

ANALISIS PENYAKIT PARU-PARU MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS PADA RUMAH SAKIT ALOEI SABOE KOTA GORONTALO

Olha Musa¹ dan Alang²

¹0lh4mu54@gmail.com, ²virus.stimik@gmail.com
^{1,2}STMIK Ichsan Gorontalo

Abstrak

Penyakit paru-paru pada manusia adalah salah satu penyakit yang banyak terjadi pada saat sekarang ini. Adapun beberapa jenis penyakit paru-paru diantaranya Tuberkulosis (TBC), Asma, Bronkhitis, pneumonia, Emfisema dan kanker paru-paru. Metode yang digunakan pada penelitian ini adalah *K-Nearest Neighboryang* merupakan salah satu metode yang cocok digunakan dalam pengklasifikasian penyakit paru-paru. Kelebihan KNN memiliki beberapa kelebihan yaitu bahwa tangguh terhadap training data yang noisy dan efektif apabila data latih nya besar. Solusi pada permasalahan penyakit paru-paru yaitu dapat suatu analisis yang terkomputerisasi dan diiringi dengan perkembangan teknologi informasi supaya proses pengambilan keputusan menjadi lebih mudah, efektif dan efisien. Tujuannya diharapkan metode KNN menghasilkan dugaan atau hasil diagnosa yang sama dengan diagnosa yang dilakukan oleh seorang ahli pada rumah sakit aloei saboe kota gorontalo.

Kata kunci : Penyakit paru-paru, K-NN, Analisis Paru-paru

1. Pendahuluan

Penyakit paru-paru pada manusia adalah salah satu penyakit yang banyak terjadi pada saat sekarang ini. Hal ini dikarenakan pola gaya hidup masyarakat saat ini yang cenderung sibuk dengan padatnya jadwal maupun tingkat mobilitas yang tinggi, dapat mempengaruhi kesehatan [1].

Paru-paru sebagai pompa satu-satunya untuk sistem pernapasan adalah organ yang sangat penting bagi berlangsungnya kehidupan. Sebagai bagian dari organ penting, paru-paru termasuk organ yang berukuran yang cukup besar dan hampir memenuhi rongga dada kita. Banyak orang menggunakan paru-paru dan sistem saluran pernapasannya bukan untuk mengisap oksigen dari udara bersih, melainkan mengisap asap hasil pembakaran tembakau, cengkeh, dan bahan-bahan psikotropika berbahaya lainnya yang tidak perlu disangkal lagi merupakan racun yang merusak paru-paru [2].

Algoritma Nearest Neighbor Retrieval (knearest neighbor atau K-NN) adalah sebuah algoritma untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma nearest neighbor. Algoritma Nearest Neighbor berdasarkan pada proses pembelajaran menggunakan analogi / learning by analogi. Training sampelnya dideskripsikan dalam bentuk atribut numerik n-dimensi. Tiap sampel mewakili sebuah titik pada ruang ndimensi. Dengan cara ini, semua training sampel disimpan pada pola ruang n-dimensi. Ketika diberikan "unknown" sampel, knearest neighbor classifier mencari pola ruang K training sampel yang paling dekat "unknown" sampel tersebut. K training sampel ini adalah k nearest neighbor dari unknown sampel. Unknown sampel ditetapkan dengan class yang paling umum diantara k nearest neighborsnya. Ketika $k = 1$, unknown sampel ditetapkan dengan class dari training sampel yang paling dekat dengan pola ruangnya [4].

Sedangkan kelemahan dari K-NN adalah K-NN perlu menentukan nilai dari parameter K (jumlah dari tetangga terdekat) dan pembelajaran berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan dan atribut mana yang harus digunakan untuk mendapatkan hasil yang terbaik.

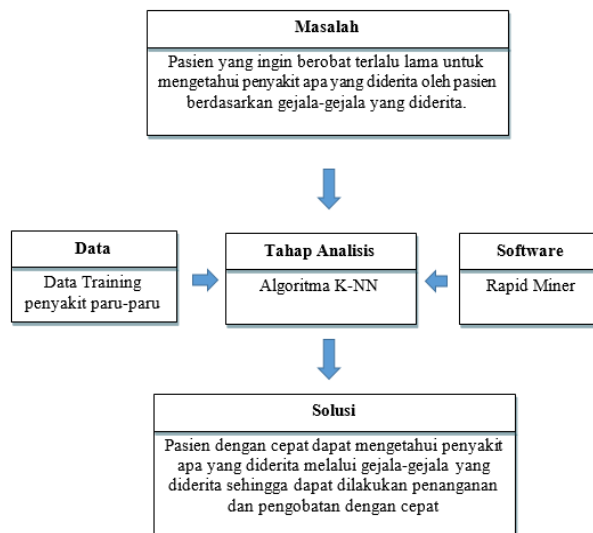
Solusi pada permasalahan penyakit paru-paru yaitu mendapatkan suatu analisis yang terkomputerisasi dan diiringi dengan perkembangan teknologi informasi supaya proses pengambilan keputusan menjadi lebih mudah, efektif dan efisien.

