

# KLASIFIKASI PEYAKIT PARU-PARU MENGGUNAKAN METODE NAIVE BAYES STUDI KASUS RS.PKU MUHAMMADIYAH UJUNG PANGKAH GRESIK

**Rizki Wana Dwi Putra<sup>1</sup>, Eko Prasetyo<sup>2</sup>, Umi Chotijah<sup>3</sup>**

<sup>1</sup> Mahasiswa Teknik Informatika Universitas Muhammadiyah Gresik

<sup>2</sup> Dosen Universitas Muhammadiyah Gresik

Fakultas Teknik, Universitas Muhammadiyah Gresik

Jl. Sumatra 101 Gresik Kota Baru (GKB), Randuagung, 61121 Telp. (031) 3951414 ,

Faks. (0561) 740186

Email: [riskiwana27@gmail.com](mailto:riskiwana27@gmail.com)<sup>1</sup>, [Eko1979@yahoo.com](mailto:Eko1979@yahoo.com)<sup>2</sup>, [umi.chotijah@umg.ac.id](mailto:umi.chotijah@umg.ac.id)<sup>3</sup>

## INTISARI

Penyakit Paru-Paru adalah penyakit yang berhubungan dengan sistem pernapasan pada manusia yang dapat berakibat buruk apabila tidak segera di tangani penderita akan merasakan gejala penyakit paru-paru apabila penderita mengalami keluhan seperti nyeri dada yang bertambah batuk, batuk kering serta disertai dahak, demam, nafsu makan yang berkurang. Penelitian ini menerapkan teknik *data mining* dengan menggunakan metode klasifikasi *Naive Bayes* untuk mengklasifikasikan jenis-jenis penyakit paru-paru. Atribut yang digunakan adalah umur, berat badan, sistole, diastole, batuk dan sesak. Data yang digunakan adalah data penderita penyakit paru-paru di RS. PKU Muhammadiyah Ujung Pangkah Gresik tahun 2015, 2016, 2017, 2018 sebanyak 160 data. Pengujian sistem dilakukan dengan tiga kali percobaan. Akurasi terbaik yang didapat pada pengujian dengan diperoleh 82%.

**Keywords :** Health Center RS PKU Muhammadiyah Ujung Pangkah Gresik, *Data Mining, Classification, Naive Bayes*

## I. PENDAHULUAN

Penyakit paru-paru merupakan penyakit yang berhubungan dengan sistem pernapasan pada manusia yang dapat berakibat buruk apabila tidak segera ditangani. Penyakit ini dapat menyebabkan kematian bagi penderitanya. Pasien penyakit paru-paru pada gejala awal tidak mengetahui bahwa dirinya mengidap penyakit paru-paru.

Menurut Badan Statistik kesehatan Dunia yang dikeluarkan oleh WHO tahun 2008, PPOK diperkirakan menjadi penyebab kematian ketiga pada tahun 2030 dan angka kecacatan karena penyakit PPOK akan meningkat dari urutan ke-9 menjadi urutan ke 5 pada tahun 2020 (MERIEADELAIDE, 2013). Pada tahun 2015 PPOK merupakan penyebab kematian ke -4 di dunia. Hasil RISKESDAS (2013) menyatakan bahwa prevalensi PPOK di

Indonesia sebesar 3,7%. Prevalensi PPOK di Indonesia tidak terlalu tinggi. Prevalensi penyakit paru-paru diperkirakan akan meningkat sehubungan dengan penyakit infeksi ke penyakit *degenerative* serta kebiasaan merokok dan polusi udara.

Sistem klasifikasi ini dapat membantu seorang petugas medis di rumah sakit dalam mengklasifikasikan penyakit paru-paru. Tujuan dari pembuatan sistem ini untuk mengetahui jenis-jenis penyakit paru-paru dimana data yang dimasukan adalah data dari pasien sendiri. Maka dari itu diperlukan sebuah sistem yang dapat membantu dokter dan perawat dalam mengklasifikasi jenis-jenis penyakit paru-paru.

## 2. LANDASAN TEORI

### 2.1 Penyakit Paru-Paru

Paru-paru merupakan salah satu organ pernapasan yang berfungsi untuk melakukan repirasi, yaitu mengubah gas (CO<sub>2</sub>) menjadi gas oksigen (O<sub>2</sub>) dan air (H<sub>2</sub>O). sebelum sampai ke paru-paru, udara yang dihirup manusia akan melewati hidung, pangkal laring, kemudian menuju ke kedua bronki utama (bronkus) dan akan disalurkan ke bronki yang paling kecil (bronkioli), selanjutnya udara dimasukan ke dalam jutaan kantong udara (alveoli) yang berada dalam paru-paru. Paru-paru memiliki lapisan pelindung (pleura) yang juga berfungsi untuk membantu kontraksi dalam rongga dada. Penyakit paru-paru adalah kondisi paru-paru dimana terjadi peradangan atau pengumpulan cairan (darah atau nana) atau masuknya bakteri, virus atau jamur ke dalam

paru-pari yang kemudian menyebabkan paru-paru tidak berfungsi dengan baik (normal).

### 2.2 Jenis-Jenis Penyakit Paru-Paru

#### 2.2.1 Tuberculosis

Penyakit *Tuberculosis* (TB) merupakan penyakit menular yang menyebar melalui udara ketika seseorang dengan infeksi TB aktif batuk, bersin, menyebarkan butiran ludah mereka melalui udara. Orang dengan sistem kekebalan tubuh lemah memiliki resiko yang jauh lebih besar untuk terjangkit TB.

#### 2.2.2 Penyakit Paru Obstruktif Kronis (PPOK)

PPOK merupakan istilah yang digunakan untuk menggambarkan penyakit paru yang sangat buruk, seperti asma dalam jangka waktu panjang (tidak ada perubahan), bronchitis menahun/kronis, emfisema (kondisi udara di paru-paru (alveoli) mengalami kerusakan yang sangat buruk).

### 2.3 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 bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar. Menurut (Kusrini & Emha Taufiq Lutfi, 2009) Hal penting yang terkait didalam *data mining* adalah :

## 2.4 Pengelompokan Data Mining

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

- Deskripsi
- Estimasi
- Prediksi
- Klasifikasi
- Clustering
- Asosiasi

## 2.5 Naive Bayes

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

Dalam bentuk praktiknya parameter untuk perhitungan model *Naive Bayes* menggunakan metode *maximum likelihood* atau kemiripan tertinggi.

Prediksi *naive Bayes* didasarkan pada teorema Bayes dengan formula untuk klasifikasi sebagai berikut (Prasetyo, 2012) :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i | Y)}{P(X)}$$

Sedangkan *Naive Bayes* dengan fitur kontinu memiliki formula :

$$P(X|Y) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(x-\mu)^2}{2\sigma^2}$$

Keterangan :

$P(Y | X)$  = probabilitas data dengan vektor  $X$  pada kelas  $Y$ .

$P(Y)$  = probabilitas awal kelas  $Y$

$\prod_{i=1}^q P(X_i | Y)$  = probabilitas independen kelas  $Y$  dari semua fitur dalam vektor  $X$

$\mu$  = Mean atau nilai rata-rata dari atribut dengan fitur kontinu

$\sigma$  = Deviasi standar

## 3. PERANCANGAN SISTEM

### 3.1 Analisis Sistem

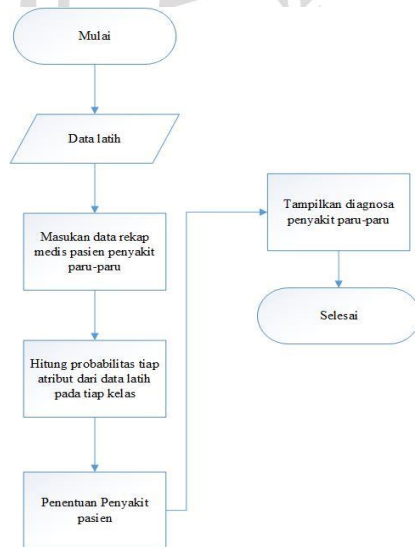
Penyakit Paru-Paru adalah penyakit yang berhubungan dengan sistem pernapasan pada manusia yang dapat berakibat buruk apabila tidak segera di tangani penderita akan merasakan gejala penyakit paru-paru apabila penderita mengalami keluhan seperti nyeri dada yang bertambah batuk, batuk kering serta disertai dahak, demam, nafsu makan yang berkurang dari data di Rumah Sakit PKU Muhammadiyah Ujung Pangkah Gresik peningkatan penderita penyakit paru-paru setiap tahunnya mengalami kenaikan. Dari Data WHO menunjukkan hampir 3,7% penduduk Indonesia mengalami penyakit paru-paru. Tingkat Keparahan penyakit paru-paru terbagi beberapa status yaitu Tuberculosis (TBC) dan Penyakit Paru Obstruktif Kronis (PPOK).

### 3.2 Hasil Analisis

Hasil analisa yang di dapat adalah sistem klasifikasi ini diharapkan dapat membantu tenaga medis dan pasien untuk lebih akurat dan lebih cepat dalam mendiagnosa jenis penyakit paru-paru yang diderita pasien karena *variabel* yang digunakan untuk mendiagnosa tidak

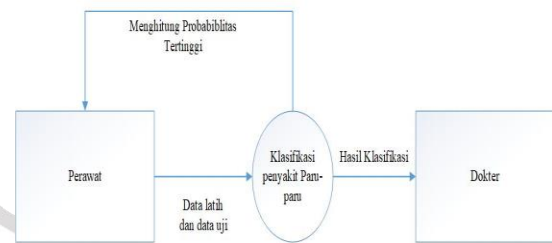
sebanyak yang dibutuhkan dalam uji coba. Pembuatan aplikasi data mining ini menggunakan metode *Neive Bayes* diperlukan data pembelajaran, data berupa umur, berat badan, diastole, sitole, batuk dan sesak.

