

## IMPLEMENTASI METODE NAÏVE BAYES PADA APLIKASI PREDIKSI PENYEBARAN WABAH PENYAKIT ISPA (Studi Kasus: Wilayah Kota Pontianak)

Muchtarul Bari<sup>1</sup>, Sampe Hotlan Sitorus<sup>2</sup>, Uray Ristian<sup>3</sup>

<sup>1,2,3</sup>Jurusan Rekayasa Sistem Komputer, Fakultas MIPA Universitas Tanjungpura

Jalan Prof. Dr. H. Hadari Nawawi Pontianak

Telp./Fax : (0561) 577963

e-mail: <sup>1</sup>muchtarulbari.siskom@gmail.com, <sup>2</sup>sitorus.hotland@gmail.com,

<sup>3</sup>eristian@siskom.untan.ac.id

### Abstrak

*ISPA merupakan penyakit saluran pernapasan pada bagian atas ataupun bawah yang disebabkan oleh virus atau bakteri yang menyerang hidung, trakea (saluran pernapasan) bahkan paru-paru yang menyebabkan fungsi pernapasan menjadi terganggu. Jika tidak segera ditangani, infeksi ini dapat menyebar ke seluruh sistem pernapasan dan menyebabkan tubuh tidak mendapatkan cukup oksigen. Kondisi ini bisa berakibat fatal, bahkan sampai berujung pada kematian. ISPA juga merupakan penyakit yang sangat mudah menular, karena itu diperlukan suatu sistem yang dapat memprediksi penyebaran wabah ISPA sehingga dapat mencegah meluasnya wabah penyakit ISPA. Dengan menerapkan metode Naïve Bayes Classifier pada Aplikasi Prediksi Penyebaran Wabah Penyakit ISPA, aplikasi ini dapat digunakan sebagai alat bantu untuk memprediksi penyebaran wabah penyakit ISPA berdasarkan data-data penyebaran wabah ISPA yang sudah terjadi. Aplikasi dibangun berbasis dekstop. Data-data yang dipakai dalam penelitian merupakan data dari tahun 2012-2016 yang didapat dari Dinas Kesehatan Kota Pontianak. Berdasarkan hasil pengujian didapat persentase keberhasilan sebesar 82,97% dalam memprediksi penyebaran wabah penyakit ISPA untuk tahun 2017.*

**Kata kunci:** ISPA, Naïve Bayes Classifier, Prediksi, Densitas Gauss, dekstop

### 1. PENDAHULUAN

Perkembangan teknologi informasi (IT) sekarang ini sangat pesat dan telah digunakan diberbagai aspek kehidupan baik di bidang pemerintahan, perbankan, sosial budaya, industri, pendidikan, bahkan kesehatan. Perkembangan teknologi informasi (IT) ini sangat membantu pekerjaan manusia, baik dari sisi kemudahan kerja, efisiensi, waktu penyelesaian pekerjaan, dan akurasi hasil pekerjaan. Demikian halnya dibidang kesehatan, perkembangan teknologi informasi (IT) memiliki peranan yang sangat signifikan dalam penyelesaian berbagai permasalahan.

Salah satu jenis penyakit yang mendapatkan perhatian penerapan IT adalah ISPA (Infeksi Saluran Pernapasan Akut) karena penyebarannya cukup luas di

wilayah Indonesia. ISPA merupakan penyakit saluran pernapasan pada bagian atas ataupun bawah yang disebabkan oleh virus atau bakteri yang menyerang hidung, trakea (saluran pernapasan) bahkan paru-paru yang menyebabkan fungsi pernapasan menjadi terganggu[1]. ISPA juga merupakan jenis penyakit yang sangat mudah menular terutama kepada orang-orang yang memiliki kelainan sistem kekebalan tubuh, orang-orang lanjut usia, dan anak-anak yang sistem kekebalan tubuhnya belum terbentuk sepenuhnya.

Untuk mengatasi hal tersebut diperlukan suatu sistem yang dapat memberikan prediksi terkait penyebaran wabah penyakit ISPA. Pada penelitian sebelumnya [2] terkait dengan Sistem Pakar diagnosa ISPA menggunakan metode Sugeno dan Naïve Bayes diperoleh hasil 70 data yang

sama dengan 71 data rekam medis pasien ISPA yang sudah diketahui sebelumnya. Selain itu pada penelitian sistem pakar diagnosis penyakit ISPA berbasis android dengan *Speech Recognition* menggunakan metode *Naive Bayes Classifier* [3] mampu mengklaster jawaban dari user berdasarkan rule yang dibuat sehingga mendapatkan hasil diagnosa.

