

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Adapun beberapa penelitian sebelumnya yang memuat masalah sejenis dengan metode analisis yang sama dengan masalah penelitian yang sedang diteliti, yaitu :

Tabel 2. 1 Perbandingan dengan Penelitian Sebelumnya

Penulis	Judul Penelitian	Objek	Metode	Hasil
Muhammad Yusril Huffandi, Elin Haerani, Fadhulah Syafria, Lola Oktavia (2022)	Klasifikasi Penyakit Paru – Paru Menggunakan Metode <i>Naive Bayes Classifier</i>	Penyakit Paru – Paru	<i>Naive Bayes Classifier</i>	Metode <i>Naive Bayes Classifier</i> efektif dalam mengklasifikasi penyakit paru – paru dengan nilai akurasi sebesar 97,06%.
Anna Hendri Soleliza Jones, Muhchromin	Implementasi Metode <i>Cart</i> untuk	Diagnosis Penyakit	<i>Cart (Classification and</i>	Metode <i>Cart</i> dapat digunakan sebagai metode

Penulis	Judul Penelitian	Objek	Metode	Hasil
Sucron Makmun (2021)	Klasifikasi Diagnosis Penyakit Hepatitis Pada Anak	Hepatitis Pada Anak	<i>Regression Tree</i>	pengklasifikasian untuk diagnosis penyakit hepatitis dengan tingkat akurasi hasil diagnosis sistem sebesar 94%.
Nabilla Yolanda Paramitha, Aang Nuryaman, Ahmad Faisol, Eri Setiawan dan Dina Eka Nurvazly (2023)	Klasifikasi Penyakit Stroke Menggunakan Metode <i>Naive Bayes</i>	Penyakit Stroke	<i>Naive Bayes</i>	Evaluasi model menggunakan <i>confusion matrix</i> pada data penyakit stroke dengan proporsi pembagian data 80:20 menghasilkan akurasi yang lebih tinggi dibandingkan proporsi lainnya.

Penulis	Judul Penelitian	Objek	Metode	Hasil
Raudlatul Munawarah, Oni Soesanto, M.Reza Faisal (2016)	Penerapan Metode <i>Support Vector Machine</i> Pada Diagnosa Hepatitis	Diagnosa Hepatitis	<i>Support Vector Machine</i>	Metode <i>Support Vector Machine</i> dapat digunakan untuk mendiagnosa penyakit hepatitis dengan tingkat akurasi cukup tinggi dengan menggunakan fungsi RBF memiliki tingkat akurasi cenderung lebih tinggi dibandingkan dengan menggunakan fungsi <i>kernel linier</i> .
Quswatun Hasanah,	Analisis Algoritma	Pasien Penderita	<i>Gaussian Naive Bayes</i>	Algoritma <i>Gaussian Naive</i>

Penulis	Judul Penelitian	Objek	Metode	Hasil
Hardian Oktavianto, Yeni Dwi Rahayu (2022)	<i>Gaussian Naive Bayes</i> Terhadap Klasifikasi Data Pasien Penderita Gagal Jantung	Gagal Jantung		<i>Bayes</i> mendapatkan nilai akurasi tertinggi ada pada 10 -fold pada tahap 9 dengan skenario 4 sebesar 69%, presisi tertinggi ada pada 4 -fold pada tahap 4 skenario 2 sebesar 65,73% serta nilai <i>recall</i> tertinggi ada pada 10 – fold pada tahap ke 4 yaitu 95,91%.

Muhammad Yusril Huffandi, Elin Haerani, Fadhulah Syafria, Lola Oktavia (2022) dalam “Klasifikasi Penyakit Paru – Paru Menggunakan Metode *Naive Bayes Classifier*”. Penelitian ini dilakukan pada tanggal 3 Agustus 2022 sampai dengan 3 September 2022 di RSUD Mayjen HA Thalib Kota Sungai. Data penelitian ini

menggunakan data rekam medis pasien penyakit paru – paru dari bulan Juli sampai Agustus sebanyak 134 data pasien. Berdasarkan pengujian menggunakan *RapidMiner* dan pemisahan data berupa 34 data *testing* dan 100 data *training* dengan perbandingan data 7:3 diperoleh nilai akurasi 97,06%.

Anna Hendri Soleliza Jones, Muhchromin Sucron Makmun (2021) dalam “Implementasi Metode Cart untuk Klasifikasi Diagnosis Penyakit Hepatitis Pada Anak”. Penelitian ini menggunakan data dua tahun terakhir dari RSUD Sei Bahar yaitu sebanyak 240 data. Penelitian ini bertujuan untuk membangun sistem aplikasi berbasis komputer dalam menentukan klasifikasi diagnosis penyakit hepatitis dengan menggunakan metode *CART*. Hasil klasifikasi menggunakan metode *CART* dengan menggunakan 35 data uji dan analisis rekomendasi dari pakar dengan nilai tingkat akurasi sebesar 94%.

Nabilla Yolanda Paramitha, Aang Nuryaman, Ahmad Faisol, Eri Setiawan dan Dina Eka Nurvazly (2023) dalam “Klasifikasi Penyakit Stroke Menggunakan Metode Naive Bayes”. Pada penelitian ini data yang digunakan merupakan data sekunder yang diperoleh dari situs *Kaggle* yang berjudul “*Brain Stroke Prediction Dataset*”. Dengan jumlah data sebanyak 200 dan variable sebanyak 11. Penelitian ini bertujuan untuk mengklasifikasi pasien penyakit stroke menggunakan metode Naïve Bayes untuk beberapa proporsi data *training* dan data *testing* yang berbeda. Hasil klasifikasi dengan nilai akurasi sebesar 80% diperoleh ketika proporsi data *training* dan data *testing* 80:20.

