



Penerapan Algoritma C4.5 Untuk Menentukan Tingkat Kemampuan Siswa Mengikuti Media Pembelajaran Google Classroom (Studi kasus : UPT SDN Panggungrejo 01)

Mochamad Fahmi Ubaidillah¹, Abdi Pandu Kusuma², Udkhiati Mawaddah³

^{1,2,3}Teknik Informatika, Universitas Islam Balitar. Jl. Majapahit No.2- 4 Blitar, Indonesia

email:Mochamadfahmi07@gmail.com¹,

pans.uib1blitar@gmail.com²,udkhiati.mawaddah@gmail.com³

ABSTRAK

Kata Kunci:

Siswa
Kemampuan Siswa
Algoritma C4.5
Pohon Keputusan
RapidMiner

Teknologi sekarang dilihat penting bagi keberadaan manusia karena membantu berbagai tanggung jawab, termasuk yang berkaitan dengan pekerjaan dan pendidikan. Pembelajaran di UPT SDN Panggungrejo 01, pembelajaran tidak hanya dilakukan tatap muka tetapi juga dilakukan pembelajaran dengan menggunakan media pembelajaran *google classroom*. Permasalahan yang terjadi di UPT SDN Panggungrejo 01 terdapat kendala yaitu koneksi *smartphone* yang bisa membuat proses pembelajaran tidak optimal. Dari masalah tersebut perlu mengetahui tingkat kemampuan siswa. *Data mining* dapat digunakan untuk menyelesaikan masalah klasifikasi data yang besar, yang akan membantu dalam pengambilan keputusan. Pada penelitian ini menerapkan algoritma C4.5 untuk melakukan klasifikasi kemampuan siswa, adanya penerapan algoritma ini diharapkan dapat mengevaluasi guru dan meningkatkan efektivitas dalam pembelajaran menggunakan aplikasi Pembelajaran. Hasil klasifikasi dengan algoritma C4.5 mendapatkan 12 rules dengan mendapatkan 9 keputusan "mampu" dan 3 keputusan "tidak mampu". Hasil uji menggunakan *rapidminer* keputusan yang terbentuk hampir sama dengan perhitungan yang dilakukan dengan manual. Perbedaannya hanya pada satu penambahan daun di keputusan mampu predikat baik pada atribut Ilmu Pengetahuan Alam & Sosial yang dihasilkan oleh *rapidminer*.

ABSTRACT

Keyword:

Students
Students Proficiency
C4.5 Algorithm
Decision Tree
Rapidminer

Technology is now seen as essential to human existence as it assists in a variety of responsibilities, including those related to work and education. Learning at UPT SDN Panggungrejo 01, learning is not only carried out face to face but also learning is carried out using *google classroom* learning media. The problem that occurs at UPT SDN Panggungrejo 01 is that there is an obstacle, namely the *smartphone* connection which can make the learning process not optimal. From these problems it is necessary to know the level of ability of students. *Data mining* can be used to solve large data classification problems, which will assist in making decisions. The results of the classification with the C4.5 algorithm get 12 rules by getting 9 "able" decisions and 3 "unable" decisions. The results of the test using *rapidminer*, the decisions that are formed are almost the same as the calculations done manually. The difference is only one additional leaf in the decision to be able to get a good title on the Natural & Social Sciences attributes produced by *rapidminer*.

PENDAHULUAN

Bidang pendidikan sangat dipengaruhi oleh kemajuan teknologi digital. Para tenaga pendidikan selalu berusaha melakukan inovasi-inovasi baru, misalnya dalam kaitannya dengan bahan ajar dan kebutuhan instruktur. Materi pembelajaran berbasis internet pada dasarnya adalah pembelajaran tatap muka yang didukung oleh sumber belajar daring. Seiring dengan kemajuan tersebut, pembelajaran daring juga mengalami berbagai kemajuan; itu adalah salah satu ciri dari transisi digital saat ini di bidang pendidikan

Teknologi sekarang dipandang penting bagi keberadaan manusia karena membantu berbagai tanggung jawab, termasuk yang berkaitan dengan pekerjaan dan pendidikan. Siswa dapat menggunakan teknologi sebagai sarana pembelajaran atau mediator untuk bertukar ilmu dengan siswa lain menggunakan berbagai saluran, seperti grup WhatsApp, Zoom, *Google classroom*, dan Google Meeting. Agar siswa tetap terlibat dan bersemangat dalam kegiatan belajar mengajar ini, pendidik dapat menggunakan media pembelajaran yang disebutkan di atas untuk memberikan penjelasan mata pelajaran yang menarik dan tidak membosankan

Pentingnya bagi UPT SDN Panggungrejo 01 dalam untuk bahan evaluasi guru dalam melaksanakan pengembangan pembelajaran menggunakan media *google classroom* dalam bidang Pendidikan. Mengetahui tingkat kemampuan siswa mengikuti media pembelajaran menggunakan aplikasi *google classroom* di UPT SDN Panggungrejo 01. Media *online* memainkan peran penting dalam pembelajaran siswa. Maka dari itu penelitian ini menerapkan Algoritma C4.5 dalam menentukan tingkat kemampuan siswa mengikuti media pembelajaran *google classroom*. Adanya penerapan algoritma ini diharapkan dapat mengevaluasi guru dan meningkatkan efektivitas dalam pembelajaran menggunakan aplikasi Pembelajaran.

METODE

Data Mining

Data mining adalah bidang teknologi pembelajaran komputer (*Machine Learning*) di mana satu atau lebih pekerjaan ditangani. Teknik ini berguna untuk analisis pengetahuan otomatis. Pembelajaran induktif adalah jenis pemahaman berbeda yang berkembang dari pengamatan terkait, seperti konsep yang dipelajari sebelumnya. *Data mining* merupakan langkah dalam proses *Knowledge Discovery in Database* (KDD), yang mengacu pada penerapan metodologi ilmiah dalam *data mining*. Sebuah teknik untuk mengumpulkan pengetahuan dari database yang ada dikenal sebagai penemuan pengetahuan dalam database (KDD). Ada beberapa tahapan KDD yaitu *Selection Data*, *Pre-Processing/Cleaning*, *Transformation Data*, *Data mining* dan Evaluasi. Ada tabel dalam database yang berinteraksi satu sama lain dan terkait. Informasi yang dihasilkan oleh metode yang disebutkan di atas dapat digunakan sebagai dasar keahlian untuk membuat keputusan [1]. *Data mining* adalah suatu metode yang dapat menghasilkan informasi dalam bentuk *database* dan dapat

meningkatkan Pengetahuan yang diperoleh dari data yang besar, yang akan membantu dalam pengambilan keputusan.

Klasifikasi

Sesuai dengan jumlah kelas yang dibutuhkan, metode pemrosesan data yang disebut klasifikasi mengkategorikan apapun ke dalam kelas yang berbeda [2]. Untuk membuat kategori untuk hal-hal yang labelnya tidak diketahui, klasifikasi juga memerlukan pencarian model atau fungsi yang menyampaikan atau membedakan konsep atau data kategorikal [3].

Algoritma C4.5

Algoritma C4.5 atau pohon keputusan adalah metode klasifikasi *Data mining* yang umum digunakan. Seperti disebutkan sebelumnya, Metode ini banyak disukai karena dapat mengkategorikan dan menggambarkan bagaimana sifat-sifat berhubungan satu sama lain. Algoritma C4.5 adalah salah satu dari banyak teknik yang dapat diterapkan pada konstruksi pohon keputusan. Data numerik dan diskrit dapat ditangani dengan algoritma C4.5 Gain ratio digunakan oleh algoritma C4.5. Nilai informasi dalam satuan bit dari kumpulan objek pertama-tama harus dihitung untuk menentukan rasio penguatan, khususnya yang melibatkan konsep entropi [4].

A. Konsep Entropi

Entropi (S) adalah ukuran berapa banyak bit yang dianggap diperlukan dalam ruang sampel S dapat digunakan mengekstrak kelas (+ atau -) dari beberapa data acak. Jumlah bit yang diperlukan untuk menunjuk suatu kelas disebut sebagai entropi. Lebih banyak ekstraksi digunakan semakin rendah nilai entropi.

Satu jenis *orisinalitas* (ketidakaslilan) S dievaluasi menggunakan entropi. Dalam ruang sampel S, entropi ditentukan pada persamaan 2.1 sebagai berikut :

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2.1)$$

Keterangan :

S : himpunan kasus

N : jumlah partisi S

P_i : proporsi dari S_i terhadap S

B. Konsep Gain

Selama *atribut* A cukup untuk memenuhi data S, mendapatkan informasi menghasilkan gain (S,A). Keuntungan mengacu pada informasi yang didapat dari data output atau variabel dependen S yang dikumpulkan berdasarkan *atribut* A. (S,A).

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.2)$$

Keterangan :

S : Himpunan kasus

A : *Atribut*

N : Jumlah partisi *atribut* A

$|S_i|$: Jumlah kasus pada partisi ke i

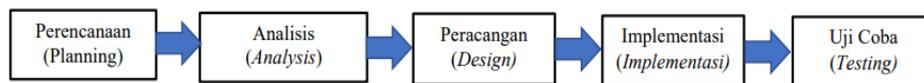
$|S|$: Jumlah kasus dalam S

Menerapkan kumpulan data yang besar dapat dibagi menjadi sekumpulan catatan yang lebih mudah dikelola dengan menggunakan teknik kategorisasi aturan kumpulan keputusan pemodelan ke dalam pohon keputusan. *Node* internal pohon keputusan mewakili *atribut* yang diuji, setiap cabang untuk hasil pengujian *atribut*, dan *node* daun untuk kelas tertentu. Dalam sebuah pohon keputusan terdapat 3 jenis *node*, yaitu :

- a. Simpul akar adalah simpul paling atas dan tidak memiliki input maupun output.
- b. *Node* internal adalah cabang dari *node* dengan minimal dua output dan satu input.
- c. *Node* akhir yang hanya memiliki satu masukan dan tidak ada keluaran dikenal sebagai *node* terminal atau *node* daun [5].

Rancangan Penelitian

Pada rancangan penelitian ini penulis menguraikan metodologi dan kerangka penelitian kerja yang digunakan dalam menyelesaikan masalah peneliti. Rancangan penelitian dapat dilihat dalam rancangan flowchart pada gambar 1.



Gambar 2.1 Rancangan Penelitian

1. Perencanaan (*Planning*)

Pada tahap perancangan, pada tahap ini peneliti mencari ruang lingkup masalah yang akan dijadikan penelitian dengan melakukan observasi awal guna mendapatkan masalah yang relevan untuk diteliti, kemudian menentukan judul dan merumuskan tujuan penelitian, serta menetapkan batasan penelitian untuk menjaga fokus penelitian agar tetap pada kaidah masalah yang telah ditetapkan sebelumnya.

2. Analisis (*Analysis*)

Untuk mendapatkan beberapa teori yang digunakan sebagai bahan referensi untuk memecahkan masalah yang ada di penelitian ini, diperlukan kajian studi literatur yang diperoleh dari buku - buku ilmiah, jurnal - jurnal ilmiah serta artikel lainnya sebagai pedoman untuk melakukan penelitian yang sesuai dengan permasalahan agar lebih terarah.

3. Perancangan (*Design*)

Pada tahap perancangan, peneliti mulai dengan menentukan kriteria *atribut* apa saja yang diperlukan untuk proses perhitungan. Data ini diperoleh dari arsip data nilai semester ganjil

tahun 2022 Serta data diperoleh dari wawancara yang dilakukan dengan guru kelas UPT SDN Panggungrejo 01. Berikut penjabarannya:

- a. *Range* nilai. 0-70 kurang, 71-80 Cukup, 81-100 Baik
- b. Label atribut. Keterangan (>80) maka mampu jika (<80) maka belum mampu

4. Implementasi (*Impelementiton*)

Pada tahapan ini melakukan pengolahan data dari hasil pengumpulan data dan kemudian melakukan perhitungan menggunakan algoritma C4.5 dengan menghitung nilai gain dan entropy dari data yang telah diklasifikasi. Setelah melakukan perhitungan, maka akan didapatkan nilai entropy dan gain pada masing– masing atribut. Kemudian menentukan nilai gain yang tertinggi untuk dijadi node pertama atau awal yang akan digunakan sebagai dasar untuk pembentukan mdoel pohon keputusan.:

5. *Testing* (Pengujian)

Pada tahap ini dilakukan pengujian hasil dari perhitungan algoritma C4.5 dan pembuatan pohon keputusan. Pengujian tersebut menggunakan *software RapidMiner* yang bertujuan untuk mengetahui kesesuaian antara model pohon keputusan menggunakan perhitungan manual dengan model pohon keputusan yang dihasilkan oleh *software RapidMiner*

HASIL DAN PEMBAHASAN

Pada penelitian ini peneliti menerapkan Data mining untuk melakukan menentukan tingkat kemampuan siswa dalam mengikuti media pembelajaran google classroom menggunakan algoritma C4.5. Dalam proses pengumpulan data diperoleh sebanyak 79 data siswa dengan *range* nilai jika 0-70 = kurang, 71-80 = cukup dan 81-100 = baik dan untuk menentukan keputusan mampu (> 80) dan belum mampu (<80) dari hasil koordinasi dengan tempat penelitian, kemudian data yang terkumpul selanjutnya akan diproses untuk dilakukan penerapan data mining menggunakan perhitungan dengan algoritma C4.5. Tahap pengujian menggunakan RapidMiner, kemudian dijalankan sehingga menghasilkan nilai untuk mengukur tingkat keberhasilan penerapan algoritma C4.5 dalam menentukan tingkat kemampuan siswa. Berikut data sampel dari UPT SDN Panggungrejo 01 pada

Tabel 1.

Tabel 1 Sampel data UPT SDN Panggungrejo 01

No	Bahasa Indonesia	PAI	PPKN	Ilmu pengetahuan Alam & Sosial	Nilai Rata- rata
1	79	79	83	81	80.5
2	80	77	82	79	79.5
3	82	87	84	85	84.5
4	82	87	84	85	82
5	78	77	80	81	78
6	81	79	83	78	81
..
..
75	80	91	80	80	82.75

76	80	81	81	81	80.75
77	79	79	80	79	79.25
78	80	82	81	84	81.75
79	80	78	81	79	79.5

3.1. Proses Pembahasan Algoritma C4.5

Perhitungan algoritma C4.5 dimulai dengan memilih atribut akar terlebih dahulu dengan mencari jumlah kasus keseluruhan, jumlah kasus keputusan mampu dan jumlah kasus keputusan Belum mampu. Menghitung entropy dari semua kasus yang dibagi berdasarkan mata pelajaran atribut PAI, PPKN, Ilmu Pengetahuam Alam & Sosial, Bahasa Indoensia. Setelah itu dilakukan perhitungan gain untuk masing – masing atribut. Hasil perhitungan ditunjukkan oleh tabel berikut ini.

Tabel 2 Perhitungan Node 1

Node 1		Jumlah (S)	Mampu	Belum Mampu	Entropy	Gain
Atribut	Total	79	44	35	0.990617	
Bahasa Indonesia						0.129589
	Baik	23	20	3	0.558629	
	Cukup	56	24	32	0.985228	
	Kurang	0	0	0	0	
PAI						0.196509
	Baik	34	29	5	0.602431	
	Cukup	45	16	29	0.938932	
	Kurang	0	0	0	0	
PPKN						0.299029
	Baik	41	35	6	0.600609	
	Cukup	38	9	29	0.789749	
	Kurang	0	0	0	0	
Ilmu Pengetahuam Alam & Sosial						0.224138
	Baik	30	27	3	0.468996	
	Cukup	49	18	31	0.948613	
	Kurang	0	0	0	0	

Perhitungan entropy total, entropy atribut dan gain.berikut adalah perhitungannya.

$$Entropy\ total = (-44/79)*Log_2(44/79)+(-35/79)*log_2(35/79)$$

$$Entropy\ total = 0.990617$$

Kemudian menghitung nilai gain pada baris atribut Bahasa Indonesia sebagai berikut.

$$Gain = (0.990617) - ((23/79) \times (0.558629) + (56/79) \times (0.985228))$$

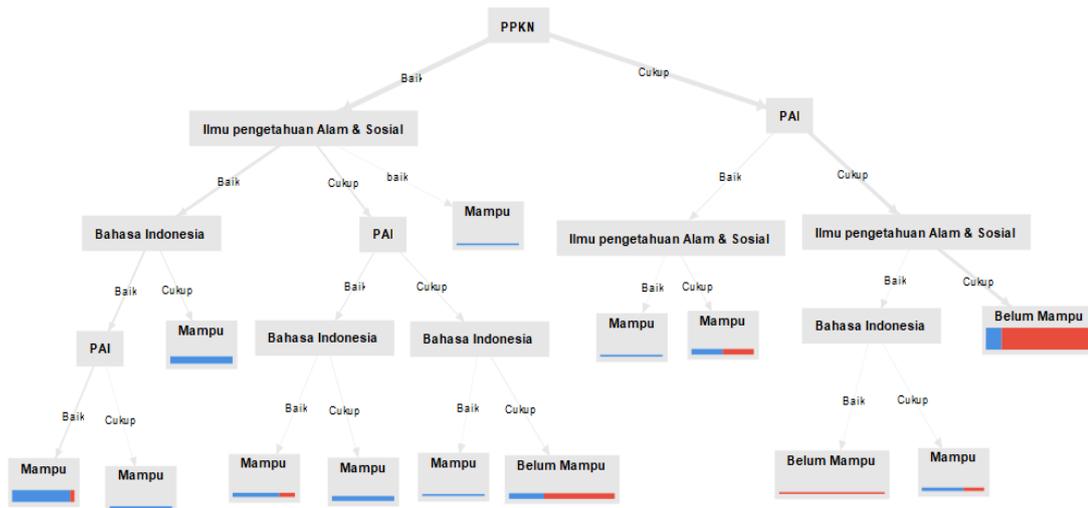
$$Gain = 0,129589$$

Dari hasil perhitungan pada tabel 1 diperoleh atribut yang menjadi node (akar) adalah atribut PPKN memiliki gain tertinggi yaitu 0.299029, dimana terdiri dari 3 predikat atribut yaitu Baik, Cukup, dan Kurang. Berdasarkan nilai entropy dari ketiga predikat atribut diatas, predikat atribut

- i) Jika PPKN = Cukup, PAI = Baik dan Ilmu pengetahuan Alam & Sosial = Cukup maka nilai keputusan “Mampu”
- j) Jika PPKN = Cukup, PAI = Cukup, Ilmu pengetahuan Alam & Sosial = Baik dan Bahasa Indonesia = Baik maka nilai keputusan “Belum Mampu”
- k) Jika PPKN = Cukup, PAI = Cukup, Ilmu pengetahuan Alam & Sosial = Baik dan Bahasa Indonesia = Cukup maka nilai keputusan “Mampu”
- l) Jika PPKN = Cukup, PAI = Cukup dan Ilmu pengetahuan Alam & Sosial = Cukup maka nilai keputusan “Belum Mampu”

Pengujian Algoritma C4.5 dengan RapidMiner

Setelah analisis kemampuan siswa terhadap media pembelajaran *Google Classroom* di UPT SDN Panggungrejo 01 menggunakan algoritma C4.5 dihitung dan pembuatan pohon keputusan menggunakan perhitungan manual, maka tahap berikutnya adalah menguji dari analisis dan perhitungan tersebut menggunakan aplikasi RapidMiner. *Rapidminer* ini digunakan untuk pengujian algoritma pohon keputusan apakah sama atau tidak dengan hasil perhitungan manual Adapun hasil dari pembahasan pohon keputusan menggunakan aplikasi RapidMiner dapat dilihat pada gambar 3.



Gambar 3. Pohon Keputusan Hasil Analisis



Gambar 4. Hasil Rule Pohon Keputusan Algoritma C4.5

Berdasarkan pengolahan data menggunakan software RapidMiner didapat Setelah dilakukan perhitungan dan pengujian data pada masing-masing atribut dengan algoritma C4.5, maka akan didapatkan pola pohon keputusan akhir. Dari gambar dapat terlihat bahwa pohon keputusan hasil analisis menggunakan aplikasi RapidMiner hampir sama dengan hasil analisis dengan perhitungan manual. Dengan menggunakan *software RapidMiner* dapat kita ambil sebuah kesimpulan bahwa hasil pengujian sangat baik karena rule yang dihasilkan hampir sama. Perbedaannya hanya pada satu penambahan daun keputusan mampu predikat baik di atribut Ilmu Pengetahuan Alam & Sosial yang dihasilkan oleh *RapidMiner*.

SIMPULAN

Algoritma C4.5 dapat digunakan sebagai metode klasifikasi dan pengambilan keputusan dalam menentukan tingkat kemampuan siswa dengan memperhatikan model pohon keputusan dari empat atribut yang terbentuk dari rules berdasarkan hasil perhitungan Algoritma C4.5 dengan memperhatikan nilai Gain yang paling tinggi dari atribut yang dipakai yaitu root node atribut PPKN dengan nilai gain sebesar 0.103196. Dari Hasil Pengujian dengan *rapidminer* dengan aplikasi *rapidminer* menghasilkan *rule* pohon keputusan yang terbentuk hampir sama yaitu menghasilkan 12 rules dengan 9 keputusan mampu dan 3 keputusan belum mampu, sedikit ada penambahan satu daun di predikat baik keputusan mampu pada atribut Ilmu Pengetahuan Alam & Sosial yang dihasilkan oleh *rapidminer*. Maka pengujian tersebut mampu melakukan klasifikasi berdasarkan mata pelajaran yang ditentukan. oleh karena itu penggunaan *rapidminer* juga sangat membantu dalam pengujian yang dilakukan.

DAFTAR RUJUKAN

- [1] Mardi, Y. (2017). *Data mining : Klasifikasi Menggunakan Algoritma C4.5*. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- [2] Arifin, M. F., & Fitriyah, D. (2018). Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada. *IncomTech*, 8(2), 87–102. <https://doi.org/10.22441/incomtech.v8i1.2198>
- [3] Febriani, S., & Sulistiani, H. (2021). Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5. *89Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(4), 89–95.
- [4] Riandari, F., & Simangunsong, A. (2019). Penerapan algoritma c4. 5 untuk mengukur tingkat kepuasan mahasiswa. *Terakreditasi Dikti*, 3(2), 1-7.
- [5] Yunus, M. F., Santi, I. H., & Kirom, S. (2022). PENERAPAN ALGORITMA C4. 5 UNTUK MENENTUKAN KUALITAS TELUR AYAM HIBRIDA. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 908-913.