Uraian di atas melatar belakangi penulis melakukan penelitian untuk membangun sebuah aplikasi yang dapat memprediksi tingkat penyebaran wabah penyakit ISPA menggunakan metode *Naive Bayes*. Hasil penelitian ini diharapkan dapat membantu Pemerintah Kota Pontianak dalam melakukan persiapan untuk menghadapi wabah penyakit ISPA.

## 2. LANDASAN TEORI

### 2.1. Metode *Naive Bayes*

Metode *Naive Bayes* merupakan metode probabilistik pengklasifikasian sederhana berdasarkan *Teorema Bayes* dimana pengklasifikasian dilakukan melalui *training set* sejumlah data secara efisien[4]. Metode *Naive Bayes* menggunakan peluang bersyarat dalam melakukan prediksi.

Sebelum menjelaskan tentang peluang bersyarat pertama perlu diketahui tentang definisi peluang. Peluang dapat didefinisikan sebagai kemungkinan munculnya suatu kejadian. Teori peluang adalah bagian integral dari ilmu statistik, dan merupakan salah satu bagian terpenting dalam teori statistik inferensial. Statistik inferensial berkaitan dengan metode pendugaan dan penarikan kesimpulan terhadap karakteristik suatu populasi berdasarkan informasi yang diperoleh dari sampel. Cara menghitung peluang dapat dilihat pada Persamaan (1) dimana peluang kejadian A dilambangkan dengan  $P(A)$  yaitu jumlah peluang A dari semua anggota titik contoh S.

$$P(A) = \frac{n(A)}{n(S)} \quad (1)$$

Keterangan:

$P(A)$  :Peluang kejadian A  
 $n(A)$  :Jumlah kejadian A  
 $n(S)$  :Jumlah titik contoh S (jumlah semesta)

Peluang bersyarat adalah peluang munculnya suatu kejadian dengan syarat bahwa suatu kejadian lain telah muncul terlebih dahulu. Peluang terjadinya kejadian A dengan syarat bahwa kejadian B telah terjadi dihitung dengan Persamaan (2).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} ; P(B) > 0 \quad (2)$$

Keterangan:

$P(A|B)$  :Peluang kejadian A dengan syarat kejadian B telah terjadi  
 $P(A \cap B)$  :Peluang kejadian A irisan B  
 $P(B)$  :Peluang kejadian B

*Naive Bayes* mengasumsikan bahwa nilai dari sebuah input atribut pada kelas yang diberikan tidak tergantung dengan nilai atribut yang lain[5]. *Teorema Bayes* merupakan teorema yang memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya[6]. Teorema tersebut dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, dimana Persamaan (3) merupakan persamaan *Teorema Bayes*.

$$P(C|F) = \frac{P(F|C) \cdot P(C)}{P(F)} \quad (3)$$

Keterangan:

F :Data dengan class yang belum diketahui  
 C :Hipotesis data F merupakan suatu class spesifik  
 $P(C|F)$  :Probabilitas hipotesis C berdasar kondisi F (posteriori probability)  
 $P(C)$  :Probabilitas hipotesis C (prior probability)  
 $P(F|C)$  :Probabilitas F berdasarkan kondisi pada hipotesis C  
 $P(F)$  :Probabilitas F

Untuk menjelaskan metode *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, *Teorema Bayes* pada Persamaan (3) disesuaikan menjadi Persamaan (4).

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n | C)}{P(F_1 \dots F_n)} \quad (4)$$

Dimana variabel C mempresentasikan kelas, sementara variabel  $F_1 \dots F_n$  mempresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka Persamaan (4) menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (*Posterior*) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, Persamaan (4) dapat pula ditulis secara sederhana seperti pada Persamaan (5).

$$\text{Posterior} = \frac{\text{Prior} \times \text{likelihood}}{\text{evidence}} \quad (5)$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai-nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut persamaan Bayes tersebut dilakukan dengan menjabarkan  $(C|F_1 \dots F_n)$  menggunakan aturan perkalian seperti pada Persamaan (6).

$$\begin{aligned} P(C|F_1 \dots F_n) &= P(C)P(F_1 \dots F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C) \dots P(F_n|C, F_1 \dots F_{n-1}) \end{aligned} \quad (6)$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi (*naif*), bahwa masing-masing petunjuk ( $F_1, F_2 \dots F_n$ ) saling bebas (*independen*) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan seperti pada Persamaan (8).

$$\begin{aligned} P(F_i|F_j) &= \frac{P(F_i \cap F_j)}{P(F_j)} \\ &= \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \end{aligned} \quad (7)$$

Untuk  $i \neq j$ , sehingga

$$P(F_i|C, F_j) = P(F_i|C) \quad (8)$$

Dari Persamaan (8) dapat disimpulkan bahwa asumsi independensi naif tersebut membuat syarat peluang menjadi sederhana, sehingga perhitungan menjadi mungkin untuk dilakukan. Selanjutnya, penjabaran  $P(C|F_1 \dots F_n)$  dapat disederhanakan menjadi Persamaan (9).

$$\begin{aligned} P(C|F_1 \dots F_n) &= P(C)P(F_1|C)P(F_2|C) \dots \\ &= P(C) \prod_{i=1}^n P(F_i|C) \end{aligned} \quad (9)$$

Keterangan:

$$P(C) \prod_{i=1}^n P(F_i|C) \quad \text{: perkalian rating antar atribut}$$

Persamaan (9) merupakan model dari teorema *Naïve Bayes* yang selanjutnya akan digunakan dalam proses prediksi wabah penyakit ISPA. Untuk menangani data numerik yang ada pada data penelitian, metode *Naïve Bayes* menggunakan persamaan *Densitas Gauss* seperti pada Persamaan (10)[7].

$$P(F_i=f_i|C=c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(f_i-\mu_{ij})^2}{2\sigma_{ij}^2}} \quad (10)$$

Keterangan:

- P : Peluang
- $F_i$  : Atribut ke i
- $f_i$  : Nilai atribut ke i
- C : Kelas yang dicari
- $c_j$  : Sub kelas C yang dicari
- $\mu$  : Mean
- $\sigma$  : Standar Deviasi

Untuk menghitung nilai *mean* (nilai rata-rata) maka digunakanlah Persamaan (11).

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad \text{atau} \quad \mu = \frac{x_1+x_2+x_3+\dots+x_n}{n} \quad (11)$$

Keterangan:

$\mu$  : *Mean* (nilai rata-rata)

$x_i$  : Nilai sampel ke- $i$

$n$  : Jumlah sampel

Sedangkan untuk menghitung nilai Standar Deviasi (simpangan baku), digunakanlah Persamaan (12).

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \quad (12)$$

Keterangan:

$\sigma$  : standar deviasi

$\mu$  : *mean* (nilai rata-rata)

$x_i$  : nilai  $x$  ke- $i$

$n$  : jumlah sampel

## 2.2. Sistem Pendukung Keputusan

Sistem Pendukung Keputusan (SPK) adalah sebuah sistem yang mampu memberikan kemampuan pemecahan masalah maupun kemampuan pengkomunikasian untuk masalah dengan kondisi semi terstruktur dan tak terstruktur. Sistem pendukung keputusan bertujuan untuk menyediakan informasi, membimbing, memberikan prediksi serta mengarahkan kepada pengguna agar dapat melakukan pengambilan keputusan dengan lebih baik.

Bonczek, dkk., (1980) dalam buku *Decision Support System And Intelligent Systems* mendefinisikan sistem pendukung keputusan sebagai sistem berbasis komputer yang terdiri dari tiga komponen yang saling berinteraksi yaitu sistem bahasa (mekanisme untuk memberikan komunikasi antara pengguna dan komponen sistem pendukung keputusan lain), sistem pengetahuan (respositori pengetahuan domain masalah yang ada pada sistem pendukung keputusan atau sebagai data atau sebagai prosedur), dan sistem pemrosesan masalah (hubungan antara dua komponen lainnya, terdiri dari satu atau lebih kapabilitas manipulasi masalah umum yang diperlukan untuk pengambilan keputusan).

## 2.3. ISPA

ISPA merupakan singkatan dari Infeksi Saluran Pernapasan Akut. ISPA

meliputi saluran pernapasan bagian atas dan saluran pernapasan bagian bawah. Penyakit ISPA menginfeksi saluran pernapasan selama sampai dengan 14 hari. Yang dimaksud dengan saluran pernapasan adalah organ mulai dari hidung sampai gelembung paru, beserta organ-organ disekitarnya seperti : sinus, ruang telinga tengah dan selaput paru[8].

Program Pemberantasan Penyakit ISPA membagi penyakit ISPA dalam 2 golongan yaitu pneumonia dan yang bukan pneumonia. Pneumonia dibagi atas derajat beratnya penyakit yaitu pneumonia berat dan pneumonia tidak berat. Sebagian besar dari infeksi saluran pernapasan hanya bersifat ringan seperti batuk pilek dan tidak memerlukan pengobatan dengan antibiotik, meskipun demikian anak akan menderita pneumonia bila infeksi paru ini tidak diobati dengan antibiotik dan dapat mengakibatkan kematian. Penyakit batuk pilek seperti rinitis, faringitis, tonsilitis dan penyakit saluran napas bagian atas lainnya digolongkan sebagai bukan pneumonia. Etiologi dari sebagian besar penyakit saluran napas bagian atas ini ialah virus dan tidak dibutuhkan terapi antibiotik. Faringitis oleh kuman *Streptococcus* jarang ditemukan pada balita. Bila ditemukan harus diobati dengan antibiotik penisilin, semua radang telinga akut harus mendapat antibiotik.

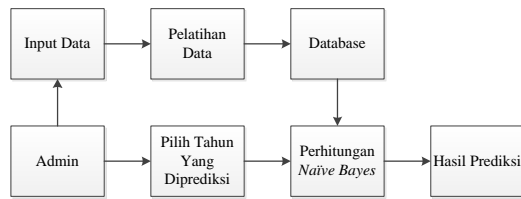
## 3. METODE PENELITIAN

Proses pertama dimulai dengan studi pustaka yang terkait dengan metode *Naïve Bayes*, ISPA, dan teori penunjang lainnya. Selanjutnya dilakukan analisis kebutuhan dan perancangan sistem yang kemudian diintegrasikan menjadi suatu sistem sehingga berfungsi sebagaimana mestinya. Selanjutnya dilakukan pengujian untuk mengetahui kinerja sistem. Setelah dilakukan pengujian, dilakukan analisa terhadap hasil dari pengujian dan diterapkan pada aplikasi prediksi setelah sistem bekerja dengan baik.

### 3.1. Diagram Blok Sistem

Pada perancangan sistem, digunakan diagram blok yang menjelaskan alur sistem.

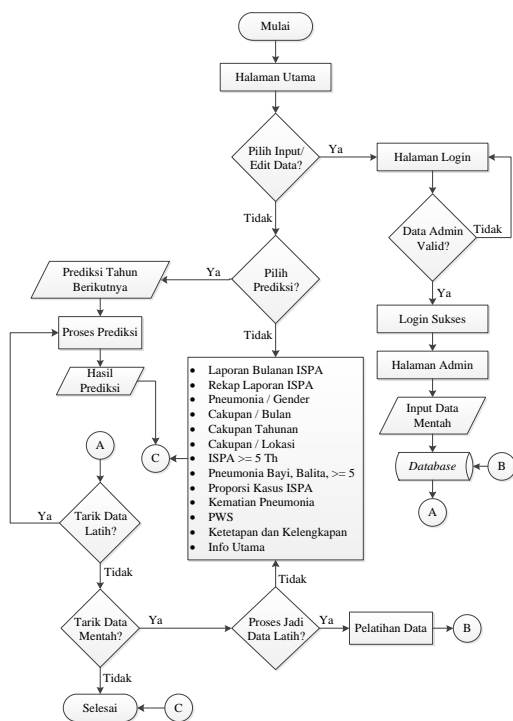
Diagram blok sistem dapat dilihat pada gambar 1.



Gambar 1. Diagram Blok Sistem

Pada tahap pertama admin memasukkan data mentah yang diperoleh dari Dinas Kesehatan Kota Pontianak, kemudian diolah menjadi data latih dengan menggunakan Naive Bayes dan disimpan di *database*. Untuk mendapatkan hasil prediksi dari tahun yang dipilih admin, sistem akan menarik data latih dari *database* untuk dimasukkan ke dalam perhitungan *Naive Bayes*.

### 3.2. Perancangan Flowchart Sistem

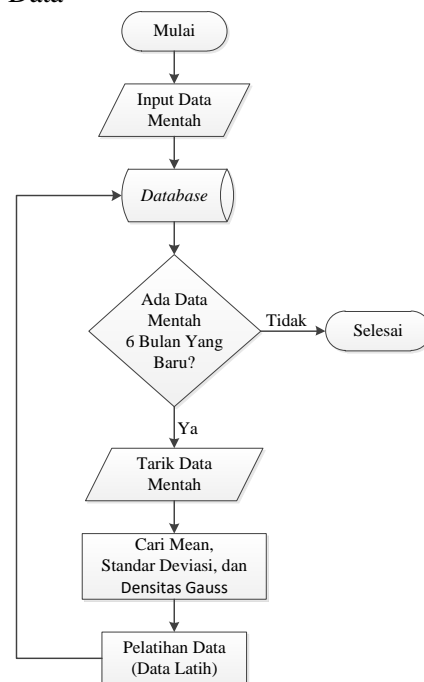


Gambar 2. Flowchart Sistem

Pada dasarnya aplikasi ini hanya diperuntukkan untuk satu pengguna yaitu admin yang ditunjuk untuk menangani pengolahan data ISPA kepala bidang P3PL di Dinas Kesehatan Kota Pontianak. Pada Gambar 2 dapat dilihat bahwa admin dapat

mengakses data-data pada aplikasi sedangkan pengguna selain admin hanya dapat melihat rekapan data saja. Data mentah pada *Flowchart* Sistem merupakan data awal dari tahun 2010 – 2016 yang didapatkan dari Dinas Kesehatan Kota Pontianak. Halaman *login* akan muncul apabila pengguna ingin memasukkan data baru ataupun mengubah data yang sudah ada. Proses prediksi merupakan proses perhitungan menggunakan *Naive Bayes* untuk mencari data tahun berikutnya.

### 3.3. Perancangan Flowchart Pelatihan Data

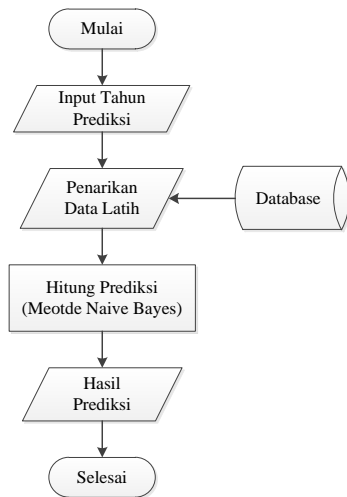


Gambar 3. Flowchart Pelatihan Data

Dari Gambar 3 dapat dilihat *database* memiliki dua jenis data yaitu data mentah dan data latih. Data mentah merupakan data asli yang langsung di masukkan oleh admin, sedangkan data latih merupakan data mentah yang sudah di proses agar dapat digunakan dalam perhitungan *Naive Bayes*. Sebelum menjadi data latih, data mentah di proses dengan cara mencari nilai *mean* dan standar deviasinya, yang kemudian dilanjutkan menggunakan persamaan *Densitas Gauss*. Ketika admin memasukkan dan menyimpan data mentah ke *database*, data tersebut tidak langsung digunakan dalam proses pelatihan data,

tetapi dikumpulkan sampai data tersebut berjumlah 6 bulan.

### 3.4. Perancangan *Flowchart* Prediksi Data



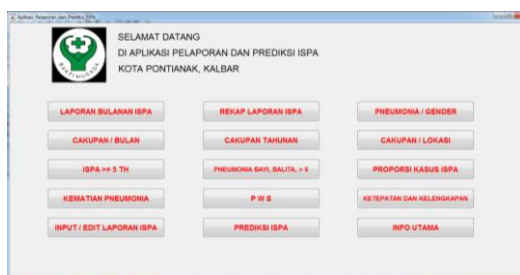
Gambar 4. *Flowchart* Prediksi Data

Pada Gambar 4 dapat dilihat untuk melakukan prediksi, sistem memerlukan data latih yang merupakan hasil pemrosesan dari data mentah. Data-data latih kemudian dimasukkan ke dalam perhitungan *Naive Bayes* yang kemudian didapatkanlah hasil berupa data prediksi.

## 4. HASIL DAN PEMBAHASAN

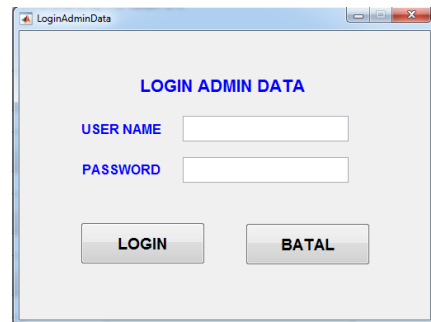
### 4.1. Tampilan Aplikasi Pelaporan dan Prediksi ISPA

Saat membuka aplikasi, halaman yang akan dijumpai pertama oleh pengguna merupakan halaman utama aplikasi. Pada halaman ini terdapat berbagai *menu-menu* yang dapat di akses pengguna kecuali *menu* “INPUT/EDIT LAPORAN ISPA” yang membutuhkan akses admin untuk mengaksesnya. Tampilan halaman utama aplikasi dapat dilihat pada Gambar 5.



Gambar 5. Tampilan Halaman Utama Aplikasi

Pada saat pengguna memilih *menu* “INPUT/EDIT LAPORAN ISPA” maka akan muncul *pop up window* yang meminta *password* admin seperti pada Gambar 6.

