

Penerapan Algoritma ID3 Decision Tree Pada Klasifikasi Penyakit Diabetes

Fida Maisa Hana

Program Studi Sistem Informasi, Universitas Muhammadiyah Kudus

e-mail: fidamaisa@umkudus.ac.id

Intisari

Kadar glukosa darah yang melebihi normal adalah tanda diabetes melitus (DM), suatu kondisi yang kompleks dan kronis yang membutuhkan perawatan medis yang berkelanjutan untuk menurunkan risiko komplikasi melalui pengontrolan glikemik. Pada tahun 2015, ada 10 juta orang di Indonesia yang menderita diabetes. Jumlah penderita diabetes ini meningkat setiap tahun. International Diabetes Federation memperkirakan bahwa jumlah orang di Indonesia yang menderita diabetes akan meningkat menjadi 16.2 juta pada tahun 2040. Untuk mengurangi risiko komplikasi pada masa yang akan datang, diabetes harus dideteksi sejak dini. Salah satu metode yang bisa digunakan untuk memprediksi penyakit diabetes adalah algoritma ID3 Decision Tree. Algoritma ID3 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Dari hasil Pengujian menghasilkan akurasi yang cukup besar yaitu sebesar 98,08 % dengan Precision sebesar 96,88% % dan Recall sebesar 100%.

Kata kunci— Data Mining, Klasifikasi, ID3 Decision Tree, Diabetes Mellitus

Abstract

A Blood glucose levels that exceed normal are a sign of diabetes mellitus (DM), a complex and chronic condition that requires ongoing medical care to reduce the risk of complications through glycemic control. In 2015, there are 10 million people in Indonesia who suffer from diabetes. The number of diabetics is increasing every year. The International Diabetes Federation estimates that the number of people in Indonesia who suffer from diabetes will increase to 16.2 million in 2040. To reduce the risk of complications in the future, diabetes must be detected early. One method that can be used to predict diabetes is the ID3 Decision Tree algorithm. The ID3 algorithm is an algorithm used to form a decision tree. From the test results, it produces a fairly large accuracy of 98.08% with a Precision of 96.88% and a Recall of 100%.

Keywords—Data Mining, Classification, ID3 Decision Tree, Diabetes Mellitus

PENDAHULUAN

Diabetes mellitus adalah gangguan metabolis yang kronis di mana tubuh pasien tidak menghasilkan cukup insulin atau tubuh pasien tidak dapat memanfaatkan insulin dengan baik, yang menyebabkan gula darah berlebihan dalam tubuh, yang sering menyebabkan komplikasi pada organ tubuh [1]. Saat kadar glukosa darahnya melebihi batas normal, pasien didiagnosa diabetes [2]. Karena diabetes melitus adalah kondisi yang sangat kompleks, perawatan medis yang berkelanjutan sangat penting untuk mengurangi efek komplikasi, termasuk pengecekan glikemik [3]. Pada tahun 2015, ada 10 juta orang di Indonesia yang menderita diabetes. Jumlah penderita diabetes ini meningkat setiap tahun. International Diabetes Federation memperkirakan bahwa jumlah orang di Indonesia yang menderita diabetes akan meningkat menjadi 16.2 juta pada tahun 2040 [4]. Untuk menangani masalah ini, penyakit diabetes harus dideteksi sejak dini. Diharapkan bahwa diagnosis dini dapat mengurangi risiko komplikasi pada pasien diabetes di kemudian hari. Pencatatan diabetes banyak dilakukan untuk mencegah penyakit ini.

Teknik klasifikasi data mining adalah salah satu metode pencatatan yang dapat digunakan. Data mining dapat digunakan untuk mengekstrak informasi implisit dan berharga dari data [5]. Menurut [6], data mining adalah suatu proses yang menghasilkan hubungan yang memiliki arti, pola, dan kecondongan dengan memanfaatkan teknik identifikasi pola saat melihat volume data yang besar disimpan. Data mining biasanya menggunakan metode seperti deskripsi atau penggambaran, prediksi atau ramalan, clustering, klasifikasi dan asosiasi, dan estimasi [6]. Klasifikasi adalah proses membuat fungsi atau model yang menjelaskan kelas pada data untuk memprediksi kelas objek yang labelnya belum diberikan [7]. Pada penelitian ini, teknik klasifikasi digunakan untuk memprediksi siapa yang menderita diabetes dan siapa yang tidak.

Perhitungan proses klasifikasi dapat dilakukan dengan beberapa algoritma. Decision tree adalah salah satu teknik klasifikasi yang paling umum digunakan dan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode ini menggunakan representasi struktur pohon, atau pohon, di mana setiap node menunjukkan atribut, cabang menunjukkan nilai atribut, dan daun menunjukkan kelas [8].

Model algoritma dari decision tree adalah ID3, C4.5, dan CART, ketiganya memiliki performa yang berbeda-beda [9]. Peneliti telah melakukan penelitian [10] menggunakan algoritma decision tree C4.5 untuk klasifikasi diabetes menghasilkan akurasi yang cukup besar yaitu 97,12 % Precision sebesar 93,02% %, dan Recall sebesar 100,00%. Algoritma C4.5 adalah pengembangan dari algoritma ID3 [6].

Iterative Dichotomiser Three (ID3) adalah algoritma decision tree learning (algoritma pembelajaran pohon keputusan) yang paling dasar. Algoritma ini menguji semua kemungkinan pohon keputusan. Dengan algoritma ID3,

pembentukan pohon klasifikasi dilakukan dalam dua tahap. Pertama, nilai entropy dihitung dan kedua, nilai information gain dihitung untuk setiap variabel [11]. Uji cross-validation telah dilakukan pada klasifikasi metode ID3, yang menunjukkan akurasi sebesar 91.81%. Hasil dari Pohon keputusan metode ID3 membantu memprediksi identifikasi ketunaan pada ABK [12].

Bersadarkan latar belakang diatas, peneliti menggunakan algoritma ID3 untuk proses klasifikasi penyakit diabetes.

LANDASAN TEORI

1. Data Mining

Data mining adalah teknik yang digunakan untuk menggali informasi yang belum ditemukan secara manual dari kumpulan data [13]. Istilah "pengolahan data" sudah dikenal sejak tahun 1990 karena mengolah data sangat penting untuk berbagai bidang, seperti kesehatan, akademik, dan bahkan industri [14]. Data mining adalah perpaduan ilmu dari penelitian basis data, statistik, dan kecerdasan buatan. Menurut [6], ramalan atau prediksi, penggambaran atau deskripsi, klasifikasi, estimasi, asosiasi, dan clustering adalah enam metode umum yang digunakan dalam data mining untuk menemukan hubungan yang mengandung arti, pola, dan hubungan dengan mengolah kelompok data. Penjabaran dari ke enam metode tersebut adalah sebagai berikut:

- a. Deskripsi
Tujuan operasi adalah untuk menciptakan kriteria dan aturan yang mudah dipahami untuk pola yang terbentuk berulang-ulang pada sekelompok data.
- b. Prediksi atau ramalan
Teknik ramalan mirip dengan klasifikasi, tetapi data kelas sesuai dengan perilaku atau nilai yang diprediksi.
- c. Klasifikasi
Metode klasifikasi, juga dikenal sebagai pengelompokan, memungkinkan data diidentifikasi dengan karakteristik tertentu dan kemudian dikelompokkan dalam kelas-kelas tertentu.
- d. Estimasi
Teknik estimasi, juga dikenal sebagai menerka, mirip dengan prediksi, tetapi yang diestimasi adalah angka numerik.
- e. Clustering
Pengelompokan atau klasterisasi data dengan kelas data yang memiliki karakteristik yang sama di satu kelompok dan yang berbeda di kelompok lain dikenal sebagai klasterisasi.
- f. Asosiasi

- g. Teknik asosiasi adalah mencari fitur yang terlihat dalam situasi tertentu atau membuat aturan yang menghubungkan antara kombinasi item.

2. Klasifikasi

Klasifikasi adalah proses mencari karakteristik sebuah objek dan memasukkannya ke dalam salah satu kelas yang sudah diartikan terlebih dahulu. Proses klasifikasi adalah proses menghitung data sebelumnya (juga disebut training data dengan data baru atau testing data). Proses ini akan menghasilkan kemungkinan dalam pengujian data [6]. Semua dataset yang digunakan harus memiliki label atau atribut tujuan untuk diklasifikasikan. Tujuan klasifikasi adalah untuk menemukan objek kelas pada setiap persoalan data. Tugas klasifikasi dimulai dengan satu set data yang memiliki kelas yang dikenal.

3. ID3

Decision tree adalah salah satu teknik klasifikasi yang paling umum digunakan dan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode ini menggunakan representasi struktur pohon, atau pohon, di mana setiap node menunjukkan atribut, cabang menunjukkan nilai atribut, dan daun menunjukkan kelas [8].

Model algoritma dari decision tree adalah ID3, C4.5, dan CART, ketiganya memiliki performa yang berbeda-beda [9]. ID3 adalah algoritma decision ID3 learning (algoritma pembelajaran pohon) yang paling dasar. Algoritma ini melakukan pencarian rakus/menyeluruh pada semua pohon keputusan yang mungkin. ID3 (Iterative Dichotomizes 3) adalah algoritma induksi pohon keputusan yang dikembangkan oleh J. Ross Quinlan. Algoritma ID3 digunakan dengan fungsi rekursif, atau fungsi yang memanggil dirinya sendiri. Tujuan algoritma ini adalah untuk membangun pohon keputusan (pohon keputusan) dari atas ke bawah. Ini dimulai dengan pertanyaan, "atribut mana yang pertama kali harus dicek dan diletakkan pada root?" Untuk menjawab pertanyaan ini, algoritma mengevaluasi semua atribut yang ada dengan menggunakan ukuran statistik, yang biasanya digunakan adalah information gain [15].

Secara ringkas, berikut adalah tahapan bagaimana Algoritma ID3 bekerja:

1. Menyiapkan data training
2. Hitung nilai Entropy dengan rumus :

$$Entropy(S) = \sum_i^c - P_i \log_2 P_i$$
$$Entropy(S) = (- P_+ \log_2 P_+) - (P_- \log_2 P_-) \quad (1)$$

3. Setelah mendapatkan nilai Entropy akan mencari Information Gain dari setiap atribut untuk mendapatkan nilai Information Gain yang paling tinggi.
4. Rumus dari Information Gain yaitu :

$$Gain(S,A) = Entropy(S) - \sum_{v \in \text{nilai}(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

5. Nilai tertinggi dari Information Gain yang akan membentuk menjadi simpul yang pertama dan akan menempati posisi paling atas.
6. Proses perhitungan Information Gain diulangi sampai semua data yang termasuk dalam kelas yang sama.
7. Sehingga akan terbentuk pohon keputusan.
8. Maka akan terbentuk Rule-Rule.

4. Confusion Matrix

Confusion Matrix merupakan sebuah hasil evaluasi dari sebuah klasifikasi data mining yang diwujudkan dalam sebuah tabel [14]. *Confusion matrix* adalah metode yang banyak dipakai untuk menghitung nilai akurasi. Pengukuran kinerja menggunakan *confusion matrix* memiliki empat istilah sebagai gambaran dari hasil klasifikasi. Adapun keempat istilah tersebut yaitu :

- a. *False Positive* (FP), yaitu data negatif tapi terprediksi sebagai data positif.
- b. *False Negative* (FN), yaitu data positif yang terprediksi sebagai data negatif.
- c. *True Positive* (TP), yaitu data positif yang terprediksi benar.
- d. *True Negative* (TN), yaitu data negatif yang terprediksi dengan benar.

dengan klasifikasi yang sebenarnya. Bentuk *Confusion Matrix* secara umum dapat dicermati pada tabel dibawah ini:

Tabel 1. Tabel *Confussion Matrix* [14]

<i>Classification</i>		<i>Predicted class</i>	
		<i>Class : Yes</i>	<i>Class : No</i>
<i>Observed Class</i>	<i>Class Yes</i>	A(True Positive)	B(False Negative)
	<i>Class No</i>	C(False Positive)	D(True Negative)

Untuk menghitung akurasi digunakan rumus sebagai berikut:

$$Akurasi = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (3)$$

Algoritma klasifikasi pastinya berusaha untuk menghasilkan model yang menghasilkan akurasi yang baik. Kinerja model dari algoritma klasifikasi ditentukan pada saat model dihadapkan pada data *testing*, karena rata-rata model

yang dipakai dapat memprediksi dengan benar pada semua data yang menjadi data *trainingnya*.

Sensitivitas atau Recall adalah rasio prediksi benar positif dipadukan dengan keseluruhan data yang benar positif atau mengukur proporsi positif asli yang diramal secara benar sebagai positif. Dalam sensitivitas berkaitan dengan kecakapan pengujian untuk mengenali hasil yang positif dari sejumlah data yang seharusnya positif. Untuk menghitung sensitivitas atau recall menggunakan persamaan dibawah ini:

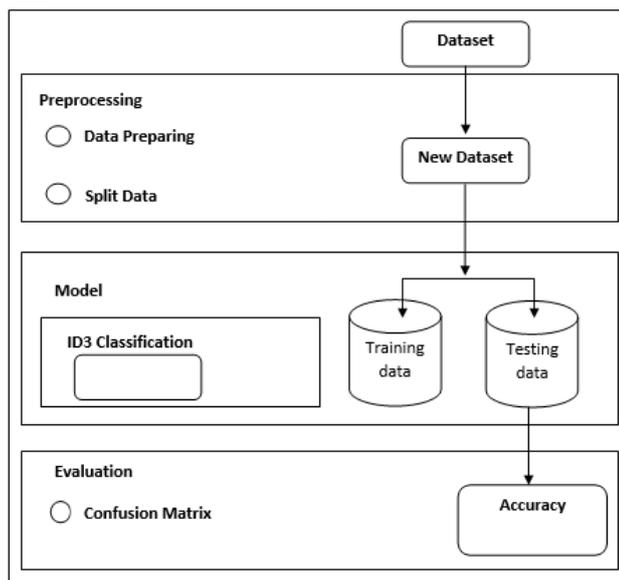
$$Sensitivitas = \frac{TP}{TP+FN} \quad (4)$$

Sedangkan precision adalah rasio ramalan benar positif dipadukan dengan semua hasil yang diprediksi positif. *Precision* menggambarkan matrik untuk menghitung kemampuan sistem dalam menghasilkan data yang penting. *Precision* pada data mining adalah hasil jumlah data yang true positive dibagi dengan jumlah data yang dikenali sebagai positif. Untuk menghitung precision menggunakan persamaan dibawah ini:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

METODE PENELITIAN

Penelitian ini menggunakan metode eksperimen untuk melakukan penelitian ini. Secara keseluruhan, penelitian ini melakukan komputasi pada dataset menggunakan algoritma klasifikasi ID3. Kerangka penelitian dari penelitian ini dapat dilihat pada gambar di bawah ini:



Gambar 1 Kerangka Penelitian

Langkah pertama adalah menyiapkan kumpulan data, kemudian membagi data menjadi data uji dan data latih dengan persentase 80% data latih dan 20% data uji. Data training berperan sebagai model atau generator model, sedangkan data test berperan sebagai model tester. Model yang digunakan adalah klasifikasi dengan algoritma ID3. kemudian dihitung akurasi menggunakan *confusion matrix*.

A. Sumber Data

Pada penelitian ini menggunakan data dari sumber dataset UCI Machine Learning Repository yang bisa ditemukan di alamat web <https://archive.ics.uci.edu/ml/machine-learning-databases/00529/>. Dataset yang digunakan adalah *Early stage diabetes risk prediction dataset* dimana file tersebut bernama *diabetes_data_upload.csv*.

Penelitian ini menggunakan 17 variabel dan total 520 data responden yang mencakup informasi tentang gejala yang menyebabkan diabetes. Kumpulan data ini dibuat dari kuesioner langsung kepada orang-orang yang baru saja menderita diabetes atau yang belum menderita diabetes tetapi memiliki sedikit atau lebih gejala. Data dikumpulkan dari pasien dengan menggunakan kuesioner langsung dari Rumah Sakit Sylhet, Sylhet, Bangladesh.

B. Split Data Otomatis

Pada tahap ini, 520 data diabetes dibagi menjadi dua, yaitu data *training* dan data *testing* dengan prosentase 80% untuk data *training* dan 20% untuk data *testing*. Data *training* bertindak sebagai pembentuk pola atau model dan data *testing* sebagai penguji model.

C. Metode yang Diusulkan

Penelitian ini menggunakan metode klasifikasi data mining dengan algoritma ID3. Langkah-langkah dalam klasifikasi menggunakan algoritma ID3 adalah sebagai berikut:

1. Siapkan data training terlebih dahulu
2. Selanjutnya menghitung nilai Entropy
3. Setelah mendapatkan nilai Entropy, tahap selanjutnya adalah mencari Information Gain dari setiap atribut untuk mendapatkan nilai Information Gain yang paling tinggi.
4. Nilai tertinggi dari Information Gain yang akan membentuk menjadi simpul yang pertama dan akan menempati posisi paling atas.
5. Proses perhitungan Information Gain diulangi sampai semua data yang termasuk dalam kelas yang sama. Sehingga akan terbentuk pohon keputusan, maka dari pohon keputusan akan terbentuk Rule-Rule.

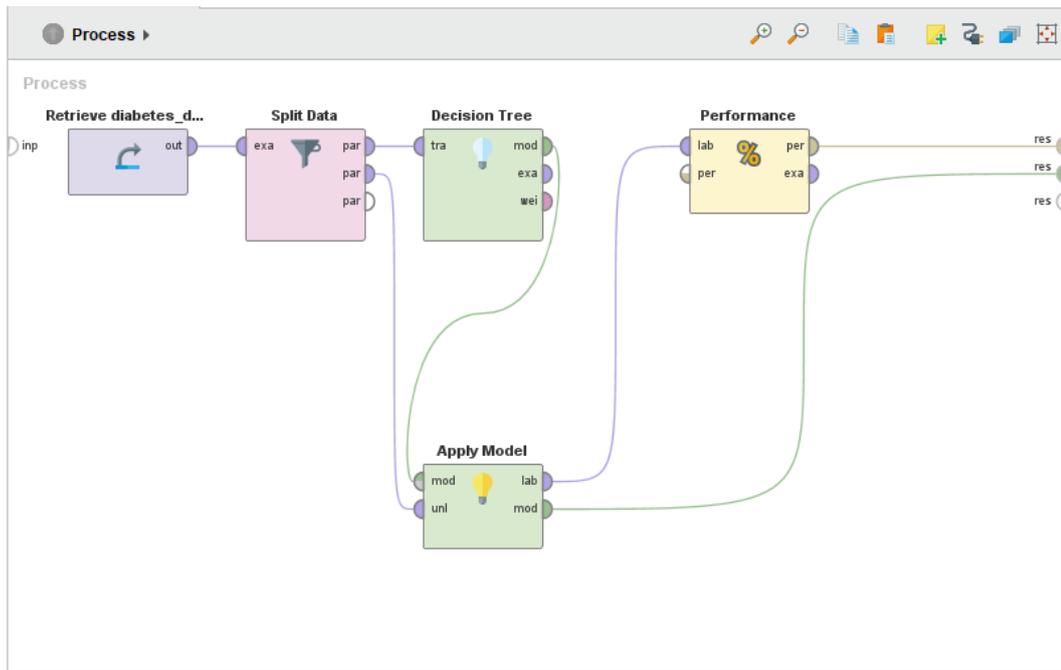
HASIL DAN PEMBAHASAN

Variabel data penelitian yang digunakan pada penelitian ini dapat dilihat pada tabel 2 yakni sebagai berikut.

Tabel 2. Variabel data penelitian

No.	Atribut	Value
1	Umur	1) 20–35, 2) 36–45, 3) 46–55, 4) 56–65, 5) diatas 65
2	Jenis Kelamin	1.Pria, 2.Wanita
3	<i>Polyuria</i>	1.Ya, 2.Tidak
4	<i>Polydipsia</i>	1.Ya, 2.Tidak
5	<i>Suddenweight loss</i>	1.Ya, 2.Tidak
6	<i>Weakness</i>	1.Ya, 2.Tidak
7	<i>Polyphagia</i>	1.Ya, 2.Tidak
8	<i>Genital thrush</i>	1.Ya, 2.Tidak
9	<i>Visual blurring</i>	1.Ya, 2.Tidak
10	<i>Itching</i>	1.Ya, 2.Tidak
11	<i>Irritability</i>	1.Ya, 2.Tidak
12	<i>Delayed healing</i>	1.Ya, 2.Tidak
13	<i>Partial paresis</i>	1.Ya, 2.Tidak
14	<i>Muscle stiffness</i>	1.Ya, 2.Tidak
15	<i>Alopecia</i>	1.Ya, 2.Tidak
16	Obesitas	1.Ya, 2.Tidak
17	Kelas	1.Positif 2.Negatif

Dari Tabel 2 Ada 16 variabel dataset gejala dan 1 variabel class penentu klasifikasi. Selain itu, dari total 520 data, data tersebut dibagi menjadi dua dengan persentase 80% pelatihan dan 20% pengujian. Jadi data latih berjumlah 416 data dan data uji berjumlah 104 data. Dari data yang diperoleh, pohon keputusan dilatih oleh algoritma ID3 untuk membuat alur keputusan. Penerapan dan pengujian dataset dilakukan pada rapidminer. Berikut ini merupakan proses klasifikasi penderita biabetes menggunakan *RapidMiner 9.7*:



Gambar 2 Proses klasifikasi pada *RapidMiner 9.7*

Langkah pertama dalam penelitian ini adalah menyiapkan kumpulan data kemudian membagi data tersebut menjadi data uji dan data latih dengan perbandingan 80% untuk data latih dan 20% untuk data uji. Selanjutnya, klasifikasi dilakukan dengan menggunakan algoritma ID3 dan akurasi dihitung dengan menggunakan *confussion matrix*. Dari hasil pengujian menghasilkan *performance vector* dan tabel *Confusion Matrix* seperti dibawah ini:

Tabel 3. Performance Vector

PerformanceVector

```

PerformanceVector:
accuracy: 98.08%
ConfusionMatrix:
True:  Positive      Negative
Positive:    62       0
Negative:    2        40
precision: 95.24% (positive class: Negative)
ConfusionMatrix:
True:  Positive      Negative
Positive:    62       0
Negative:    2        40
recall: 100.00% (positive class: Negative)
ConfusionMatrix:
True:  Positive      Negative
Positive:    62       0
Negative:    2        40
AUC (optimistic): 1.000 (positive class: Negative)
AUC: 0.500 (positive class: Negative)
AUC (pessimistic): 0.969 (positive class: Negative)
    
```

Tabel 4. Hasil Akurasi

Accuracy: 98.08%

	True Positive	True Negative	Class Precision
Pred. Positive	62	0	100.00 %
Pred. Negative	2	40	95.24 %
Class recall	96.88%	100.00%	

Dari tabel diatas bisa dilihat bahwa pengujian menghasilkan akurasi yang cukup besar yaitu 98.08 % *Class Precision pred. positive* sebesar 100% % dan *pred. negative* sebesar 95,24%, sedangkan *Class Recall true positive* sebesar 96,88% dan *true negative* sebesar 100%.