Sistem yang dibangun merupakan aplikasi atau *tool* klasifikasi pasien penderita Penyakit paru-paru menggunakan teknik data mining dengan metode *Neive Bayes*. Sistem ini menghasilkan nilai keluaran berupa penyakit tuberculosis dan penyakit paru obstruktif kronis. Terdapat beberapa atribut yang dibutuhkan untuk mengklasifikasikan penderita penyakit paru-paru diantaranya umur, berat badan, diastole, sitole, batuk dan sesak. Pada Gambar 3.1 akan menjelaskan bagaimana proses alur sistem pada aplikasi klasifikasi penyakit paru-paru.



**Gambar 3.1** Flowchart sistem klasifikasi penyakit paru-paru

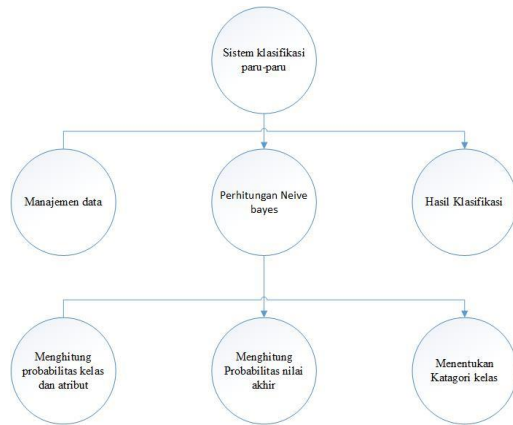
### 3.3 Diagram Konteks



**Gambar 3.2** Diagram Konteks Sistem Klasifikasi Penyakit paru-paru

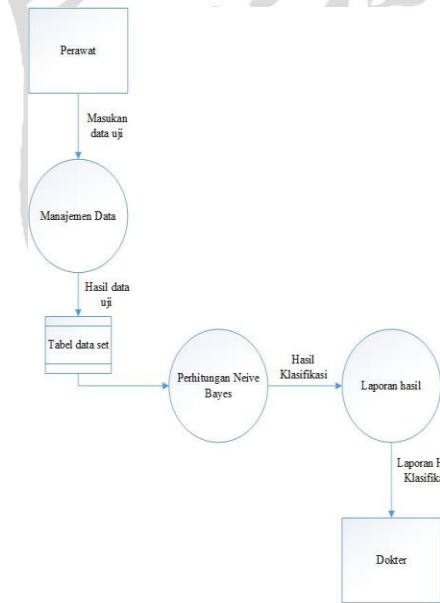
Penjelasan pada gambar 3.2 adalah *entity* dalam sistem ini adalah perawat dan dokter. Perawat memasukkan data pasien penyakit paru-paru sebagai data latih yang terdiri dari umur, berat badan, tekanan, darah sistole, tekanan darah diastole, sesak dan batuk. Data tersebut digunakan sebagai data latih atau *data training* yang akan dihitung probabilitasnya. Keluaran dari sistem untuk perawat adalah menentukan kelas dengan probabilitas akhir terbesar berdasarkan data yang telah dimasukkan. Setelah nilai kelas probabilitas ditemukan, maka perawat memasukkan data uji pasien penyakit paru-paru untuk masuk prediksi kelas sesuai kelas probabilitas terbesar diklasifikasikan dan dokter akan menerima hasil klasifikasi pasien.

### 3.4 Diagram Berjenjang



**Gambar 3.3** Diagram Berjenjang pada sistem klasifikasi penyakit paru-paru

### 3.5 DFD Level 0



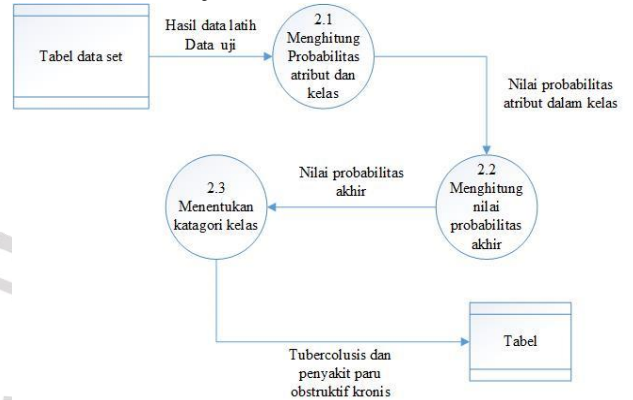
**Gambar 3.4** Diagram level 0

### 3.6 DFD Level 1

**Gambar 3.5** DFD Level 1

1. Menghitung probabilitas atribut tiap kelas yang terdiri dari atribut umur, berat badan, sitole, diastole, sesak dan batuk.