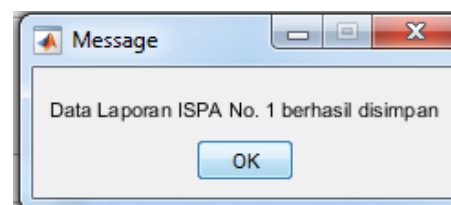


Gambar 6. Tampilan *Login* Admin

Setelah pengguna berhasil *login* sebagai admin maka jendela untuk melakukan input laporan akan keluar, seperti pada Gambar 7. Pada jendela inilah pengguna memasukkan berbagai data yang didapatkan dari lapangan. Untuk menyimpan semua data yang sudah dimasukkan pengguna harus menekan tombol “Simpan Laporan”. Saat data berhasil disimpan maka akan muncul pesan kalau data berhasil disimpan seperti pada Gambar 8.



Gambar 7. Tampilan *Menu* Input Laporan



Gambar 8. Tampilan Pemberitahuan Data Sudah Tersimpan

Apabila pengguna ingin menambah data yang akan dimasukkan, pengguna dapat menekan tombol “Tambah Laporan”. Maka jendela *input* laporan akan kembali dikosongkan seperti pada Gambar 9.



$$P(\text{Tinggi}|\text{Januari}) = \frac{0}{24} = 0$$

- 4) Cari data latih kasus Pneumonia  
Untuk data latih kasus perhitungan menggunakan persamaan (10) dikarenakan data yang bersifat kontinyu.

$$P(\text{Rendah}|0) = \frac{1}{\sqrt{2(3,14)9,08}} e^{-\frac{(0-3,509)^2}{2(9,08)^2}} = 0,123$$

$$P(\text{Sedang}|0) = \frac{1}{\sqrt{2(3,14)24,145}} e^{-\frac{(0-35,5)^2}{2(24,145)^2}} = 0,028$$

$$P(\text{Tinggi}|0) = \frac{1}{\sqrt{2(3,14)21,937}} e^{-\frac{(0-30,5)^2}{2(21,937)^2}} = 0,032$$

Dalam penelitian ini jangka waktu yang digunakan dalam prediksi yaitu enam bulan sehingga untuk memprediksi bulan Januari maka digunakanlah data bulan Juli tahun sebelumnya. Setelah semua data yang diperlukan diperoleh maka proses selanjutnya adalah melakukan perhitungan untuk memprediksi menggunakan metode Naïve Bayes. Untuk menghitung nilai peluang prediksi digunakanlah persamaan (9). Berikut adalah contoh perhitungan prediksi penderita pneumonia di puskesmas Khatulistiwa pada bulan Januari tahun 2017 menggunakan metode Naïve Bayes.

$$P(\text{Rendah}) = (0,045 \times 0,093 \times 0,123) \times 0,96 = 0,000395$$

$$P(\text{Sedang}) = (0,000 \times 0,143 \times 0,028) \times 0,02 = 0$$

$$P(\text{Tinggi}) = (0,000 \times 0,000 \times 0,032) \times 0,02 = 0$$

Dari perhitungan diatas diketahui peluang rendah lebih besar dari peluang yang lain sehingga dapat disimpulkan bahwa banyaknya penderita pneumonia di puskesmas Khatulistiwa pada bulan Januari tahun 2017 kemungkinan ialah rendah dibandingkan jumlah penderita tahunan.

Berikut contoh yang lain dari perhitungan prediksi menggunakan metode Naïve Bayes. Perhitungan yang dilakukan

yaitu untuk mencari data latih serta prediksi banyaknya penderita bukan Pneumonia di puskesmas Khatulistiwa pada bulan Januari tahun 2017.

- 1) Cari data latih kelas bukan Pneumonia

$$P(\text{Rendah}) = \frac{\text{Banyaknya data rendah}}{\text{Jumlah data}} = \frac{916}{1242} = 0,74$$

$$P(\text{Sedang}) = \frac{\text{Banyaknya data sedang}}{\text{Jumlah data}} = \frac{263}{1242} = 0,21$$

$$P(\text{Tinggi}) = \frac{\text{Banyaknya data tinggi}}{\text{Jumlah data}} = \frac{63}{1242} = 0,05$$

- 2) Cari data latih puskesmas Khatulistiwa bukan Pneumonia

$$P(\text{Rendah}|Khatulistiwa) = \frac{46}{916} = 0,05$$

$$P(\text{Sedang}|Khatulistiwa) = \frac{4}{263} = 0,015$$

$$P(\text{Tinggi}|Khatulistiwa) = \frac{4}{63} = 0,063$$

- 3) Cari data latih bulan Januari bukan Pneumonia

$$P(\text{Rendah}|Januari) = \frac{104}{916} = 0,114$$

$$P(\text{Sedang}|Januari) = \frac{11}{263} = 0,042$$

$$P(\text{Tinggi}|Januari) = \frac{0}{63} = 0$$

- 4) Cari data latih kasus bukan Pneumonia

$$P(\text{Rendah}|0) = \frac{1}{\sqrt{2(3,14)157,557}} e^{-\frac{(224-216,198)^2}{2(157,557)^2}} = 0,032$$

$$P(\text{Sedang}|0) = \frac{1}{\sqrt{2(3,14)203,125}} e^{-\frac{(224-366,087)^2}{2(203,125)^2}} = 0,022$$

$$P(\text{Tinggi}|0) = \frac{1}{\sqrt{2(3,14)280,445}} e^{-\frac{(224-579,016)^2}{2(280,445)^2}} = 0,011$$



