

Implementasi Data Mining Pada Pasien Setelah Koma dengan Menggunakan Metode Algoritma C 4.5 dan Naivy Bayes untuk Mengetahui Intelegensi (IQ) Pasien

Ferry Aditya Nugroho
Program Studi Informatika
Universitas PGRI Yogyakarta
Yogyakarta, Indonesia
ferynugroho323@gmail.com

Mustika Bintang Islamia
Program Studi Informatika
Universitas PGRI Yogyakarta
Yogyakarta, Indonesia
mustikabintang6@gmail.com

Meilany Nonsi Tentua
Program Studi Informatika
Universitas PGRI Yogyakarta
Yogyakarta, Indonesia
meilany@upy.ac.id

Abstrak— Gagar Otak atau kepala terbentur, efeknya mirip dengan suatu dosis Asam lisergat dietilamida (LSD). Pada dunia kesehatan banyak obat-obatan yang mengandung psikedelik, diprediksi kandungan psikedelik ini dapat meningkatkan kreativitas seseorang karena obat-obatan yang mengandung psikedelik dapat meningkatkan kadar serotonin. Metode Naive Bayes dan Metode C4.5 merupakan metode klasifikasi klasik yang masih banyak digunakan untuk beberapa penelitian. Perhitungan yang telah dilakukan terlihat bahwa metode C4.5 mempunyai tingkat akurasi lebih tinggi yaitu sebesar 63%.

Kata kunci— *Intelegensi, Gagar Otak*

I. PENDAHULUAN (HEADING 1)

Dalam ilmu kesehatan apabila orang mengalami kecelakaan dan sampai terjadi benturan keras disekitar kepala serta mengalami koma maka disaat orang itu sadar diperlukan penanganan khusus salah satunya yang terpenting adalah tes IQ. ditentukan bahwa tes IQ sangat penting bagi diri seseorang, dikarenakan seseorang yang pernah mengalami koma serta cedera didaerah kepala kemungkinan besar mengalami gangguan pada otak dan sistem saraf, kerusakan yang terjadi di otak dan saraf pendukung otak, dapat memengaruhi fungsi saraf dan otak.

Algoritma C4.5 merupakan salah satu metode yang digunakan dalam menyelesaikan sebuah kasus dimana metode Algoritma C4.5 yang digunakan ini dapat melakukan klasifikasi atau segmentasi atau pengelompokan data yang ada. Dalam metode Algoritma C4.5 ini juga bersifat prediktif yaitu sebuah proses yang dapat melakukan prediksi atau memperkirakan suatu kasus yang paling mungkin terjadi pada kasus yang ada. Sementara sifat klasifikasi atau pengelompokan merupakan sebuah proses atau sistematis data yang memiliki fungsi untuk mengelompokan data yang memiliki nilai sama atau satu kategori data sehingga dapat menghasilkan pola yang memiliki ukuran relatif besar hingga sangat besar dan memiliki nilai yang sama.

Metode Naive Bayes merupakan sebuah sistematis pengklasifikasian dengan memanfaatkan sebuah metode probabilitas yang ditemukan oleh ilmuwan Inggris Thomas Bayes, yaitu sebuah metode yang bertujuan untuk memprediksi peluang yang mungkin terjadi atau sering muncul dari sebuah data kasus yang terjadi paling besar sebelumnya sehingga dikenal sebagai teorema Bayes. [1] Teorema ini merupakan sebuah kombinasi dengan memperhatikan naive yang dapat diasumsikan sebagai

keadaan antara aturan/arah (atribut) saling keterkaitan. Salah satu pengimplementasian terbesar metode Naive Bayes yaitu pada bidang kesehatan. [2]

II. TINJAUAN PUSTAKA

Sebagai bahan pertimbangan dalam pembuatan karya ini penulis melakukan tinjauan pustaka sebagai berikut :

Penelitian yang berjudul “Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT. Arupadhatu Adisesanti” dari Jurusan Sistem Informasi Universitas Putra Batam [3]

Penelitian tentang Klasifikasi Penyakit Hipertensi Menggunakan Algoritma C4.5 Studi Kasus RSUD Provinsi NTB diambil dari Skripsi Audio Alief Kautsar Hartama (2017), mahasiswa Teknik Informatika Fakultas Sains dan Teknologi Universitas Sanata Dharma [4].

Klasifikasi Status Gizi Balita Menggunakan Naive Bayes Classification (Studi Kasus Posyandu Ngudi Luhur) yang diambil dari Skripsi Adrian Satria Putra (2018), mahasiswa Teknik Informatika Fakultas Sains dan Teknologi Universitas Sanata Dharma [5].

Penelitian tentang Implementasi Metode Naive Bayes Pada Sistem Diagnosa Penyakit Ikan Lele diambil dari Skripsi Siska Puspitaningsih (2018), mahasiswa Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim [6].

