

BAB 2

LANDASAN TEORI

Pada bagian telaah literatur, membutuhkan beberapa penjelasan teori-teori untuk dijadikan sebagai pedoman untuk implementasi dan penerapan pada algoritma untuk rekomendasi jenis musik dalam mendukung kesehatan mental. berikut penjelasan detail dari landasan teori pada penelitian ini.

2.1 Kesehatan Mental

Kesehatan mental merupakan suatu situasi atau keadaan psikologis yang memberikan kemampuan dalam menyesuaikan diri seseorang atau menyelesaikan suatu masalah pada dirinya sendiri maupun masalah yang dipengaruhi oleh faktor lingkungan dari luar dirinya[12]. Kesehatan mental memiliki petunjuk untuk berpikir, berperasaan, dan bertindak sesuai dengan karakter individu masing-masing dengan lebih efektif dan efisien serta dapat lebih berani dalam menjalani kehidupan[12]. Kesehatan mental dapat bekerja dengan baik jika seorang individu memberikan indikasi bahwa fungsi mental pada dirinya bekerja secara optimal[12]. Fungsi-fungsi mental tersebut dapat bekerja sesuai dengan individu seseorang, karena jika dirinya mampu melakukan suatu kegiatan yang bermanfaat bagi lingkungan hidupnya, maka individu tersebut dapat melaksanakan hubungan sosial yang efektif dan efisien dengan individu lainnya serta mampu beradaptasi pada transformasi hidup dan mampu mengendalikan atau meminimalisir kendala dalam kehidupan yang dialaminya[12].

2.2 Jenis Musik

Jenis musik didefinisikan berdasarkan instrumen yang digunakan, teknik musik, pola musik, tema, dan sebagainya.[13] Musik menjadi bagian individu yang dijadikan sebuah media refleksi yang meningkatkan konsentrasi, mengurangi depresi, dan menjaga kesehatan mental[13]. Jika memahami preferensi jenis musik dapat membantu manusia menjadi lebih mendalami apa yang didengarkan dalam dirinya serta membuka wawasan keberagaman musik yang tersedia menjadi bisa lebih dalam dan luas[13]. Ada juga berbagai bantuan teknologi untuk mengetahui saran musik yang diinginkan yaitu sistem rekomendasi musik yang berfungsi sebagai alat dalam membantu memilah data yang sudah diambil menjadi sebuah

data yang relevan seperti selera musik seseorang yang hanya melihat riwayat lagu yang didengarkan dan akan langsung merekomendasikan musik sesuai suasana dan situasi seseorang[14]. Musik juga merangsang pengeluaran dalam gelombang otak yang dikenal sebagai gelombang α yang memiliki frekuensi sekitar 8-12 *cycles per second* yang berfungsi sebagai menjaga mood, membantu pola tidur, perasaan yang tenang dan melepaskan rasa depresi berlebih[15].

2.3 Naive Bayes Classifier

Naive Bayes merupakan salah satu metode pengklasifikasian sederhana karena penerapan yang mudah dan memiliki hasil yang baik. Ketika diterapkan pada suatu penelitian. Adapun kelemahan dari *Naive Bayes* yaitu adanya asumsi atau suatu kondisi kelas yang saling bebas, sehingga dapat menyebabkan ketidakakuratan, dalam prakteknya, kebergantungan ada pada tiap variabel. Misalnya kasus pada rumah sakit: pasien, umur, keluarga, dsb. Kebergantungan diantara variabel tadi tidak dapat dimodelkan pada *Naive Bayes Classifier*[16]. Adapun rumus dari teorema pada persamaan berikut[16]:

$$P(H | X) = \frac{P(X | H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan:

X : Data dengan class yang belum diketahui

H : Hipotesis data X

P(H—X) : Probabilitas hipotesis H

P(H) : Probabilitas hipotesis H (prior probability)

P(X—H) : Probabilitas X berdasarkan kondisi H

P(X) : Probabilitas X

2.4 Feature Selection

Feature Selection merupakan proses melakukan pemilihan atribut data untuk digunakan dalam suatu pembelajaran mesin[17]. Tujuan dari proses ini untuk memilih atribut-atribut yang relevan sehingga lebih efektif bekerja untuk hasil akhir dan membuang atribut-atribut yang tidak memiliki kaitannya dengan data yang diteliti[17].

