menilai keandalan metode *Decision Tree* dalam analisis kepuasan masyarakat dan dapat menjadi referensi untuk penelitian serupa di masa mendatang.