2. Metode

Metode pada penelitian analisis penyakit paru-paru menggunakan algoritma k-nearest neighbors(K-NN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. *K-Nearest Neighbor* berdasarkan konsep '*learning by analogy*'. Data learning dideskripsikan dengan atribut numerik n-dimensi. Tiap data learning merepresentasikan sebuah titik, yang ditandai dengan c, dalam ruang n-dimensi. Jika sebuah data query yang labelnya tidak diketahui diinputkan, maka *K-Nearest Neighbor* akan mencari k buah data learning yang jaraknya paling dekat dengan data query dalam ruang n-dimensi. Jarak antara data query dengan data learning dihitung dengan cara mengukur jarak antara titik yang merepresentasikan data query dengan semua titik yang merepresentasikan data learning dengan rumus *Euclidean Distance*.

Nilai k yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, k = 1) disebut algoritma *nearest neighbor*. Ketepatan algoritma K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performaklasifikasi menjadi lebih baik [3].

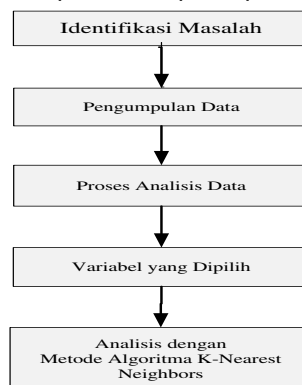
2.1 Kerangka Pikir



Gambar 1. Kerangka Pikir

2.2 Metodologi

Penelitian Analisis Penyakit Paru-Paru Pada Rumah Sakit Aloe Saboe Kota Gorontalo terdapat langkah-langkah penelitian yang mengacu pada metodologi *Algoritma K-Nearest Neighbors*. Langkah analisis penyakit paru-paru dapat dilihat pada prosedur penelitian terdapat pada Gambar 2.



Gambar 2. Prosedur penelitian Analisis Penyakit Paru-Paru dengan Metode *K-Nearest Neighbors*

3 Hasil dan Pembahasan

3.1 Hasil Luaran yang Dicapai

Hasil dari penelitian ini memperoleh dataset penyakit paru-paru. Dataset seperti berikut :

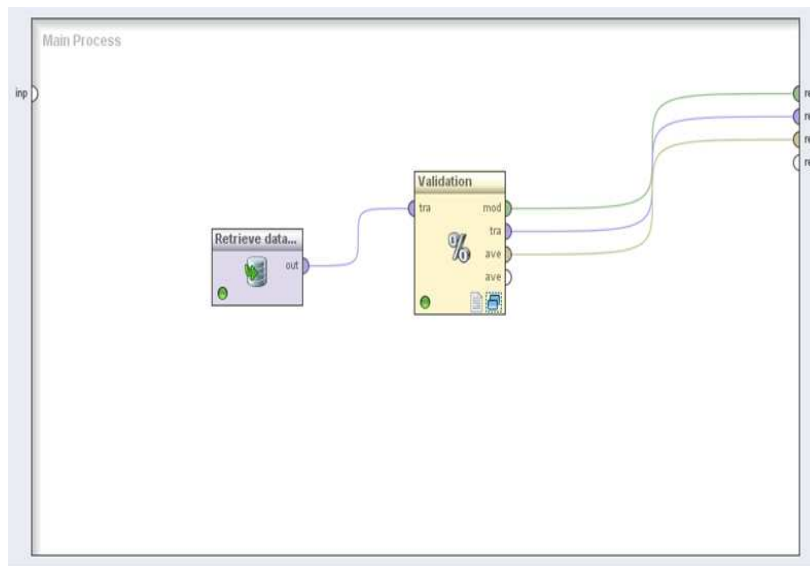
G1	G2	G3	G4	G5	G6	G7	G8	G9	HASIL
YA	YA	YA	YA	YA	YA	YA	YA	YA	POSITIF
YA	YA	YA	YA	TIDAK	YA	YA	YA	YA	POSITIF
YA	TIDAK	YA	YA	YA	YA	YA	YA	YA	POSITIF
YA	YA	YA	YA	YA	YA	YA	YA	TIDAK	POSITIF
TIDAK	YA	YA	YA	YA	YA	YA	YA	YA	POSITIF
YA	YA	YA	YA	YA	YA	YA	TIDAK	YA	POSITIF
YA	YA	YA	YA	YA	YA	TIDAK	YA	YA	POSITIF
TIDAK	YA	YA	YA	YA	YA	YA	YA	TIDAK	POSITIF
YA	YA	YA	YA	YA	TIDAK	TIDAK	YA	YA	POSITIF
YA	TIDAK	YA	YA	TIDAK	YA	YA	YA	YA	POSITIF
YA	YA	YA	YA	YA	TIDAK	YA	YA	TIDAK	POSITIF
YA	YA	YA	YA	TIDAK	YA	TIDAK	YA	YA	POSITIF
YA	TIDAK	YA	YA	YA	YA	YA	YA	TIDAK	POSITIF
YA	YA	YA	TIDAK	YA	TIDAK	YA	YA	YA	POSITIF
YA	TIDAK	YA	YA	YA	YA	YA	TIDAK	YA	POSITIF
YA	YA	TIDAK	YA	YA	TIDAK	YA	YA	YA	POSITIF
YA	YA	YA	YA	YA	YA	YA	YA	YA	POSITIF
YA	YA	YA	TIDAK	TIDAK	YA	YA	YA	YA	POSITIF
TIDAK	TIDAK	YA	YA	YA	YA	YA	YA	YA	POSITIF
YA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	YA	NEGATIF
YA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	YA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	TIDAK	YA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	TIDAK	TIDAK	YA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	TIDAK	TIDAK	TIDAK	YA	TIDAK	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	YA	TIDAK	TIDAK	TIDAK	NEGATIF
TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	YA	TIDAK	TIDAK	NEGATIF

G1	Batuk-batuk yang bisa menjadi batuk berdahak. Batuk ini berlangsung selama 21
G2	Batuk yang mengeluarkan darah.
G3	Sesak Napas
G4	Nyeri dada
G5	Tidak nafsu makan
G6	Penurunan berat badan
G7	Demam dan menggigil
G8	Berkeringat secara berlebihan pada malam hari
G9	Sering Kelelahan
HASIL	Hasil Diagnosa

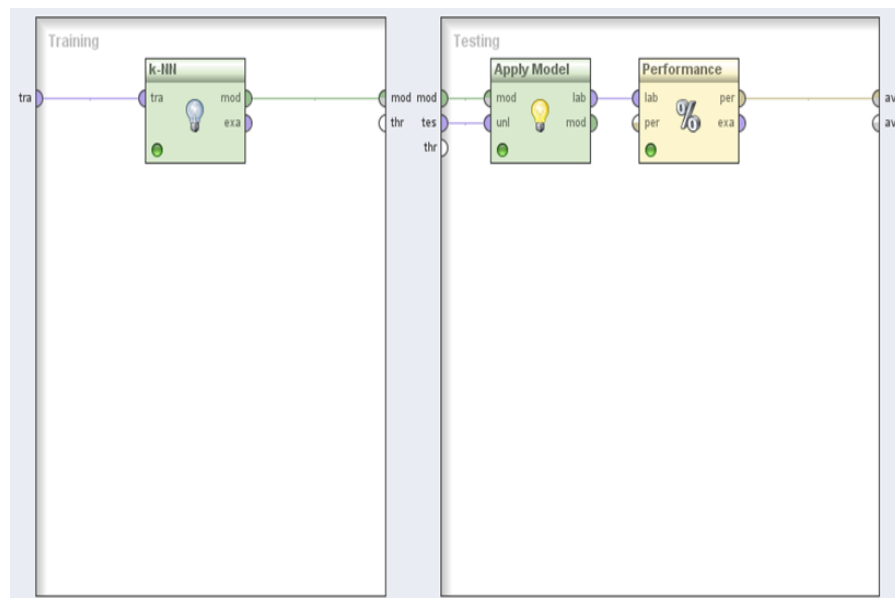
Gambar 3. Dataset Penyakit Paru-Paru

3.2 Rancangan Model Analisis

Hasil dari penelitian ini berupa model yang digunakan untuk menganalisis serta memprediksi penyakit paru-paru. Penggambaran model seperti berikut :



Gambar 4. Desain Model



Gambar 5. Desain Model

3.3 Hasil Analisis Prediksi

Adapun hasil prediksi dari model yang telah dirancang adalah sebagai berikut:

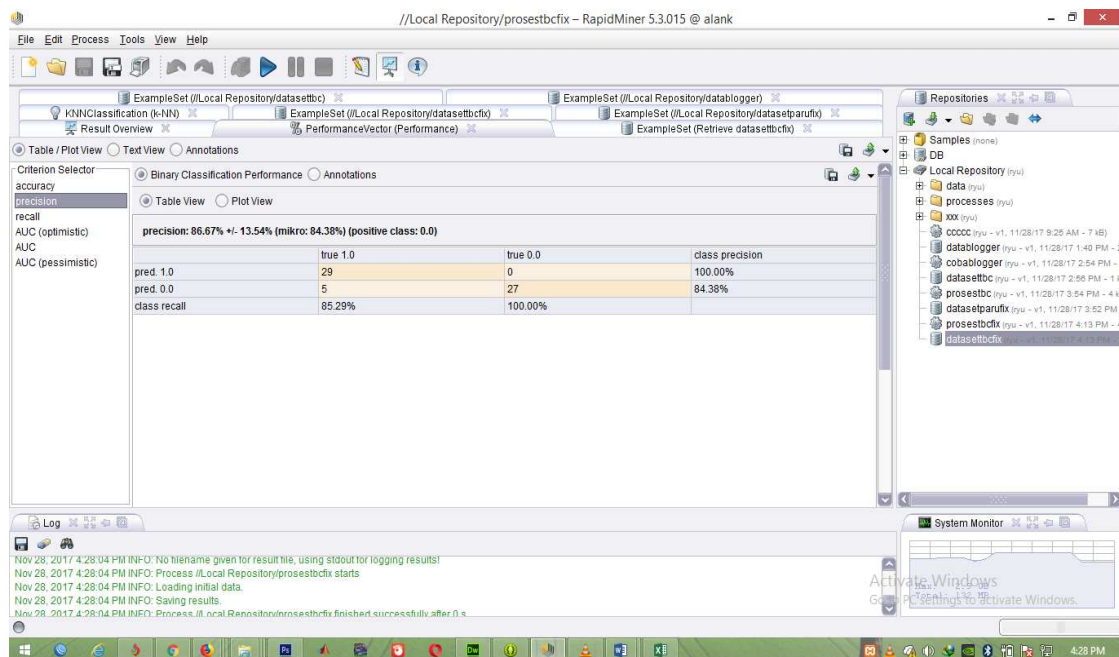
Multiclass Classification Performance
 Annotations
 📄 📥

Table View
 Plot View

accuracy: 91.90% +/- 8.12% (mikro: 91.80%)

	true 1.0	true 0.0	class precision
pred. 1.0	29	0	100.00%
pred. 0.0	5	27	84.38%
class recall	85.29%	100.00%	

Gambar 6. Hasil Analisis Prediksi Berupa Akurasi



Gambar 6. Hasil Analisis Prediksi Berupa Presisi

4. Kesimpulan dan Saran

4.1 Kesimpulan

Proses analisa penyakit paru-paru menggunakan algoritma K-Nearest neighbor , diperoleh kesimpulan sebagai berikut :

1. Hasil prediksi yang dilakukan oleh algoritma K-Nearest Neighbor menghasilkan akurasi yang cukup tinggi mencapai 91.90% dengan demikian mampu mendeteksi penyakit paru-paru secara akurat.
2. Berdasarkan dari nilai presisi yang mencapai 86.67%, algoritma K-Nearest Neighbor mampu mendeteksi penyakit paru-paru dengan tepat.

4.1 Saran

Diharapkan penelitian selanjutnya menggunakan algoritma lain untuk meningkatkan akurasi dalam hal prediksi penyakit paru-paru sehingga nantinya mampu untuk dijadikan acuan untuk menciptakan sebuah sistem pakar yang tepat dan akurat.

Daftar Pustaka

- [1] Anies, 2005, *Pencegahan Dini Gangguan Kesehatan*, PT. Elex Media Komputindo, Jakarta
- [2] Juanda, 2006, *Torch Akibat Dan Solusinya*, PT. Wangsa Jatra Lestari, Solo
- [3] Kusrini, 2008, *Konsep dan Aplikasi Sistem Pakar*, Andi, Yogyakarta.
- [4] Novita Mariana dkk, 2015, Penerapan Algoritma K-NN (Nearest Neighbor) Untuk Deteksi Penyakit (Kanker Serviks), *Jurnal Dinamika Informatika – Vol.7 No. 1*, Maret 2015 ISSN: 2085-3343