Ada beberapa metode bisa dilakukan pada *feature selection* tersebut seperti mengurangi dimensi dengan teknik *filter*, *wrapper*, dan *embedded* serta

keunggulan dan kelemahan dari setiap teknik akan memiliki variasi yang sesuai dari penerapannya[17]. Dalam pemilihan metode seleksi fitur yang tepat akan mendapatkan hasil yang optimal dan dapat meningkatkan tingkat akurasi yang lebih baik[17].

2.5 Information Gain

Information Gain merupakan metode dalam memilih atribut dalam dataset yang paling sederhana dengan hanya memberi peringkat pada atribut. Metode *Information Gain* banyak digunakan dalam penerapan kategori analisis data, teks, *microarray*, dan analisis data gambar[9]. Dalam memilih atribut dataset, pendekatan klasifikasi menjadi pemegang peranan yang sangat penting dan metode *Information Gain* dapat membantu mengurangi noise yang disebabkan pada atribut yang tidak relevan[9]. Langkah pertama yang harus dilakukan yaitu menentukan nilai atribut terbaik dengan menghitung nilai entropi. Entropi merupakan suatu proses menggunakan probabilitas tertentu[9]. Setelah melakukan perhitungan pada nilai entropi, selanjutnya dapat dihitung metode *Information Gain*nya dengan kalkulasi entropi yang didefinisikan seperti berikut[9]:

$$Entropy(S) = \sum_i^e -P_i \log_2 P_i \quad (2)$$

Keterangan:

C = angka dari nilai label klasifikasi

Pi = angka dari sampel kelas i

Selanjutnya merupakan metode *Information Gain* yang didefinisikan seperti berikut[7]:

$$Gain(S,A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v), \quad (3)$$

Keterangan:

A = atribut

v = nilai yang memungkinkan dari atribut A

$|S_v|$ = angka dari sampel nilai v

$|S|$ = jumlah dari semua data sampel

Entropy (Sv) = entropi dari sampel yang memiliki nilai v

2.6 Confusion Matrix

Confusion Matrix adalah tabel dalam menggambarkan suatu performa sebuah algoritma secara detail. Setiap baris dari *matrix* tersebut, mewakili kelas aktual dari data, dan setiap kolom mewakili kelas prediksi dari data atau sebaliknya[18].

	Prediksi Negatif	Prediksi Positif
Aktual Negatif	True Negatif (TN)	False Positif (FP)
Aktual Positif	False Negatif (FN)	True Positif (TP)

Keterangan :

- *True* Positif = Seberapa banyak data yang aktual kelasnya positif dan model juga akan memprediksi positif,
- *True* negatif = Seberapa banyak data yang aktual kelasnya negatif dan model juga akan memprediksi negatif.
- *False* Positif = Seberapa banyak data yang aktual kelasnya negatif, tetapi model memprediksi positif.
- *False* Negatif = Seberapa banyak data yang aktual kelasnya positif, tetapi model memprediksi negatif.

Dari keempat data tersebut, terdapat data lain yang akan berguna dalam mengukur suatu performa pada model seperti[15]:

1. *Accuracy* = Total Keseluruhan dalam menghitung seberapa benar model pada klasifikasi. Formula *accuracy* menggunakan persamaan 4.

$$\frac{TP + TN}{\text{Total}} \quad (4)$$

2. *Precision* = Ketika model memprediksi positif, maka perhitungan seberapa sering prediksi tersebut akan benar. Formula *precision* menggunakan persamaan 5.

$$\frac{TP}{FP + TP} \quad (5)$$

3. *Recall* (*Sensitivity* atau *True Positive Rate*) = Ketika kelas aktual positif, maka perhitungan seberapa sering model akan memprediksi akan positif. Formula *recall* menggunakan persamaan 6.

$$\frac{TP}{FN + TP} \quad (6)$$

4. *F1-Score* = Rata-rata harmonik dari model *precision* dan *recall*. Formula *f1-score* menggunakan persamaan 7.

$$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA