

Penerapan Metode Naïve Bayes Dalam Mengklasifikasi Penyakit ISPA di Puskesmas

Sonia Indhira✉, Yuhandri, Billy Hendrik

Fakultas Ilmu Komputer, Universitas Putra Indonesia, Padang, 25221, Indonesia

soniaindhira@gmail.com

Abstract

The rapid development of technology in Indonesia has marked a significant transition in a new era of human life. Technological advancements have impacted various fields, particularly in the healthcare sector, such as in hospitals. One of the diseases that is commonly encountered is Acute Respiratory Infection (ARI), which affects the sinuses, throat, airways, or lungs. The occurrence of cough and cold in toddlers in Indonesia is estimated to be 3 to 6 times per year, meaning that a toddler on average experiences these symptoms 3 to 6 times a year. In addition to viruses and bacteria, many other factors contribute to infections, including exposure to cigarette smoke and weakened immune systems in children and the elderly, making them more vulnerable to attacks from viruses or bacteria that cause infection symptoms. ARI is one of the most common diseases in children, both in developing and developed countries. This study aims to classify ARI cases at the Community Health Center (Puskesmas). The method used in this research is Naïve Bayes, which is chosen for its simplicity and efficiency in classification tasks. The data used in this study consists of ARI cases from the Muaro Bodi Community Health Center, totaling 2,678 cases. From 21 respondents, the data processing results using the Naïve Bayes method show that 6 respondents had severe ARI, 8 respondents had mild ARI, and 7 respondents did not have ARI. The ultimate goal of the ARI classification process is to achieve a higher percentage for ARI cases compared to non-ARI cases. Based on the results, with an accuracy rate of 96.97%, The final objective of this ARI classification process is to have a higher percentage of ARI cases compared to non-ARI cases. Therefore, it can be concluded that the Naïve Bayes model is considered suitable for making reliable decisions in terms of prediction and classification.

Keywords: information technology, machine learning, naïve bayes, ARI, community health center.

Abstrak

Perkembangan teknologi indonesia yang semakin pesat telah menandakan adanya perubahan transisi pada era kehidupan manusia yang baru. Kemajuan teknologi memberikan dampak terhadap berbagai macam bidang, khususnya pekerjaan didalam dunia kesehatan seperti di Rumah Sakit. Salah satu penyakit yaitu Infeksi Saluran Pernapasan Akut atau yang kita sering sebut ISPA. ISPA menjadi salah satu penyakit terbanyak yang diderita oleh anak-anak, baik dinegara berkembang maupun dinegara maju. Penelitian ini bertujuan untuk mengklasifikasi penyakit ISPA Pada Puskesmas Muaro Bodi. Puskesmas Muaro Bodi adalah satu satunya Puskesmas yang ada di Kecamatan IV Nagari Kabupaten Sijunjung. Metode yang digunakan dalam penelitian ini adalah Naïve Bayes. Naïve Bayes dipilih karena merupakan salah satu dari metode pengklasifikasian yang simpel dan efisien. Data yang digunakan dalam penelitian ini merupakan data dari kasus ISPA yang ada pada Puskesmas Muaro Bodi dengan total mencapai 2.678 kasus. Dari 21 Responden didapat hasil pengolahan menggunakan metode Naïve Bayes yang terdiri dari 6 Responden ISPA Berat, 8 Responden ISPA Ringan dan 7 Responden Tidak ISPA. Tujuan akhir dari proses Klasifikasi penyakit ISPA dengan presentase kelas Penyakit ISPA lebih besar dari kelas Tidak ISPA. Berdasarkan hasil yang diperoleh dengan tingkat akurasi mencapai 96,97%, maka dapat dinyatakan model proses Naïve Bayes termasuk layak digunakan sebagai penentuan hasil keputusan yang baik dalam hal prediksi dan klasifikasi. Penelitian ini dapat berkontribusi untuk membantu masyarakat dalam memprediksi penyakit ISPA yang dideritanya, dan membantu pihak Puskesmas dalam memantau pasiennya.

Kata kunci: teknologi informasi, machine learning, naïve bayes, ISPA, puskesmas

KomtekInfo is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



1. Pendahuluan

Perkembangan teknologi indonesia yang semakin pesat telah menandakan adanya perubahan transisi pada era kehidupan manusia yang baru. Hal ini tidak dapat dipungkiri bahwa kehidupan manusia telah memasuki era digital 5.0 yang memberikan dampak dan efek dengan bertujuan untuk mempermudah segala

pekerjaan manusia [1]. Perkembangan teknologi yang semakin pesat ini diikuti dengan arus globalisasi yang semakin pesat juga. Kemajuan teknologi memberikan dampak terhadap berbagai macam bidang, khususnya pekerjaan didalam dunia kesehatan seperti di Rumah Sakit [2]. Infeksi Saluran Pernapasan Akut atau yang kita sering sebut ISPA merupakan infeksi yang terjadi pada bagian sinus, tenggorokan, saluran udara, atau

paru-paru. Selain itu dari virus dan bakteri banyak factor-faktor penyebab terjadinya infeksi yaitu salah satunya dari asap rokok, system kekebalan tubuh pada anak-anak, dan lanjut usia yang sangat lemah membuat mereka semakin rentan terhadap serangan virus atau bakteri yang menyebabkan gejala infeksi. ISPA menjadi salah satu penyakit terbanyak yang diderita oleh anak-anak, baik dinegara berkembang maupun dinegara maju [3]. Tingkat keparahan penyakit ISPA bervariasi dari faringitis ringan sampai pneumonia bahkan sampai pada kematian [4]. Kejadian penyakit batuk pilek pada balita di Indonesia diperkirakan 3 sampai 6 kali per tahun, yang berarti seorang balita rata-rata mendapat serangan batuk pilek sebanyak 3 sampai 6 kali setahun [5]. ISPA adalah penyakit saluran pernapasan atas atau bawah, biasanya menular, yang dapat menimbulkan berbagai spektrum penyakit yang berkisar dari penyakit tanpa gejala atau infeksi ringan sampai penyakit yang parah dan mematikan, tergantung pada patogen penyebabnya, faktor lingkungan, dan faktor pejamu [6]. Sistem Pakar merupakan kemampuan komputer untuk mengubah pengetahuan dari manusia ke dalam komputer sehingga dapat membantu mengatasi masalah yang hanya dapat diselesaikan oleh para ahli [7]. Sistem pakar yang dalam menyelesaikan permasalahan pada penyakit mulut dan kuku salah satunya menggunakan metode *Naïve Bayes*. *Naïve Bayes* yaitu metode probabilitas dengan memprediksi peluang yang akan datang berdasarkan pengalaman di masa lalu[8].

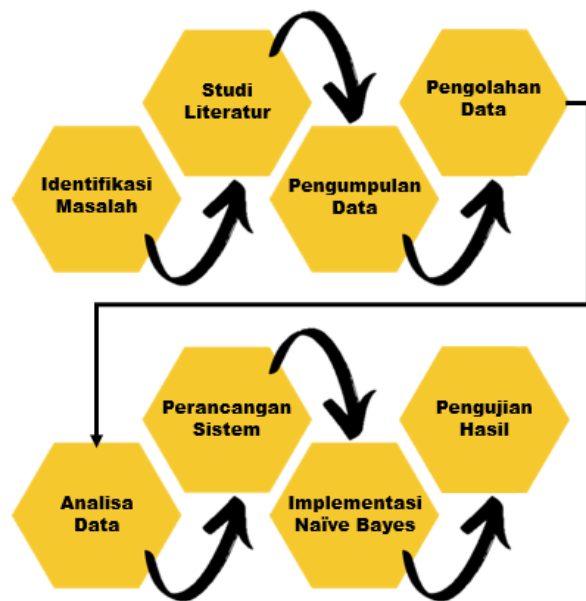
Metode yang digunakan didalam penelitian ini adalah metode *Naïve Bayes*. *Naïve Bayes* merupakan algoritma yang dapat mengklasifikasi suatu variabel tertentu dengan menggunakan metode probabilitas dan statistic[9]. Bergantung pada model probolitasnya, *Naïve Bayes* classifier dapat dilatih untuk melakukan supervised learning dengan sangat efektif. *Naïve bayes* tidak membutuhkan jumlah data training yang banyak[10]. *Naïve Bayes* merupakan metodeklasifikasi sederhana yang menghitung semuaprobabilitas berdasarkan teorema Bayes yang dikombinasikan dengan kombinasi nilai frekuensi database. [11]

Penelitian sebelumnya yang dilakukan oleh Indra Gunawan dan Yusra Fernando pada tahun 2021 didalam penelitian Nugroho & Prambodo tentang implementasi metode *Naïve Bayes* dalam penerapan sistem pakar untuk mendiagnosa penyakit kulit pada kucing menyimpulkan dari hasil penerapan sistem pakar mampu mendiagnosa penyakit kulit pada kucing sebesar 80% dari data uji sampel [12]. Penelitian lainnya oleh yuliana ddk yaitu Sistem yang dibangun dengan menerapkan metode *Naïve Bayes* Classifier mampu mendiagnosis jenis penyakit ISPA secara akurat dengan persentase 90% berdasarkan data dan gejala yang dialami pasien [13]. Diagnosis penyakit Karies pada gigi manusia dengan metode *Naïve Bayes* oleh Indah dkk dapat mengidentifikasi penyakit Karies pada manusia dan memberikan solusi yang tepat untuk menangani kerusakan pada gigi, dengan tingkat akurasi sistem sebesar 83,61% [14].

Keunggulan penelitian ini dibandingkan dengan penelitian lain adalah penerapan *Naïve Bayes* dibuat khusus untuk mengklasifikasi penyakit ISPA pada Puskesmas. Fokusnya adalah untuk mengetahui gejala, penyebab dan solusi. Pada sistem ini penyakit ISPA dapat diklasifikasi berdasarkan gejala-gejala yang ada pada pasien. Sistem ini memuat basis pengetahuan, halaman cek ISPA, halaman hasil diagnosa. Pada Puskesmas Muaro Bodi belum ada sistem untuk penerapan metode *Naïve bayes* dalam membatu mengklasifikasi penyakit ISPA. Mengklasifikasi penyakit ISPA memerlukan kriteria gejala dan penyebab agar dapat menentukan solusi dari penyakit ISPA yang diderita oleh pasien. Penelitian ini bertujuan untuk membatu pasien ISPA dalam mengetahui gejala, penyebab dan solusi dari penyakit ISPA yang diderita oleh pasien.

2. Metodologi Penelitian

Metodologi Penelitian ini dilakukan dengan metode *Naïve Bayes* dengan pembentukan kerangka kerja di mana kerangka kerja ini perlu agar penelitian bisa dilakukan dengan terstruktur dan terarah sesuai tujuan yang diharapkan. Kerangka kerja merupakan tahapan-tahapan proses penelitan yang terurut berdasarkan langkah-langkah yang saling berkaitan. Langkah-langkah tersebut dimulai dengan proses Identifikasi Masalah, Studi iteratur, Pengumpulan Data (Observasi, Wawancara dan Kuisisioner), Pengolahan Data, Analisa Data, Membangun Sistem, Impelementasi dan Pembahasan Hasil.