5) Prediksi peluang bukan Pneumonia

$$P(\text{Rendah}) = (0,05 \times 0,114 \times 0,032) \times 0,74 = 0,0000937$$

$$P(\text{Sedang}) = (0,015 \times 0,042 \times 0,022) \times 0,21 = 0,0000038$$

$$P(\text{Tinggi}) = (0,063 \times 0,000 \times 0,011) \times 0,05 = 0,0000027$$

Peluang rendah lebih besar dari peluang sedang dan peluang tinggi sehingga bisa disimpulkan banyaknya penderita bukan Pneumonia di puskesmas Khatulistiwa pada bulan Januari tahun 2017 adalah rendah.

4.3. Pengujian Sistem Aplikasi Prediksi

Pengujian Sistem dilakukan untuk mengetahui presentasi terbaik pada aplikasi ini. Pada sistem akan diuji hasil prediksi untuk bulan Januari tahun 2017. Hasil yang didapatkan dari pengujian pada sistem dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Tabel Hasil Pengujian Pneumonia Bulan Januari Tahun 2017

Puskesmas	Data Real	Output Aplikasi	Validasi
Khatulistiwa	Rendah	Rendah	Benar
St.Hilir	Rendah	Rendah	Benar
St.Tengah	Rendah	Rendah	Benar
St.Hulu	Rendah	Rendah	Benar
T.biru	Rendah	Rendah	Benar
Tj.Hulu	Rendah	Rendah	Benar
Saigon	Rendah	Rendah	Benar
Bj.Serasan	Rendah	Rendah	Benar
Prt.Mayor	Rendah	Rendah	Benar
T.Sampit	Rendah	Rendah	Benar
Kp.Dalam	Rendah	Rendah	Benar
Kp.Bangka	Rendah	Rendah	Benar
Paris II	Rendah	Rendah	Benar
Purnama	Rendah	Rendah	Benar
Gg. Sehat	Tinggi	Rendah	Salah
Kp. Bali	Rendah	Rendah	Benar
Alianyang	Rendah	Rendah	Benar
Pal Tiga	Rendah	Rendah	Benar
Karya Mulya	Rendah	Rendah	Benar
Perumnas I	Tinggi	Rendah	Salah
Perumnas II	Rendah	Rendah	Benar
Komyos	Rendah	Rendah	Benar
Pal Lima	Rendah	Rendah	Benar

Tabel 2. Tabel Hasil Pengujian Bukan Pneumonia Bulan Januari Tahun 2017

Puskesmas	Data Real	Output Aplikasi	Validasi
Khatulistiwa	Rendah	Rendah	Benar
St.Hilir	Tinggi	Tinggi	Benar
St.Tengah	Rendah	Rendah	Benar
St.Hulu	Sedang	Rendah	Salah
T.biru	Rendah	Rendah	Benar
Tj.Hulu	Rendah	Rendah	Benar
Saigon	Sedang	Rendah	Salah
Bj.Serasan	Rendah	Rendah	Benar
Prt.Mayor	Rendah	Rendah	Benar
T.Sampit	Rendah	Rendah	Benar
Kp.Dalam	Sedang	Rendah	Salah
Kp.Bangka	Sedang	Rendah	Salah
Paris II	Rendah	Rendah	Benar
Purnama	Sedang	Rendah	Salah
Gg. Sehat	Rendah	Rendah	Benar
Kp. Bali	Rendah	Rendah	Benar
Alianyang	Tinggi	Rendah	Salah
Pal Tiga	Sedang	Rendah	Salah
Karya Mulya	Sedang	Rendah	Salah
Perumnas I	Rendah	Rendah	Benar
Perumnas II	Rendah	Rendah	Benar
Komyos	Rendah	Rendah	Benar
Pal Lima	Sedang	Sedang	Benar