KESIMPULAN

Penelitian ini menggunakan Algoritma klasifikasi *Decision tree* ID3 untuk pengklasifikasian seseorang terkena penyakit diabetes atau tidak. Dari 520 data dibagi menjadi 416 sebagai data *training* dan 104 sebagai data *testing*. Dari tabel diatas bisa dilihat bahwa pengujian menghasilkan akurasi yang cukup besar yaitu 98.08 % *Class Precision pred. positive* sebesar 100% % pred negative 95,24%, sedangkan *Class Recall true positive* sebesar 96,88% dan *true negative* 100%.

DAFTAR PUSTAKA

- [1] Khairani, “Pengetahuan Diabetes Mellitus Dan Upaya Pencegahan Pada Lansia Di Lam Bheu Aceh Besar,” *Pengetah. Diabetes Mellit. Dan Upaya Pencegah. Pada Lansia Di Lam Bheu Aceh Besar*, vol. 3, no. 3, pp. 58–66, 2012.
- [2] Nurlina, “Jurnal Media Keperawatan : Politeknik Kesehatan Makassar Jurnal Media Keperawatan : Politeknik Kesehatan Makassar,” *J. Media Keperawatan Politek. Kesehat. Makassar*, vol. 10, no. 01, pp. 59–66, 2019.
- [3] ADA, “Classification and Diagnosis of Diabetes Mellitus,” *Stand. Med. Care Diabetes*, vol. 39, no. January, 2016, doi: 10.1016/B978-0-323-18907-1.00038-X.
- [4] Suwarno and AA Abdillah, “Penerapan Algoritma Bayesian Regularization Backpropagation Untuk Memprediksi Penyakit Diabetes,” *J. MIPA*, vol. 39, no. 2, pp. 150–158, 2017.
- [5] H. Rahmawan and A. SN2, “Penentuan Rekomendasi Pelatihan Pengembangan Diri Bagi Pegawai Negeri Sipil Menggunakan Algoritma C4 . 5 Dengan Principal Component Analysis,” *J. TEKNO KOMPAK*, vol. 14, no. 1, pp. 5–10, 2020.
- [6] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*, vol. 9780470908. 2014.
- [7] P. P. Putra and A. S. Chan, “Pengembangan Aplikasi Perhitungan Prediksi Stock Motor Menggunakan Algoritma C 4.5 Sebagai Bagian dari Sistem Pengambilan

- Keputusan (Studi Kasus di Saudara Motor),” *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 24, 2018, doi: 10.35314/isi.v3i1.296.
- [8] M. Kamber and J. Han, *Data Mining: Concepts and Techniques : Concepts and Techniques*. 2018.
- [9] M. Yusa, E. Utami, and E. Luthfi. Taufiq, “Evaluasi Performa Algoritma Klasifikasi Decision Tree Id3 ,” *InfoSys J.*, vol. 4, no. 1, pp. 23–34, 2016, [Online]. Available: <http://e-journal.potensi-utama.ac.id/ojs/index.php/INFOSYS/article/view/136>.
- [10] F. M. Hana, “Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4 . 5,” 2020.
- [11] L. Irawan, L. H. Hasibuan, and F. Fauzi, “Analisa Prediksi Efek Kerusakan Gempa Dari Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data Mining Orange,” *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 14, no. 2, pp. 189–201, 2020, doi: 10.47111/jti.v14i2.1079.
- [12] F. Hafidh, M. Y. Kurniawan, and R. I. Yazidah Anwar, “Identifikasi Ketunaan Anak Berkebutuhan Khusus dengan Algoritma Iterative Dichotomiser 3 (ID3),” *J. Buana Inform.*, vol. 12, no. 2, pp. 78–87, 2021, doi: 10.24002/jbi.v12i2.4488.
- [13] R. Dita Merawati, “Penerapan Data Mining Penentu Minat Dan Bakat Siswa Smk Dengan Metode C4 . 5,” *J. Algor*, vol. 1, no. 1, pp. 28–37, 2019.
- [14] F. Gorunescu, *Data Mining: Concepts, Models, and Techniques*. Springer, 2011.
- [15] J. R. Quinlan, “Learning decision tree classifiers,” *ACM Comput. Surv.*, vol. 28, no. 1, pp. 71–72, 1996, doi: 10.1145/234313.234346.