Raudlatul Munawarah, Oni Soesanto, M.Reza Faisal (2016) dalam “Penerapan Metode Support Vector Machine Pada Diagnosa Hepatitis”. Peneliti menggunakan uji coba menggunakan data *testing* dengan kedua kernel dengan data *training* 100 data positif dan 100 data negatif. Berdasarkan dari hasil uji coba dapat disimpulkan metode *Support Vector Machine* dapat digunakan untuk mendiagnosa penyakit hepatitis dengan tingkat akurasi 68 – 83% untuk fungsi kernel linier dan 70 – 80% untuk fungsi kernel RBF.

Quswatun Hasanah, Hardian Oktavianto, Yeni Dwi Rahayu (2022) dalam “Analisis Algoritma Gaussian Naive Bayes Terhadap Klasifikasi Data Pasien Penderita Gagal Jantung”. Peneliti menggunakan data dari *kaggle* sebanyak 1000 orang dengan jumlah variabel sebanyak 30. Pada tahap uji coba penelitian ini menggunakan *Cross Fold Validation* dengan $k=2, 4, 5,$ dan 10, kemudian penentuan kriteria hasil klasifikasi terhadap *output* data asli (nilai aktual) menggunakan *confusion matrix*. Berdasarkan dari hasil penelitian validasi terbaik pada fold 10 tahap 9 dengan skenario 4 yaitu sebesar 69%, untuk nilai presisi terbaik terdapat pada fold 4 tahap 4 dengan skenario 2 sebesar 65,73% dan nilai *recall* tertinggi pada fold 10 tahap 4 skenario 4 sebesar 95,91%.

2.2 Dasar Teori

2.2.1 Hepatitis

Penyakit hepatitis merupakan penyakit peradangan hati karena infeksi virus yang menyerang dan menyebabkan kerusakan pada sel – sel dan fungsi organ hati. Penyakit hepatitis merupakan penyakit cikal bakal dari kanker hati. Penyakit hepatitis dapat merusak fungsi organ hati sebagai penetral racun dan sistem

pencernaan makanan dalam tubuh yang mengurai sari-sari makanan untuk kemudian disebar ke seluruh organ tubuh yang sangat penting bagi manusia. Hepatitis atau peradangan hati merupakan salah satu dari banyaknya jenis penyakit hati, yang lainnya seperti pembengkakan hati (*fatty live*) dan kanker hati (*cirrhosis*). Di Indonesia, pada tahun 2007 penyakit hati merupakan salah satu dari sepuluh besar penyakit penyebab kematian terbesar di Indonesia (Departemen Kesehatan RI,2009).

2.2.2 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

Data mining merupakan bidang dari beberapa keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar.

Data mining mengolah data mentah yang tersimpan didalam basis data sehingga menghasilkan informasi yang berguna. Data mentah merupakan data yang disimpan hanya sebagai dokumentasi. Pengolahan data mentah dapat dimanfaatkan untuk pembuatan laporan, mengontrol kebutuhan, menemukan informasi dan pengetahuan baru. Data mining menelusuri data pada database untuk membangun model dan menggunakannya untuk mengenali pola data lain yang tidak tersimpan dalam basis data.

2.2.3 Algoritma Naïve Bayes

Naïve Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan asal Inggris bernama Thomas Bayes, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes.

Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yang sangat kuat (naif) akan independensi dari masing – masing kondisi atau kejadian. Teorema tersebut dikombinasikan dengan naïve dimana diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya.

Persamaan dari teorema Bayes adalah :

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

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

P(H|X) : Probabilitas hipotesis H berdasarkan kondisi X (posteriori probability)

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

P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H

P(x) : Probabilitas atas X

2.2.4 Confusion Matrix

Confusion Matrix merupakan pengujian yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran proses klasifikasi. Dengan confusion matrix dapat dianalisa seberapa baik *classifier* dapat mengenali *record* atau kelas – kelas yang berbeda. Tabel confusion matrix dapat ditunjukkan pada tabel berikut ini :

Tabel 2. 2 Confusion Matrix

Kelas	Terklasifikasi Positive	Terklasifikasi Negative
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

Keterangan :

- TP adalah *True Positive*, yaitu jumlah data positif namun terklasifikasi dengan benar.
- TN adalah *True Negative*, yaitu jumlah data negative namun klasifikasi dengan benar.
- FN adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah.
- FP adalah *False Possitive*, yaitu jumlah data positif namun terklasifikasi salah.

Berdasarkan nilai *True Negative* (TN), *False Negative* (FN), *True Positive* (TP), *False Negative* dapat diperoleh nilai akurasi, presisi dan *recall* :

1. Akurasi

Nilai akurasi menggambarkan seberapa akurat model dalam mengklasifikasi dengan benar. Nilai akurasi dapat diperoleh dengan persamaan berikut :

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

2. Presisi

Nilai presisi menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Presisi dapat diperoleh dengan persamaan berikut :

$$Presisi = \frac{TP}{FP + TP} \times 100\%$$

3. Recall

Nilai recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Recall dapat diperoleh dengan persamaan berikut :

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

2.2.4 Flask

Flask adalah micro framework dari bahasa pemrograman python. Flask disebut microframework karena dia tidak butuh tools atau library. Dia juga tidak ada layer abstraksi, validasi form, atau bahkan komponen lainnya dari pihak ketiga

yang menyediakan fungsi pada umumnya. Namun, Flask bisa menambahkan ekstensi dan fitur di aplikasi ketiga. Flask dibangun tujuan bisa membuat website dengan mudah. Flask akan mengurus tentang masalah server, orm database dan juga terdapat sebagai macam extensions untuk memudahkan website.