III. METODE PENELITIAN

Penelitian ini menggunakan survey deskriptif dengan menggunakan metode Algoritma C4.5 dan naive bayes. Penelitian dilaksanakan di salah satu rumah sakit. Penelitian ini menggunakan metode kuantitatif dengan jenis penelitian korelasional; yaitu mencari korelasi antara variabel bebas dengan variabel terikat. Menurut Jalaludin Rahmat (2004:90), [7]

Pada penggunaan metode korelasi ini bertujuan untuk meneliti keterkaitan antara variasi pada satu faktor dengan variasi yang terdapat pada faktor lain. Penelitian ini disebut penelitian korelasional, karena peneliti ingin mengetahui tingkat hubungan antara tingkat intelegensi seseorang (IQ) dan pasien koma dengan hasil intelegensi setelah pulih dari koma atau gagar otak.

IV. PEMBAHASAN DAN HASIL

Sebagian besar masyarakat sering berfikir bahwa semakin tinggi tingkat IQ, maka orang tersebut akan semakin pintar. Padahal hal tersebut tidak menjadi pengaruh. Terdapat banyak faktor yang dapat menentukan kecerdasan intelegensi seseorang. Ada beberapa tingkatan dalam IQ yaitu sebagai berikut.

- **Idiot (IQ 0-29)**
Tingkat IQ yang paling rendah adalah idiot dimana seseorang mengalami kesulitan dalam berkomunikasi dan berinteraksi dengan orang lain bahkan bergantung pada orang lain.
- **Imbicile (IQ 30-40)**
Pada tingkat Imbicile ini adalah dimana seseorang mampu berkomunikasi tetapi orang tersebut setara dengan anak-anak umur 3-7 tahun.
- **Moron atau Mentally Retarded (IQ 50-69)**
Pada tingkat ini seseorang bisa diajari berhitung sederhana bahkan mereka bisa membaca dan menulis.
- **Dull atau Borderline (70-79)**
IQ pada tingkat ini sudah tergolong lebih baik dari tiga diatas tetapi masih dibawah normal karena memerlukan usaha ekstra dalam pengembangannya.
- **Normal (80-119)**
Tingkat normal adalah dimana seseorang bisa melakukan sesuatu sendiri dengan normal tetapi ditingkatan ini ada seseorang yang tidak bisa cepat dalam memahami sesuatu ada yang mampu mengikuti dengan baik meski terkadang kesulitan dan ada yang mampu berada ditingkat yang lebih atas dibanding lainnya.
- **Cerdas atau Superior (120-129)**
Pada tingkat ini dimana seseorang biasanya memiliki nilai yang lebih tinggi dibanding lainnya dalam pendidikan dan lebih cepat dalam memahami sesuatu.
- **Sangat Cerdas atau Very Superior (130-139)**
Kemampuan-kemampuan yang dimiliki pada tingkat ini seperti membaca, menulis, dan menahami sesuatu lebih baik dari sebelumnya. Seseorang yang berada pada fase ini akan memiliki kesehatan, kekuatan dan ketangkasan yang cukup baik.
- **Genius (140+)**
Tingkat IQ ini adalah tingkatan tertinggi dimana seseorang mampu memahami suatu hal dengan sangat cepat tanpa harus bersusah payah untuk belajar ekstra. Contohnya seperti Albert Einstein.[8]

Pada kenyataannya kecerdasan dan kemampuan otak seseorang dapat diasah dan dapat berkembang jadi meski ada tingkatan IQ seperti diatas bukan berarti menjadi patokan untuk menilai kecerdasan seseorang.

Pada Tabel dibawah, diperoleh hasil IQ hanya 9 orang yang termasuk kategori Cerdas (4%). Sebanyak 187 orang yang termasuk kategori Normal (81%). Sementara pada kategori Dibawah Rata-rata terdapat 36 orang (15%), Sedangkan pada kategori Genius tidak ditemukan.

Tabel I. TABEL HASIL IQ

Kategori Intelligence Qoutien (IQ)		
IQ	N	%
Genius > 140	0	0
Cerdas > 120-132	9	4
Normal > 80-119	187	81
Dibawah rata-rata <79	36	15
TOTAL	232	100

- Jumlah informasi yang dapat dikaitkan atau dihasilkan dalam suatu kejadian dapat membuat persoalan informasi menjadi nilai kuantitas, sehingga dapat diterapkan pada proses perhitungan aritmatika. Secara umum, dapat diberikan sebuah nilai distribusi probabilitas $P = (p_1, p_2, \dots, p_n)$ dan sampel membawa informasi oleh distribusi ini, juga disebut entropy P yang menghasilkan pada fungsi s dibawah ini [9] :

$$Entropi (S) = \sum_{i=1}^n p_i \log_2 p_i \tag{1}$$

Keterangan :
 S : Himpunan (dataset)
 n : banyaknya record
 pi : probabilitas yang didapat dari jumlah

- **Informasi Gain G (p,T)**
Informasi gain adalah seleksi pengukuran atribut untuk memilih atribut yang akan di test pada tiap simpul (node). Atribut dengan informasi gain tertinggi akan terpilih sebagai atribut yang akan di test dari suatu node. Dalam perhitungan gain bisa terjadi atau tidak suatu missing value. Untuk langkah selanjutnya dengan menghitung gain (2), adalah sebagai berikut:

$$Gain (A) = Entropi (S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi (S_i) \dots \tag{2}$$

Keterangan :
 S : himpunan (dataset)
 A : atribut yang akan dipakai
 n : jumlah partisi atribut A
 |Si| : jumlah kasus pada partisi ke-i
 |S| : jumlah kasus dalam S

- Ulangi langkah ke-b dan ulangi hingga semua record terpartisi
- Proses partisi akan berhenti saat :
 - Semua record dalam node N mendapat kelas yang sama
 - Tidak ada atribut didalam data record yang dipartisi lagi sampai tidak ada record di dalam cabang yang kosong.