2. Menghitung nilai probabilitas akhir dari data uji.



3. Setelah probabilitas akhir akan ditentukan katagori kelas berupa tuberculosis dan penyakit paru obstruktif kronis

## 4. HASIL DAN ANALISA

### 4.1Keberhasilan

Terdapat 3 variasi percobaan dimana ketiganya memiliki jumlah data latih dan data uji yang berbeda-beda pula. Tabel 4.4 menunjukkan hasil perbandingan pengujian dari percobaan pertama sampai percobaan ketiga.

No	Data latih	Data uji	Jumlah benar	Jumlah salah	Akurasi
Ke 1	120	40	32	8	80%
Ke 2	110	50	41	9	82%
Ke 3	100	60	47	13	78.3%

Berdasarkan pengujian yang telah dilakukan didapat hasil bahwa pengujian kedua menghasilkan akurasi terbaik dibandingkan

dengan jenis pengujian lain dengan nilai akurasi tertinggi mencapai 82%.

## 4.2 Pengujian Satu Data

Dilakukan satu pengujian data uji, contoh kasus dari permasalahan sebagai berikut:

Seorang pasien memiliki usia = 60 berat badan = 80, sistole = 120, diastole = 70, batuk = Ya sesak = Tidak Bagaimana klasifikasi tingkat keparahan hipertensi pasien tersebut?

Diketahui :

Usia = 60                      Sistole = 120  
Batuk = Ya  
BB = 80                      Diastole = 70  
Sesak = Tidak

Masukkan data tersebut pada menu prediksi , kemudian uji klasifikasi tersebut akan diproses sesuai dengan perhitungan *Naive Bayes*.

Edit Data Latih Pasien [Kembali](#)

Umur Pasien

Berat Badan

Sistol

Diastol

Batuk

Sesak

Kelas

[Simpan](#)

Gambar 4.1 Pengujian Satu Data

Perhitungan dan Hasil Klasifikasi [Klasifikasi Semua Data](#)

Show  entries Search

No	Usia	Berat Badan	Sistol	Diastol	Batuk	Sesak	Kelas Awal	Kelas Sistem
1	60	88	120	70	ya	tidak	TBC	TBC

Gambar 4.2 Hasil

## 5. PENUTUP

### 5.1 Kesimpulan

Hasil yang didapatkan dari penelitian ini dapat disimpulkan sebagai berikut.

1. Sistem pengklasifikasian penyakit paru-paru dapat membantu tenaga medis dalam melakukan pengklasifikasian penyakit paru-paru berdasarkan katagori ke dalam 2 kelas.
2. Dalam peneitian ini menggunakan tiga kali pengujian. Pengujian kedua memiliki akurasi terbaik di bandingkan jenis pengujian yang lain dengan data latih 110 dan data uji 50 diperoleh akurasi 82% dan laju *error* 18%

### 5.2 Saran

Saran yang akan diberikan untuk mengembangkan penelitian ini adalah Proses klasifikasi *Neive Bayes* sangat bergantung dari data latih yang digunakan, penggunaan data latih sebaiknya menggunakan proporsi data yang sama dari setiap kelasnya serta memastikan kelas dari data latih tersebut dengan benar

## DAFTAR PUSTAKA

- Baridwan. (2010). *DEFINISI SISTEM*. Baridwan.
- DAMAYANTI, N. (2017). Klasifikasi penyakit paru dengan metode ANN. *NURITA DAMAYANTI, 1*.
- Gonunesco. (2011). *Langkah Proses Dalam Data Mining*. Bandung.
- Jogyianto, H. (2005). Analisis dan Desain: Sistem Informasi Pendekatan Terstruktur Teori dan Praktek. *ANDI*.
- Kusrini, & Emha Taufiq Lutfi. (2009). *algoritma data mining*. yogyakarta: andi offset.

MERIEADELAIDE. (2013). HASIL RISKESDAS.

*MERIEADELAIDE, 1.*

Prasetyo, E. (2014). Data Mining - Mengolah

Data Menjad Informasi menggunakan

Matlab. *Yogyakarta: Andi Offset.*

Prasetyo, E. (2014). Data Mining Konsep dan

Aplikasi Menggunakan Matlab.

*Yogyakarta : Andi Offset.*

Sistem Klasifikasi Penyakit Asma Menggunakan

Algoritma Neive bayes. (2017). *Muhdi*

*Abdullah, 1.*