4.4. Pembahasan

Dari penelitian ini diketahui bahwa dalam mengimplementasikan metode Naïve Bayes data-data harus terlebih dahulu dikelompokkan. Data-data dikelompokkan ke dalam tiga kelompok yaitu tinggi, sedang, dan rendah. Pengelompokkan data dilakukan dengan cara mencari nilai tertinggi tiap kategori dalam kurun waktu 6 bulan untuk tiap tahun. Nilai tertinggi tersebut kemudian digunakan untuk menentukan kelompok dari tiap data (Fi) dalam kurun waktu 6 bulan tersebut dengan aturan sebagai berikut.

$$\begin{aligned} \text{Tinggi} & : F_i \geq 67\% \\ \text{Sedang} & : 67\% > F_i \geq 33\% \\ \text{Rendah} & : F_i < 33\% \end{aligned}$$

Dalam melakukan prediksi aplikasi ini menggunakan data uji 6 bulan terakhir, pada pengujian ini data uji yang digunakan yaitu dari bulan Juli sampai bulan Desember tahun 2016. Pengujian dilakukan dengan cara memprediksi 6 bulan kedepan yaitu bulan Januari sampai bulan Juni tahun 2017. Dari pengujian aplikasi didapatkanlah

persentase keberhasilan sebesar 82,97%. Berikut adalah proses perhitungan persentase keberhasilan pengujian.

$$\begin{aligned}\% \text{ Keberhasilan} &= \frac{\text{Jumlah output benar}}{\text{Jumlah data real}} \\ &= \frac{229}{276} \times 100\% = 82,97\%\end{aligned}$$

$$\begin{aligned}\% \text{ Kegagalan} &= \frac{\text{Jumlah output salah}}{\text{Jumlah data real}} \\ &= \frac{47}{276} \times 100\% = 17,03\%\end{aligned}$$

## 5. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Berdasarkan pada proses pembuatan aplikasi prediksi wabah penyakit ISPA ini didapatkanlah beberapa kesimpulan sebagai berikut:

1. Untuk mengimplementasikan metode *Naive Bayes* data-data harus terlebih dahulu dikonversi menjadi bentuk diskrit dengan cara dikelompok-kelompokkan.
2. Dengan menerapkan metode *Naive Bayes* pada aplikasi prediksi maka didapatkanlah tingkat akurasi sebesar 82,97%.

### 6.2. Saran

Berdasarkan pada permasalahan-permasalahan yang ditemukan dalam proses pembuatan aplikasi prediksi wabah penyakit ISPA ini terdapat beberapa saran untuk penelitian lebih lanjut, antara lain:

1. Diperlukan lebih atribut data seperti atribut cuaca, polusi udara, dan lain-lain untuk melakukan prediksi yang lebih baik.
2. Disarankan dalam menggunakan metode *Naive Bayes* data-data pada bagian kategori harus bersifat diskrit atau sudah memiliki kelompok-kelompok

*balita*. Jakarta: Departemen Kesehatan RI.

- [2] Pramudyo, E. R., Bettiza, M., & Sallu, S. (2016). *Sistem Pakar Diagnosa Penyakit Infeksi Saluran Pernafasan Akut (ISPA) Dengan Menggunakan Metode Sugeno Dan Naive Bayes (Studi Kasus Rumah Sakit Otorita Batam)*. Jurnal Elektronik Tugas Akhir Mahasiswa.
- [3] Marlina, M., Saputra, W., Mulyadi, B., Hayati, B., & Jaroji. (2017, Mei). *Aplikasi Sistem Pakar Diagnosis Penyakit ISPA Berbasis Speech Recognition Menggunakan Metode Naive Bayes Classifier*. Jurnal Teknologi Informasi & Komunikasi Digital Zone, 8, 58-70.
- [4] Hadiyani, E. P. (2013). *Rancang Bangun Sistem Pendukung Keputusan Untuk Pemilihan Anggota Terbaik AIESEC Surabaya Dengan Menggunakan Metode Naive Bayes*. Fakultas Sains dan Teknologi.
- [5] Pang-Ning, T., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Pearson Education. Boston.
- [6] Bustami. (2014, Januari). *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi*. Jurnal Informatika, 8(1), 884-898.
- [7] Aminudin, M. (2011). *Peramalan Cuaca Kota Surabaya Tahun 2011 Menggunakan Metode Moving Average Dan Klasifikasi Naive Bayes*. Jurnal Data Mining, 6.
- [8] Rasmaliah. (2004). *Infeksi Saluran Pernafasan Akut (ISPA) Dan Penanggulangannya*. USU digital library(6), 1-8.

## DAFTAR PUSTAKA

- [1] Depkes RI. (2002). *Pedoman pemberantasan penyakit infeksi saluran pernapasan akut untuk penanggulangan pneumonia pada*