

PENERAPAN DATA MINING METODE NAIVE BAYES UNTUK DIAGNOSA PENYAKIT PNEUMONIA PADA BALITA

Muhtajuddin Danny¹, Donny Maulana², Taufik Hidayat

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa

¹utat@pelitabangsa.ac.id, ²donny.maulana@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

Abstrak

Terdapat beberapa kasus kematian pada anak khususnya balita yang meninggal akibat terkena penyakit *pneumonia*, menurut *World Health Organization (WHO)* memperkirakan kematian balita karena *pneumonia* di seluruh dunia sebesar 15%, dan pada tahun 2015 diperkirakan 922.000 kematian balita yang disebabkan oleh *pneumonia*. Saat ini teknologi dapat memberikan informasi yang cepat dan akurat khususnya di lingkungan kesehatan baik untuk tim kesehatan, dokter, perawat bahkan untuk pasien sendiri agar lebih mudah mengontrol kesehatan mereka. *Data mining* berhubungan dengan pencarian data untuk menemukan pola atau pengetahuan dari data keseluruhan, kumpulan data yang besar dapat menghasilkan sebuah data yang hasilnya dapat memberikan informasi pengetahuan yang baru. Pada penelitian ini akan dibahas tentang perancangan *data mining* menggunakan algoritma *Naïve Bayes Classifier* untuk menghitung probabilitas kemungkinan seorang pasien dengan gejala-gejala tertentu apakah mengidap penyakit *pneumonia* atau tidak sehingga dapat memberikan kontribusi kepada tim medis di lingkungan kesehatan untuk mengetahui dan menindak lanjut pasien yang terkena penyakit *pneumonia*.

Kata kunci: *Data Mining, Naïve Bayes Classifier, Pneumonia, Balita*

Abstract

There are several cases of death in children, especially children under five who die from pneumonia, according to the World Health Organization (WHO) estimates under-five deaths due to pneumonia world wide by 15%, and in 2015 an estimated 922,000 under-five deaths caused by pneumonia. Nowadays technology can provide fast and accurate information, especially in health environments for health teams, doctors, nurses and even patients themselves so that they can control their health more easily. Data mining is related to the search for data to find patterns or knowledge from the overall data, large data sets can produce a data that results can provide new knowledge information. In this study will be discussed about data mining design using the Naïve Bayes Classifier algorithm to calculate the probability of the possibility of a patient with certain symptoms whether or not pneumonia can contribute to the medical team in the health environment to know and follow-up patients affected by the disease pneumonia.

Keyword: Data mining, Naïve Bayes Classifier, Pneumonia, Toddlers

1. Pendahuluan

Dengan perkembangan teknologi yang semakin pesat di semua bidang baik dari sektor pemerintahan, pendidikan, pertanian, perikanan, dan khususnya di lingkungan kesehatan. Saat ini teknologi dapat memberikan informasi yang cepat dan akurat baik untuk tim kesehatan, dokter, perawat bahkan untuk pasien sendiri agar lebih mudah mengontrol kesehatan mereka. Berbagai bidang di kesehatan menghasilkan data dalam jumlah yang besar tiap tahunnya. Data tersebut rata-rata tidak bisa memberikan informasi secara cepat dan langsung sehingga informasi yang didapatkan memakan waktu yang lama (Tarigan, Rini, & Puspita, 2017).

Pneumonia adalah infeksi pernapasan akut pada bagian kantung kecil paru- paru (alveoli). Alveoli akan dipenuhi nanah dan cairan sehingga bernapas menjadi sakit karena asupan oksigen terbatas. Hal ini sebagian besar disebabkan oleh agent infeksi dan dapat menyebar dengan cara yang berbeda seperti

batuk dan bersin. Pneumonia salah satu penyakit infeksi saluran pernapasan akut (ISPA) penyebab kematian utama pada anak usia dibawah lima tahun. World Health Organization (WHO) memperkirakan kematian balita karena pneumonia di seluruh dunia sebesar 15%, lebih dari dua juta anak meninggal karena pneumonia, hal ini menunjukkan bahwa satu dari lima balita meninggal dunia karena pneumonia (Patria, 2016). Terjadinya pneumonia ditandai dengan gejala batuk dan atau kesulitan bernapas seperti napas cepat, dan tarikan dinding dada bagian bawah ke dalam.

Pada umumnya, pneumonia dikategorikan kedalam penyakit menular yang ditularkan melalui udara, dengan sumber penularan adalah penderita *pneumonia* yang menyebarkan kuman dalam bentuk *droplet* ke udara pada saat batuk atau bersin. Untuk selanjutnya, kuman penyebab *pneumonia* masuk ke saluran pernapasan melalui proses *inhalasi* (udara yang dihirup), atau dengan cara penularan langsung, yaitu percikan *droplet* yang dikeluarkan oleh penderita saat batuk, bersin, dan berbicara langsung terhirup oleh orang di sekitar penderita, atau memegang dan menggunakan benda yang telah terkena *sekresi* saluran pernapasan penderita (Anwar & Dharmayanti, 2013).

Berdasarkan uraian latar belakang diatas maka dibutuhkan suatu cara atau metode yang dapat mendiagnosa penyakit *pneumonia*, mengikuti penelitian yang telah ada sebelumnya, yaitu (Effendi, 2016) membahas tentang “Sistem Diagnosis Penyakit Infeksi Saluran Pernapasan Akut Pada Anak Menggunakan Metode *Naive Bayes Classifier*”, peneliti diatas menerapkan metode *Naive Bayes* dengan tujuan agar dapat mendiagnosa penyakit berdasarkan gejala-gejala yang diderita dengan akurat, maka dari itu judul penelitian yang akan diambil yaitu “Penerapan *Data Mining* Metode *Naive Bayes* untuk Diagnosa Penyakit *Pneumonia* pada Balita. Penelitian ini menggunakan algoritma *naive bayes classifier* sebagai salah satu algoritma klasifikasi data mining untuk mendiagnosa penyakit *pneumonia* dengan cepat dan akurat.

2. Tinjauan Studi

Dibawah ini adalah beberapa penelitian tentang data mining ataupun mendekati penelitian yang *digunakan* sebagai bahan referensi:

- 1 Visualisasi Berbasis Naive Bayes untuk Pemetaan Penyebaran Penyakit

IIO

Infeksi Saluran Pernafasan Akut (Prasetio & Haryanto, 2017). Dalam penelitian ini peneliti melakukan visualisasi terhadap penduduk yang terkena penyakit ISPA dengan kategori Pneumonia Berat, Pneumonia Ringan dan non ISPA menggunakan metode Naive Bayes sehingga dapat menentukan daerah mana yang paling banyak terkena dan daerah mana yang sedikit terkena penyakit ISPA berdasarkan dari faktor geografis lingkungan seperti daerah pabrik, aliran sungai, jalan raya dan pencemaran udara.

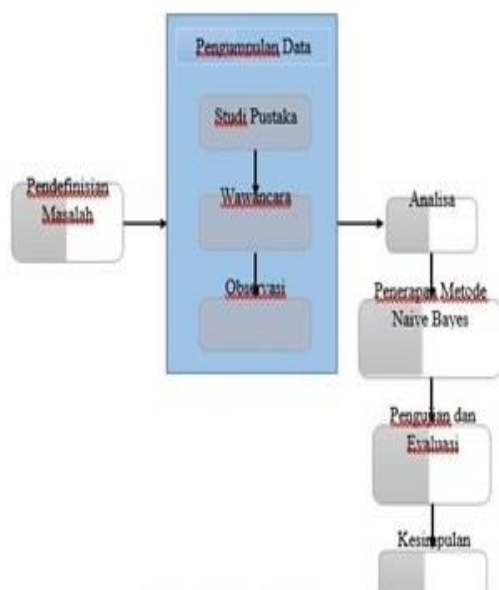
- 2 Sistem Pakar Diagnosis Penyakit Infeksi Saluran Pernafasan Akut Pada
- 3 Anak Menggunakan Metode Naive bayes Classifier (Effendi, 2016) Pada penelitian ini didapat hasil bahwa metode Naive Bayes Classifier dapat diterapkan pada sistem pakar layanan kesehatan diagnosis penyakit, khususnya penyakit ISPA pada anak dan dapat memberikan informasi mengenai gejala serta penanganan awal penyakit ISPA pada anak yang dibutuhkan oleh tim medis, orang tua dan masyarakat.
- 4 Sistem Pakar Pendeteksi Penyakit Paru Dan Saluran Pernafasan
- 5 Menggunakan Metode Naive Bayes (Rianto & Sudibyo, 2015). Dalam penelitian ini didapat hasil bahwa metode Naive Bayes dapat digunakan untuk mendiagnosa penyakit paru dan pernafasan, sistem dapat digunakan oleh user untuk mendiagnosa dan memberikan informasi tentang penyakit paru dan saluran pernafasan, selain itu sistem dengan metode ini mampu memberikan tingkat akurasi sebesar 77%.
- 6 Perancangan Data Mining untuk Klasifikasi Prediksi Penyakit ISPA
- 7 dengan Algoritma C4.5 (Tarigan et al., 2017). Pada penelitian ini mengusulkan perancangan data mining untuk memprediksi penderita penyakit ISPA untuk kategori Akut dan Tidak Akut menggunakan Klasifikasi Algoritma C4.5, karena rules yang terbentuk sederhana. Untuk pengukuran atau memilih atribut sebagai akar berdasarkan nilai gain tertinggi dari atribut-atribut yang ada. Dari semua penelitian dan metode yang digunakan diatas terbukti penggunaan metode Naive Bayes memiliki banyak kelebihan didalam hal prediksi dengan tingkat akurasi yang baik, oleh karena itu metode Naive Bayes

3. Kerangka Konsep

Pada tahap penelitian diawali dengan pendefinisian masalah, studi pustaka,

menerapkan metode yang akan di gunakan, melakukan pengujian dan evaluasi, kemudian pengambilan kesimpulan. Dibawah ini adalah tahapan penelitian sebagai berikut :

- Pendefinisian permasalahan yang berkaitan dengan penelitian kemudian merancang suatu analisa.
- Studi Pustaka : Pengumpulan data yaitu berupa literatur buku, jurnal nasional maupun internasional, kemudian penelitian yang pernah di lakukan peneliti sebelumnya.
- Wawancara : Melakukan proses wawancara tanya jawab dengan petugas kesehatan Puskesmas Cimalaka.
- Observasi : Pengamatan langsung proses kegiatan yang ada dengan hasil mendapatkan data-data penunjang untuk dilakukan penelitian.
- Analisa : Melakukan analisa terhadap data yang di dapat pada saat observasi dan penelitian yang telah dilakukan sebelumnya.
- Penerapan Metode : Metode yang digunakan yaitu *Naive Bayes Classifier* (NBC).
- Pengujian dan Evaluasi: Melakukan pengujian data menggunakan Rapid Miner Studio 7.6 dan mengevaluasinya.
- Pengambilan kesimpulan dari penelitian yang dilakukan dengan tujuan dapat memberikan hasil diagnosa yang cepat dan akurat.



Gambar 1. Diagram Pengumpulan data

4. Desain Penelitian / Hasil Metodologi

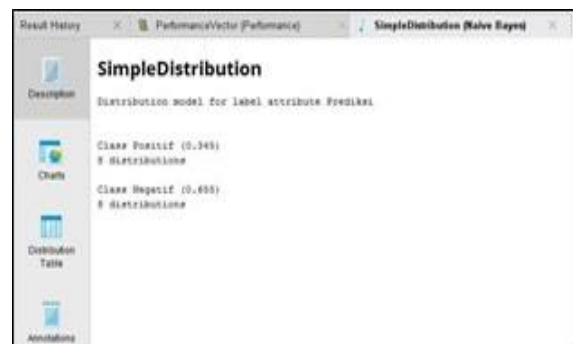
Berdasarkan permasalahan yang telah diuraikan sebelumnya, maka bahan dan peralatan yang diperlukan untuk penelitian ini meliputi :

Dalam penelitian ini bahan yang dibutuhkan adalah data sekunder berupa data balita yang mengalami gejala *pneumonia* untuk digunakan sebagai instrumentasi guna memperoleh data untuk mendiagnosa pasien. **Peralatan Penelitian** Peralatan penelitian yang digunakan diantaranya:

1. Kebutuhan Perangkat Lunak:
 - a. Microsoft Office Word 2013
Software ini digunakan untuk mengolah laporan hasil penelitian.
2. Microsoft Office Excel 2013
3. Software ini digunakan sebagai media penulisan dan pengolahan dataset, data training dan data testing.
4. Sistem operasi Microsoft Windows 10
5. Sistem Operasi yang digunakan peneliti.
6. Rapid Miner Studio 7.6
7. Tools yang digunakan dalam mengolah data serta untuk mengetahui tingkat akurasi dari data yang sedang diteliti.
8. Microsoft Visual Basic 6.0
9. Software yang digunakan untuk membuat sistem diagnosa.
10. Kebutuhan Perangkat Keras:
 - a. Prosesor AMD A8-4500 1.90 GHz
 - b. AMD Radeon 8750M 2GB DDR3
 - c. LCD 14-inchi HD
 - d. RAM 4GB DDR3
 - e. HDD 500 GB
 - f. Mouse

5. Hasil Penelitian

Menganalisa tabel data balita dalam mendiagnosa penyakit *pneumonia* dengan metode *Naive Bayes* dapat menghasilkan 2 *class* utama.



Gambar 2. Simple Distribution Model

Hasil klasifikasi dari data balita dengan menggunakan metode *Naive Bayes* membagi 2 class klasifikasi yaitu class Positif dan class Negatif, untuk nilai class Positif yaitu (0.345) dan nilai class Negatif yaitu (0.655). Tabel distribusi hasil analisa dengan metode *Naive Bayes* terhadap tabel data balita dalam mendiagnosa penyakit *Pneumonia* dapat dilihat pada tabel berikut:

No	Atribut	Value	Diagnosa			
			Positif	%	Negatif	%
1	Usia	<=3	54	78.26	107	81.68
		>3	15	21.74	24	18.32
	Total		69	100	131	100
2	Jenis Kelamin	Laki-laki	32	46.38	63	48.10
		Perempuan	37	53.62	68	51.90
	Total		69	100	131	100
3	Gejala 1	Ya	54	78.26	61	46.56
		Tidak	15	21.74	70	53.44
	Total		69	100	131	100
4	Gejala 2	Ya	62	89.86	39	29.78
		Tidak	7	10.14	92	70.22
	Total		69	100	131	100
5	Gejala 3	Ya	58	84.06	42	29.78
		Tidak	11	10.14	89	70.22
	Total		69	100	131	100
6	Gejala 4	Normal	2	2.90	122	93.12
		Tidak Normal	67	97.10	9	6.88
	Total		69	100	131	100
7	Gejala 5	Ya	67	97.10	7	5.34
		Tidak	2	2.90	124	94.66
	Total		69	100	131	100
8	Gejala 6	Normal	8	11.60	112	85.50
		Tidak Normal	61	88.40	19	14.50
	Total		69	100	131	100

Tabel 1. Tabel Distribusi

Tabel 1 menunjukkan bahwa balita yang rentan terkena *pneumonia* adalah usia <=3 tahun sebanyak 54 (78,26 %) dengan rata – rata jenis kelamin perempuan sebanyak 37 (53,62 %), sedangkan gejala yang paling berpengaruh untuk

terjadinya *pneumonia* adalah gejala 5 yaitu retraksi dinding dada sebanyak 67 orang (97,10%) dengan hasil positif menderita *pneumonia*.

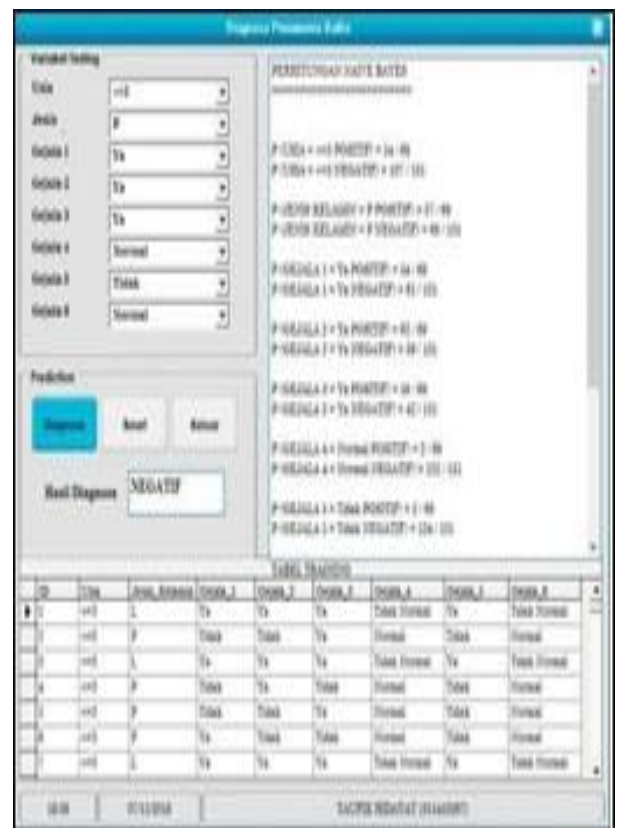
Penyebaran (Deployment)