Gambar 1. Kerangka Kerja Penelitian

Gambar 1 menunjukkan kerangka kerja penelitian yang merangkum tahapan-tahapan dalam proses pengklasifikasi penyakit ISPA menggunakan metode *Naïve Bayes*. Setiap tahapan dalam kerangka kerja ini meliputi identifikasi masalah, studi literatur, pengumpulan data, pengolahan data, analisa data, merancang sistem, implementasi *Naïve Bayes*,

pengujian hasil, di mana kerangka kerja ini perlu agar penelitian bisa dilakukan dengan terstruktur dan terarah sesuai tujuan yang diharapkan, sebagai berikut:

1. Identifikasi Masalah

Tahapan identifikasi masalah adalah langkah awal dalam penelitian yang sangat penting. Pada tahap ini, peneliti meninjau objek yang akan diteliti, mengidentifikasi permasalahan utama yang terkait dengan objek tersebut, serta menentukan ruang lingkup penelitian. Proses ini membantu memastikan bahwa penelitian fokus pada isu-isu yang relevan dan dapat menghasilkan temuan yang signifikan. Berdasarkan hal tersebut maka penulis dapat menentukan rencana kerja serta menentukan data apa saja yang dibutuhkan dalam penelitian ini.

2. Studi Literatur

Studi literatur dilakukan untuk mengumpulkan berbagai sumber tertulis yang relevan, seperti buku, jurnal, dan dokumen terkait. Peneliti mencari referensi yang berhubungan dengan kasus ISPA dan metode Naïve Bayes. Kajian ini kemudian digunakan sebagai landasan teori dan acuan dalam pengembangan penelitian.

3. Pengumpulan Data

Pengumpulan data disini adalah data yang di peroleh dari Puskesmas Muaro Bodi dari Tahun 2022 sampai Tahun 2024. Data tersebut berfungsi sebagai dasar untuk analisis dan prediksi menggunakan metode Naïve Bayes. Sumber data juga di dapat dalam penulisan ada 3 yaitu Observasi, Wawancara, dan Penyebaran Kuesioner.

- a. Observasi adalah teknik pengumpulan data yang dilakukan dengan cara mengamati secara langsung ke tempat penelitian.
- b. Wawancara adalah teknik pengumpulan data dengan cara melakukan tanya jawab secara langsung antara peneliti dan responden dan juga antara peneliti dan karyawan Puskesmas.
- c. Kuesioner adalah teknik pengumpulan data dengan cara memberikan daftar pertanyaan tertulis kepada responden untuk diisi.

4. Pengolahan Data

Pengolahan Data disini adalah dari data yang sudah dikumpulkan maka akan dilakukan pengolahan data. Hasil dari pengolahan data Observasi, Wawancara dan Kuisisioner maka dapat dilanjutkan dengan analisa data menggunakan Microsoft Excel untuk hasil dari perhitungan yang sudah di dapat. Kemudian dari perhitungan yang di dapat maka di lanjutkan ke bagian Analisa Data.

5. Analisa Data

Analisa Data pada penelitian ini yang dipergunakan adalah kumpulan hasil tanggapan dari kuesioner dengan parameter. Seluruh data yang sudah didapatkan

menggunakan metode Naïve Bayes kemudian akan di analisa dan proses klasifikasi menggunakan Microsoft Excel untuk mempermudah penulis dalam menganalisa data selanjutnya. Kemudian setelah data di analisa maka di lanjutkan ke bagian Perancangan Sistem.

6. Perancangan Sistem

Pada tahapan ini setelah proses analisis selesai, maka dilakukan tahapan perancangan proses perhitungan menggunakan rumus Naive Bayes. Perancangan dilakukan menggunakan Microsoft Excel. Perancangan disini menggunakan rumus Naïve Bayes yang bisa dilihat melalui tahapan sebagai berikut:

- a. Melakukan proses pengolahan data secara manual menggunakan Software Microsoft Excel menggunakan rumus Naive Bayes.
- b. Melakukan proses perhitungan menggunakan program yang sudah di rancang dengan menggunakan rumus Naive Bayes dengan data yang sama.

7. Implementasi Naïve Bayes

Setelah proses perancangan sistem menggunakan rumus Naïve Bayes selesai. Mengasumsikan fitur-fitur yang digunakan untuk memprediksi kelas bersifat independen satu sama lain. Proses tahapan Selanjutnya yaitu Implementasi sistem dengan menggunakan metode Naïve Bayes untuk mengkonfirmasi modul-modul perancangan, sehingga pengguna dapat memberi masukan kepada pengembang project management dimasa yang akan datang.

8. Pengujian Hasil

Pengujian Hasil disini adalah proses evaluasi kinerja model setelah dilatih dengan data tertentu. Pengujian ini melibatkan penggunaan data untuk memprediksi kelas dari data tersebut. Bagian terakhir dari tahapan penelitian ini adalah pengujian hasil. Pada tahap ini yaitu dilakukan proses menguji hasil dari perhitungan baik itu manual dengan menerapkan formula dari algoritma Naive Bayes. Berikut ini tahapan-tahapan dari proses pengujian:

- a. Baca data training.

Baca data training disini adalah langkah awal dalam proses pelatihan model. Data training adalah kumpulan data yang sudah diberi label, digunakan untuk melatih model agar dapat mengenali pola dan hubungan antara fitur-fitur input dan kelas target. Data training yang baik dan representatif sangat penting untuk membangun model yang akurat dan dapat diandalkan.

- b. Hitung jumlah dan probabilitas.

Menghitung jumlah dan probabilitas adalah langkah inti dalam proses pelatihan dan prediksi. Jumlah kemunculan setiap fitur dalam masing-masing kelas dihitung dari data training. Selanjutnya, probabilitas kondisional dihitung berdasarkan jumlah tersebut,

menggunakan Teorema Bayes. Probabilitas ini menggambarkan seberapa besar kemungkinan suatu kelas diberikan fitur tertentu, dengan asumsi bahwa semua fitur bersifat independen. Selama prediksi, model menghitung probabilitas gabungan dari semua fitur untuk setiap kelas, dan kelas dengan probabilitas tertinggi akan dipilih sebagai prediksi akhir.

1) Perhitungan Probabilitas Prior.

Perhitungan probabilitas prior adalah langkah untuk menentukan probabilitas awal dari setiap kelas sebelum mempertimbangkan fitur-fitur yang ada. Probabilitas prior ini dihitung dengan membagi jumlah instance dalam suatu kelas dengan total jumlah instance dalam data training. Probabilitas prior mencerminkan seberapa umum atau jarang suatu kelas dalam data, dan ini digunakan sebagai dasar untuk menghitung probabilitas posterior saat model melakukan klasifikasi.

$$P(C_i) = \frac{S_i}{S} \tag{1}$$

2) Menghitung Probabilitas Data pada masing-masing Kategori.

Menghitung probabilitas data pada masing-masing kategori adalah proses di mana model menentukan kemungkinan bahwa suatu data termasuk dalam setiap kelas yang ada, berdasarkan fitur-fitur yang dimilikinya. Probabilitas ini dihitung dari data training dengan membagi jumlah kemunculan fitur dalam kelas dengan total jumlah instance dalam kelas tersebut. Proses ini memungkinkan model Naive Bayes untuk membuat prediksi klasifikasi terhadap data baru.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2}$$

c. Baca Data Testing.

Baca data testing adalah proses memuat dan mempersiapkan data yang akan digunakan untuk menguji model yang sudah dilatih. Data testing adalah kumpulan data yang belum pernah digunakan dalam pelatihan model dan berfungsi untuk mengevaluasi seberapa baik model dapat memprediksi kelas yang benar. Proses ini melibatkan pengambilan data dari sumber yang sama seperti data training, kemudian memastikan bahwa data tersebut berada dalam format yang sesuai untuk digunakan oleh model.

d. Menghitung Probabilitas Data Testing.

Menghitung probabilitas data testing adalah proses di mana model menghitung probabilitas bahwa data uji termasuk dalam setiap kelas yang ada, menggunakan probabilitas prior yang sudah dihitung dari data training. Untuk setiap instance dalam data testing kemudian dipilih sebagai prediksi. Proses ini memungkinkan model untuk menentukan klasifikasi yang paling mungkin untuk setiap data baru yang dihadapi.

$$P(X_1|C) \cdot P(X_2|C) \dots P(X_n|C) = P(C|X_1, \dots, X_n) \cdot P(X_1, \dots, X_n) = P(C) \cdot \prod_{i=1}^n P(X_i|C) \tag{3}$$

$$P(A|B) = P(B|A) \cdot P(A|B) \tag{4}$$

e. Hasil Probabilitas Akhir.

Hasil probabilitas akhir menunjukkan kemungkinan terjadinya suatu peristiwa dengan tingkat ketepatan tertentu. Menganalisis hasil probabilitas akhir, kita dapat memahami seberapa besar peluang keberhasilan atau kegagalan dalam suatu eksperimen atau keputusan, membantu dalam perencanaan dan strategi yang lebih efektif. Penjabaran hasil probabilitas akhir membantu dalam mengidentifikasi risiko dan peluang secara kuantitatif, memberikan dasar yang kuat untuk membuat keputusan yang lebih terinformasi dan mengelola ketidakpastian dengan lebih baik.

3. Hasil dan Pembahasan

Hasil dan pembahasan dari penelitian ini adalah menganalisa data dengan mencari proses serta menyusun secara sistematis data yang telah diperoleh dari wawancara, observasi atau kuesioner dengan cara mengorganisasikan data tersebut ke dalam kategori dan menjabarkan ke dalam unit-unit, menyusun ke dalam pola, memilih mana yang penting dan mana yang akan dipelajari serta kemudian membuat kesimpulan agar dapat dipahami penulis maupun orang lain.

3.1. Analisa Data menggunakan metode Naïve Bayes

Kegiatan yang dilakukan dalam analisa data, penulis akan menganalisa kebutuhan apa saja yang diperlukan dalam mengklasifikasikan penyakit ISPA di Puskesmas Muaro Bodi. Dalam mengumpulkan data untuk menentukan tingkat keparahan penyakit ISPA kemudian dilakukan penyebaran kuesioner kepada Pasien yang berobat di Puskesmas Muaro Bodi tersaji pada Tabel 1.

Tabel 1. Data Training

No	Responden	Parameter Penyakit			Terindikasi
		K1	K2	K3	
1	A1	Y	Y	Y	ISPA Berat
2	A2	Y	Y	T	ISPA Ringan
3	A3	Y	Y	Y	ISPA Berat
4	A4	T	Y	Y	Tidak ISPA
5	A5	Y	Y	Y	ISPA Berat
6	A6	Y	Y	Y	ISPA Berat
7	A7	T	Y	Y	Tidak ISPA
8	A8	T	T	T	Tidak ISPA
9	A9	Y	T	Y	ISPA Ringan
..
87	A87	T	T	T	Tidak ISPA

Tabel 2. Data Testing

No	Responden	Parameter Penyakit			Terindikasi
		K1	K2	K3	
1	A1	T	T	T	?
2	A2	Y	T	Y	?
3	A3	Y	T	T	?
4	A4	T	Y	T	?
5	A5	Y	T	Y	?

6	A6	Y	Y	T	?
7	A7	Y	Y	Y	?
8	A8	T	Y	Y	?
9	A9	T	T	T	?
10	A10	Y	T	Y	?
11	A11	Y	Y	Y	?
12	A12	Y	Y	Y	?
13	A13	Y	T	Y	?
14	A14	Y	Y	Y	?
15	A15	Y	T	T	?
16	A16	Y	T	Y	?
17	A17	Y	Y	Y	?
18	A18	Y	T	Y	?
19	A19	Y	Y	Y	?
20	A20	Y	T	T	?
21	A21	Y	Y	T	?

Keterangan:

K1 = Demam Y = Ya
 K2 = Batuk T = Tidak
 K3 = Pilek

Tabel 3. Parameter Penyakit

Kategori	Pertanyaan
K1	Apakah anda mengalami Demam dalam 1 bulan terakhir ?
K2	Apakah anda mengalami Batuk dalam 1 bulan terakhir ?
K3	Apakah anda mengalami Pilek dalam 1 bulan terakhir ?

Langkah awal penulis akan melakukan perhitungan jumlah ISPA dan Tidak ISPA berdasarkan Dari 87 data set yang digunakan, diketahui kelas ISPA sebanyak 43 data, dan kelas Tidak ISPA sebanyak 44 data. Tahapan proses perhitungan menggunakan metode Naive Bayes untuk klasifikasi penyakit ISPA sebagai berikut:

a. Baca Data Training

Baca data ttraining adalah Baca data training disini adalah langkah awal dalam proses pelatihan model. Data training adalah kumpulan data yang sudah diberi label, digunakan untuk melatih model agar dapat mengenali pola dan hubungan antara fitur-fitur input dan kelas target. Data training yang baik dan representatif sangat penting untuk membangun model yang akurat dan dapat diandalkan. Seperti yang tertera pada Tabel 1 data set yang digunakan bukanlah data numerik, maka proses selanjutnya menghitung jumlah dan probabilitas dari setiap kategori.

b. Hitung Jumlah dan Probabilitas

Menghitung jumlah dan probabilitas adalah langkah inti dalam proses pelatihan dan prediksi. Jumlah kemunculan setiap fitur dalam masing-masing kelas dihitung dari data training. Selanjutnya, probabilitas kondisional dihitung berdasarkan jumlah tersebut, menggunakan Teorema Bayes. Probabilitas ini menggambarkan seberapa besar kemungkinan suatu kelas diberikan fitur tertentu, dengan asumsi bahwa

semua fitur bersifat independen. Selama prediksi, model menghitung probabilitas gabungan dari semua fitur untuk setiap kelas, dan kelas dengan probabilitas tertinggi akan dipilih sebagai prediksi akhir.

1) Perhitungan Probabilitas Prior ($P(C_i)$)

Perhitungan probabilitas prior adalah langkah untuk menentukan probabilitas awal dari setiap kelas sebelum mempertimbangkan fitur-fitur yang ada. Probabilitas prior ini dihitung dengan membagi jumlah instance dalam suatu kelas dengan total jumlah instance dalam data training. Probabilitas prior mencerminkan seberapa umum atau jarang suatu kelas dalam data, dan ini digunakan sebagai dasar untuk menghitung probabilitas posterior saat model melakukan klasifikasi. Berdasarkan dari data set Tabel 1 jumlah data yang digunakan sebanyak 87 data di mana terdapat 43 data yang menyatakan ISPA dan 44 menyatakan Tidak ISPA. Proses perhitungan Probabilitas Prior pada kelas ISPA dan Tidak ISPA dengan menggunakan rumus persamaan (1) sebagai berikut:

$$P(C_i) = \frac{S_i}{S} \quad P(C_{ISPA}) = \frac{total\ ispa}{total\ data}$$

$$P(C_{ISPA}) = \frac{43}{87} = 0,49$$

$$P(C_{bukan\ ISPA}) = \frac{44}{87} = 0,51$$

2) Menghitung Probabilitas Data pada masing-masing Kategori.

Menghitung probabilitas data pada masing-masing kategori adalah proses di mana model menentukan kemungkinan bahwa suatu data termasuk dalam setiap kelas yang ada, berdasarkan fitur-fitur yang dimilikinya. Probabilitas ini dihitung dari data training dengan membagi jumlah kemunculan fitur dalam kelas dengan total jumlah instance dalam kelas tersebut. Proses ini memungkinkan model Naive Bayes untuk membuat prediksi klasifikasi terhadap data baru. Menghitung probabilitas data pada masing-masing kategori sebagai berikut:

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

$$P(ISPA\ Ringan|Demam) = \frac{25}{49} = 0,51$$

$$P(ISPA\ Berat|Demam) = \frac{18}{49} = 0,37$$

$$P(Tidak\ ISPA|Demam) = \frac{6}{49} = 0,12$$

$$P(\text{ISPA Ringan}|\text{Batuk}) = \frac{8}{47} = 0,17$$

$$P(\text{ISPA Berat}|\text{Batuk}) = \frac{18}{47} = 0,38$$

$$P(\text{Tidak ISPA}|\text{Batuk}) = \frac{21}{47} = 0,45$$

$$P(\text{ISPA Ringan}|\text{Pilek}) = \frac{17}{55} = 0,31$$

$$P(\text{ISPA Berat}|\text{Pilek}) = \frac{18}{55} = 0,33$$

$$P(\text{Tidak ISPA}|\text{Pilek}) = \frac{20}{55} = 0,36$$

$$P(\text{ISPA Ringan}|\text{Tidak Demam}) = \frac{0}{38} = 0,00$$

$$P(\text{ISPA Berat}|\text{Tidak Demam}) = \frac{0}{38} = 0,00$$

$$P(\text{Tidak ISPA}|\text{Tidak Demam}) = \frac{38}{38} = 1,00$$

$$P(\text{ISPA Ringan}|\text{Tidak Batuk}) = \frac{17}{40} = 0,43$$

$$P(\text{ISPA Berat}|\text{Tidak Batuk}) = \frac{0}{40} = 0,00$$

$$P(\text{Tidak ISPA}|\text{Tidak Batuk}) = \frac{23}{40} = 0,58$$

$$P(\text{ISPA Ringan}|\text{Tidak Pilek}) = \frac{8}{32} = 0,25$$

$$P(\text{ISPA Berat}|\text{Tidak Pilek}) = \frac{0}{32} = 0,00$$

$$P(\text{Tidak ISPA}|\text{Tidak Pilek}) = \frac{24}{32} = 0,75$$

Tabel 4. Hasil Probabilitas Kelas

Demam = ya ispa	0,88
Batuk = ya ispa	0,55
Pilek = ya ispa	0,64
Demam = tidak ispa	0,00
Batuk = tidak ispa	0,43
Pilek = tidak ispa	0,25

Setelah semua kelas didapatkan. Kemudian masuk ke bagian hasil perhitungan. Hasil perhitungan probabilitas pada parameter penyakit dapat dilihat pada Tabel dibawah ini:

Tabel 5. Hasil Perhitungan Probabilitas Parameter Penyakit

Kategori	ISPA	Tidak ISPA	Jumlah
Demam	49	38	87
Batuk	47	40	87
Pilek	55	32	87

Tabel 6. Hasil Perhitungan Probabilitas Parameter Penyakit

Kategori	ISPA	Tidak ISPA	Jumlah
Demam	49/87 = 0,56	38/87 = 0,44	1
Batuk	47/87 = 0,54	40/87 = 0,46	1
Pilek	55/87 = 0,63	32/87 = 0,37	1

Tabel diatas adalah hasil dari perhitungan probabilitas parameter penyakit berdasarkan kategorinya. Selanjutnya masuk pada bagian baca data testing. Menggunakan 21 data responden.

c. Baca Data Testing

Baca data testing adalah proses memuat dan mempersiapkan data yang akan digunakan untuk menguji model yang sudah dilatih. Data testing adalah kumpulan data yang belum pernah digunakan dalam pelatihan model dan berfungsi untuk mengevaluasi seberapa baik model dapat memprediksi kelas yang benar. Proses ini melibatkan pengambilan data dari sumber yang sama seperti data training, kemudian memastikan bahwa data tersebut berada dalam format yang sesuai untuk digunakan oleh model. Data uji yang digunakan dalam penelitian ini terdiri dari 21 data responden..

d. Menghitung Probabilitas Data Testing

Menghitung probabilitas data testing adalah proses di mana model menghitung probabilitas bahwa data uji termasuk dalam setiap kelas yang ada, menggunakan probabilitas prior yang sudah dihitung dari data training. Untuk setiap instance dalam data testing kemudian dipilih sebagai prediksi. Proses ini memungkinkan model untuk menentukan klasifikasi yang paling mungkin untuk setiap data baru yang dihadapi.

(Perhitungan Kelas Tidak ISPA)

$$P(\text{Tidak ISPA}|Y_a) = P(\text{Demam}=\text{Tidak ISPA}) \times P(\text{Batuk}=\text{Tidak ISPA}) \times P(\text{Pilek}=\text{Tidak ISPA})$$

$$= (0,00 \times 0,43 \times 0,25)$$

$$= 0,00$$

$$P(\text{Tidak ISPA}|Y_a) = P(\text{Demam}=Y_a) \times P(\text{Batuk}=\text{Tidak}) \times P(\text{Pilek}=\text{Tidak})$$

$$= (0,88 \times 0,43 \times 0,25)$$

$$= 0,09$$

(Perhitungan Kelas ISPA Ringan)

$$P(\text{ISPA Ringan}|\text{Ya}) = P(\text{Demam}=\text{Ya}) \times P(\text{Batuk}=\text{Tidak}) \times P(\text{Pilek}=\text{Ya})$$

$$= (0,88 \times 0,43 \times 0,64)$$

$$= 0,24$$

$$P(\text{ISPA Ringan}|\text{Ya}) = P(\text{Demam}=\text{Ya}) \times P(\text{Batuk}=\text{Ya}) \times P(\text{Pilek}=\text{Tidak})$$

$$= (0,88 \times 0,55 \times 0,25)$$

$$= 0,12$$

(Perhitungan Kelas ISPA ISPA Berat)

$$P(\text{ISPA Berat}|\text{Ya}) = P(\text{Demam}=\text{Ya}) \times P(\text{Batuk}=\text{Ya}) \times P(\text{Pilek}=\text{Ya})$$

$$= (0,88 \times 0,55 \times 0,64)$$

$$= 0,31$$

Perhitungan Kelas ISPA :

$$\text{ISPA Berat} \geq 0,30$$

$$\text{ISPA Ringan} \leq 0,29-0,11$$

$$\text{Tidak ISPA} \leq 0,10-0,00$$

Setelah nilai dari perhitungan kelas ISPA didapatkan. Selanjutnya tabel perhitungan probabilitas data testing. Berikut hasil perhitungan probabilitas data testing dengan Kriteria dapat dilihat pada Tabel 7.

Tabel 7. Perhitungan Probabilitas Data Testing

No	Responden	Parameter Penyakit		
		K1	K2	K3
1	A1	0,00	0,43	0,25
2	A2	0,88	0,43	0,64
3	A3	0,88	0,43	0,25
4	A4	0,00	0,55	0,25
5	A5	0,88	0,43	0,64
6	A6	0,88	0,55	0,25
7	A7	0,88	0,55	0,64
8	A8	0,00	0,55	0,64
9	A9	0,00	0,43	0,25
10	A10	0,88	0,43	0,64
11	A11	0,88	0,55	0,64
12	A12	0,88	0,55	0,64
13	A13	0,88	0,43	0,64
14	A14	0,88	0,55	0,64
15	A15	0,88	0,43	0,25
16	A16	0,88	0,43	0,64
17	A17	0,88	0,55	0,64
18	A18	0,88	0,43	0,64
19	A19	0,88	0,55	0,64
20	A20	0,88	0,43	0,25
21	A21	0,88	0,55	0,25

Tabel di atas adalah hasil dari perhitungan data testing. Setiap responden sudah memiliki bobot dari parameter penyakit. Kemudian tentukan hasil probabilitas akhir.

e. Hasil Probabilitas Akhir

Hasil probabilitas akhir menunjukkan kemungkinan terjadinya suatu peristiwa dengan tingkat ketepatan tertentu. Probabilitas akhir merupakan pemaksimalan klasifikasi Tidak ISPA, ISPA Ringan dan ISPA Berat. Langkah ini menentukan hasil dari semua langkah yang sudah di lakukan. Hasil perhitungan dapat dilihat pada Tabel 8 dibawah:

Tabel 7. Hasil Perhitungan Data Probabilitas Akhir

No	Responden	Parameter			Terindikasi	Hasil
		K1	K2	K3		
1	A1	0,00	0,43	0,25	0,00	Tidak ISPA
2	A2	0,88	0,43	0,64	0,24	ISPA Ringan
3	A3	0,88	0,43	0,25	0,09	Tidak ISPA
4	A4	0,00	0,55	0,25	0,00	Tidak ISPA
5	A5	0,88	0,43	0,64	0,24	ISPA Ringan
6	A6	0,88	0,55	0,25	0,12	ISPA Ringan
7	A7	0,88	0,55	0,64	0,31	ISPA Berat
8	A8	0,00	0,55	0,64	0,00	Tidak ISPA
9	A9	0,00	0,43	0,25	0,00	Tidak ISPA
10	A10	0,88	0,43	0,64	0,24	ISPA Ringan
11	A11	0,88	0,55	0,64	0,31	ISPA Berat
12	A12	0,88	0,55	0,64	0,31	ISPA Berat
13	A13	0,88	0,43	0,64	0,24	ISPA Ringan
14	A14	0,88	0,55	0,64	0,31	ISPA Berat
15	A15	0,88	0,43	0,25	0,09	Tidak ISPA
16	A16	0,88	0,43	0,64	0,24	ISPA Ringan
17	A17	0,88	0,55	0,64	0,31	ISPA Berat
18	A18	0,88	0,43	0,64	0,24	ISPA Ringan
19	A19	0,88	0,55	0,64	0,31	ISPA Berat
20	A20	0,88	0,43	0,25	0,09	Tidak ISPA
21	A21	0,88	0,55	0,25	0,12	ISPA Ringan

Tabel 8. Klasifikasi Tingkat ISPA

ISPA Berat	\geq	0,30
ISPA Ringan	\leq	0,29 – 0,11
Tidak ISPA	\geq	0,10-0,00

Berdasarkan hasil klasifikasi yang terdapat pada tabel 8 dapat dilihat bahwa dalam klasifikasi tingkat penyakit ISPA apakah masuk kategori ISPA Ringan, ISPA Berat atau Tidak ISPA, dapat dilihat dari nilai probabilitas akhir Hasil Indikasi. Jika nilai probabilitas akhir lebih besar atau sama dengan 0,30 maka akan masuk kedalam kategori ISPA Berat, jika nilai probabilitas akhir lebih kecil atau sama dengan 0,29 sampai 0,11 maka akan masuk kedalam kategori ISPA Ringan, jika nilai probabilitas akhir lebih kecil atau sama dengan 0,10 sampai 0,00 maka akan masuk kedalam kategori Tidak ISPA. Dari hasil pengolahan menggunakan metode *Naive Bayes* diperoleh 6 Responden ISPA Berat, 8 Responden ISPA Ringan dan 7 Responden Tidak ISPA.