Hitung Entropy dari seluruh nilai atribut

Dengan perhitungan yang sama dilakukan perhitungan entropy terhadap keseluruhan nilai atribut sampai keseluruhan nilai atribut sampai diketahui nilai entropynya. Contoh perhitungan entropy pada atribut duration adalah:

- Entropy (duration)

$$= \left(-\left(\frac{1}{28}\right) \times \left(\ln\left(\frac{1}{28}\right) / \ln(2) \right) \right) + \left(-\left(\frac{7}{28}\right) \times \left(\ln\left(\frac{7}{28}\right) / \ln(2) \right) \right)$$

$$= (0,17169) + (0,5)$$

$$= 0,67169$$

- Entropy (plq)

$$= \left(-\left(\frac{1}{85}\right) \times \left(\ln\left(\frac{1}{85}\right) / \ln(2) \right) \right) + \left(-\left(\frac{9}{85}\right) \times \left(\ln\left(\frac{9}{85}\right) / \ln(2) \right) \right)$$

$$= (0,07540) + (0,34300)$$

$$= 0,4184$$

- Perhitungan Entropy total

Entropy (total)

$$= \left(-\left(\frac{53}{232}\right) \times \left(\ln\left(\frac{53}{232}\right) / \ln(2) \right) \right) + \left(-\left(\frac{179}{232}\right) \times \left(\ln\left(\frac{179}{232}\right) / \ln(2) \right) \right)$$

$$= 0,53739$$

Hasil perbandingan tingkat akurasi kedua metode tersebut dapat dilihat pada tabel II.

Tabel II. TINGKAT AKURASI

	Naïve Bayes	C45
Akurasi	0,52	0,63
Presisi	0,50	0,60
Recall	0,51	0,62

V. PENUTUP

Tes IQ dijadikan sebuah uji standar yang dirancang untuk menilai kecerdasan seseorang, apakah kurang, rata-rata, atau cerdas. Kebanyakan orang berusaha mendapatkan nilai yang

tinggi dalam tes IQ, karena hasil IQ yang tinggi dianggap dapat mempengaruhi kesuksesan seseorang, baik dalam pendidikan maupun karier. Dari hasil perhitungan dapat dilihat bahwa klasifikasi dengan menggunakan metode C45 mempunyai tingkat akurasi yang lebih tinggi dibandingkan metode Naïve Bayes.

DAFTAR PUSTAKA

- [1] D. Nofriansyah, K. Erwansyah, and M. Ramadhan, "Penerapan Data Mining dengan Algoritma Naive Bayes Classifier untuk Mengetahui Minat Beli Pelanggan terhadap Kartu Internet XL (Studi Kasus di CV. Sumber Utama Telekomunikasi)," no. 2, p. 12, 2016.
- [2] M. Sabransyah and Y. N. Nasution, "Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung," vol. 8, p. 8, 2017.
- [3] E. Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti," J. Online Inform., vol. 2, no. 1, p. 36, Jul. 2017, doi: 10.15575/join.v2i1.71.
- [4] A. A. K. Hartama, "Klasifikasi Penyakit Hipertensi Menggunakan Algoritma C4.5 Studi Kasus RSU Provinsi NTB," Univ. Sanata Dharma, 2017.
- [5] A. S. Putra, "Klasifikasi Status Gizi Balita Menggunakan Naive Bayes Classification (Studi Kasus Posyandu Ngudi Luhur)," Univ. Sanata Dharma, 2018.
- [6] S. Puspitaningsih, "Penelitian tentang Implementasi Metode Naive Bayes Pada Sistem Diagnosa Penyakit Ikan Lele," Univ. Islam Negeri Maulana Malik Ibrahim, 2018.
- [7] I. Veriansyah and M. G. Rindarjono, "Hubungan Tingkat Intelegensi (Iq) Dan Motivasi Belajar Geografi Dengan Hasil Belajar Siswa Kelas X Sekolah Menengah Atas Negeri Singkawang Kota Tahun Ajaran 2016/2017," vol. 4, no. 1, p. 10, 2018.
- [8] Alfa, "https://blogs.itb.ac.id/wikia/berapa-iq-tertinggi-manusia/," Oct. 14, 2019. di akses 29 Maret 2020, Jam 15:50
- [9] Tentua, M.N., Sihabuddin, A., "Improved C4.5 algorithm using the L'hospital rule and pruning on the recommendation system", International Journal of Scientific and Technology Research, Volume 8, Issue 11, November 2019