Dalam penerapan pada sistem yang dibuat, *data training* yang dimasukkan adalah sebanyak 200 *record*, sistem ini hanya diterapkan dengan alur algoritma *Naive Bayes*, berikut adalah tampilan dari aplikasi yang dibuat dengan *Visual Basic 6.0*.



Gambar 3 Tampilan Form Login

Pada gambar 4.10 dan gambar 4.11 adalah hasil diagnosa dimana balita berpotensi Positif *pneumonia* jika pada Gejala 5 yaitu Ya, dan berpotensi Negatif terkena *pneumonia* apabila pada Gejala 5 yaitu Tidak.



Gambar 4 Tampilan Awal Sistem



Gambar 5 Tampilan Diagnosa Positif



Gambar 6 Tampilan Diagnosa Negatif

6. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan untuk mendiagnosa penyakit *pneumonia* dengan metode *Naive Bayes*, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Metode *Naive Bayes* dapat digunakan untuk mengklasifikasi data agar dapat mendiagnosa Balita Positif atau Negatif terkena penyakit *Pneumonia*.

2. Proses *Data Mining* dengan metode *Naive Bayes* dapat menganalisa dan mendiagnosa penyakit *pneumonia* cepat dan akurat, dari pengujian yang dilakukan menggunakan Rapid Miner di dapat tingkat akurasi sebesar 98.50%.

Daftar Pustaka

- Anwar, A., & Dharmayanti, I. (2013). *Pneumonia pada Anak Balita di Indonesia*, 8(8), 359–365.
- Darujati, C., & Gumelar, A. B. (2012). *Pemanfaatan Teknik Supervised Untuk Klasifikasi Teks Bahasa Indonesia*, 16(1).
- Effendi, H. (2016). *Sistem Pakar Diagnosis Penyakit Infeksi Saluran Pernapasan Akut Pada Anak Menggunakan Metode Naive Bayes Classifier*, 284–290.
- Ernawati, Riyanti, E., & Indraswari, R. (2017). *Faktor-faktor yang Berhubungan dengan Perilaku Kelurahan Tandang*, 5(5), 969–979.
- Imtiyaz, M. Z., Nasrun, M., & Ahmad, U. A. (2015). *Analisi dan Iplemetasi Framework CRISP-DM untuk Mengetahui Perilaku Data Transaksi Pelanggan (Case Study: PT X)*, 2(1), 596–602.
- Novianti, B., Rismawan, T., & Bahri, S. (2016). *Implementasi Data Mining dengan Algoritma C4 . 5 untuk Penjurusan Siswa (Sudi Kasus : SMA Negeri 1 Pontianak)*, 04(3), 75–84.
- Patria, M. A. (2016). *Faktor Risiko Pneumonia pada Balita di Indonesia : Narrative Review Penelitian AkademikBidang Kesehatan Masyarakat*, 10(2), 57–62.
- Prasetyo, Y., & Haryanto, H. (2017). *Visualisasi Berbasis Naive Bayes untuk Pemetaan Penyebaran Penyakit Infeksi Saluran Pernafasan Akut Naive Bayes Based Visualization For Mapping Acute Respiratory Infection Disease*, 7(1),
- Putri, D. L., & Santoso. (2016). *Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien (Studi Kasus : Puskesmas Kajen) K-Means Algorithm Implementation for Classification of Disease Patient (Case Study: Health Centers Kajen Regency Pekalongan)*, 1.

- Rianto, M. T., & Sudiby, U. (2015). Sistem Pakar Pendeteksi Penyakit Paru dan Saluran Pernafasan Menggunakan Metode Naïve Bayes.
- Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, 7(1), 59–64.
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga, 2(3), 207–217.
- Tarigan, D. M., Rini, D. P., & Puspita, V. (2017). Perancangan Data Mining untuk Klasifikasi Prediksi Penyakit ISPA dengan, 3(1), 179–182.