Kesimpulan

Berdasarkan dari hasil pembahasan, perhitungan, dan pengujian dalam mengklasifikasi penyakit ISPA menggunakan metode *Naive Bayes* maka didapat kesimpulan yaitu hasil klasifikasi yang terdapat pada tabel 7 dapat dilihat bahwa dalam klasifikasi tingkat penyakit ISPA di Puskesmas Muaro Bodi apakah masuk

kategori ISPA Ringan, ISPA Berat atau Tidak ISPA, dapat dilihat dari nilai probabilitas akhir Hasil Indikasi. Jika nilai probabilitas akhir lebih besar atau sama dengan 0,30 maka akan masuk kedalam kategori ISPA Berat, jika nilai probabilitas akhir lebih kecil atau sama dengan 0,29 sampai 0,11 maka akan masuk kedalam kategori ISPA Ringan, jika nilai probabilitas akhir lebih kecil atau sama dengan 0,10 sampai 0,00 maka akan masuk kedalam kategori Tidak ISPA. Dari hasil pengolahan menggunakan metode Naive Bayes diperoleh 6 Responden ISPA Berat, 8 Responden ISPA Ringan dan 7 Responden Tidak ISPA, diperoleh tingkat akurasi mencapai 96,97%. Tujuan akhir dari proses Klasifikasi penyakit ISPA dengan presentase kelas Penyakit ISPA lebih besar dari kelas Tidak ISPA. Maka dapat dinyatakan model proses Naive Bayes termasuk layak digunakan sebagai penentuan hasil keputusan yang baik dalam hal prediksi dan klasifikasi.

Daftar Rujukan

- [1] Ananda Mustari, K., Assiroj, P., Hartati, B., & Samuel, F. (2024). Implementasi Data Mining Pada Instansi Pemerintahan (Systematic Literature Review). In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 3).
- [2] Wijaya, S., Rizkiyah, A. D., Kesehatan, F., Nahdlatul, U., & Surabaya, U. (2024). Penggunaan Teknologi Pada Manajemen Sumber Daya Manusia Di Rumah Sakit. *Jurnal Kesehatan Masyarakat Indonesia*, 1(3), 35–47. <https://doi.org/10.62017/Jkmi>
- [3] Hidayatullah, W., & Darmawan Bakti, L. (2023). Sistem Pakar Diagnosis Penyakit Ispa Menggunakan Metode Naive Bayes Berbasis Web Pada Puskesmas Teratak. 2(1), 32–42.
- [4] Indhira, S., & Hendrik, B. (2023). Penerapan Algoritma Decision Tree C4.5 Untuk Diagnosa Penyakit Ispa Pada Puskesmas Sabak Auh. 1(2), 6–9
- [5] Herlina Susanto Sunuh, Mk., Subagyo, I., & Bungawati, Mk. A. (2023). Pencemaran Udara Dan Ispa (Infeksi Saluran Pernapasan Akut) Penerbit Cv.Eureka Media Aksara.
- [6] Romli, I. (2021). Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Ispa. *Indonesian Journal Of Business Intelligence (Ijubi)*, 4(1), 10. <https://doi.org/10.21927/ijubi.V4i1.1727>
- [7] Putra, R. S., & Yuhandri, Y. (2021). Sistem Pakar Dalam Menganalisis Gangguan Jiwa Menggunakan Metode Certainty Factor. *Jurnal Sistim Informasi Dan Teknologi*, 227–232. <https://doi.org/10.37034/jsisfotek.V3i4.70>
- [8] Bani, A. U., & Asruddin, A. (2022). Pendeteksian Penyakit Mulut Dan Kuku Pada Sapi Dengan Menerapkan Metode Naive Bayes. *Journal Of Computer System And Informatics (Josyc)*, 3(4), 264–268. <https://doi.org/10.47065/Josyc.V3i4.1934>
- [9] Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN. *Jurnal KomtekInfo*, 1–7. <https://doi.org/10.35134/komtekinfo.v10i1.330>
- [9] Gelar Guntara, R. (2023). Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning Yolov7. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(1), 55–60. <https://doi.org/10.47233/Jteksis.V5i1.750>
- [10] Ibnu Daqiqil Id. (2021). *Machine Learning Teori, Studi Kasus Dan Implementasi Menggunakan Python*.
- [11] Nugroho, F., & Prambodo, Y. L. (2022). Pendeteksian Penyakit Limfadenopati Dengan Menerapkan Metode Naive Bayes. *Journal Of Computer System And Informatics (Josyc)*, 3(4), 199–204. <https://doi.org/10.47065/Josyc.V3i4.1997>
- [12] Yuliana, Y., Paradise, P., & Kusriani, K. (2021). Sistem Pakar Diagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Classifier Berbasis Web. *Csrid (Computer Science Research And Its Development Journal)*, 10(3), 127. <https://doi.org/10.22303/Csrid.10.3.2018.127-138>
- [13] Indah Prasasti, M., & Normawati, D. (2023). Sistem Pakar Deteksi Dini Status Stunting Pada Balita Menggunakan Metode Naive Bayes. 7(3), 1276–1286. <https://doi.org/10.30865/Mib.V7i3.6443>
- [14] Widiastuti, N., Hermawan, A., & Avianto, D. (2023b). Implementasi Metode Naive Bayes Untuk Klasifikasi Data Blogger. *Jipi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 8(3), 985–994. <https://doi.org/10.29100/Jipi.V8i3.3713>
- [15] Li, C., Chen, L., Wu, S., Mo, Y., & Chen, L. (2021). In Weather Website. Application Research of Naive Bayes Algorithm Based On Dikk In Weather Website. <https://doi.org/10.1109/Hpcc-Dss-Smartcity-Depends>