

Model Prediksi Kelulusan Peserta Pelatihan Online Menggunakan Algoritma Decision Tree C4.5

Abdul Rohman¹, Sri Mujiyono²

^{1,2}S1 Teknik Informatika, Universitas Ngudi Waluyo

Email: abdulrohman15@gmail.com¹, mujiyn80@gmail.com²

Abstrak— Pelatihan online metode pembelajaran yang sangat efektif untuk menghindari penyebaran covid-19 sehingga menghindari terjadinya kerumunan masa. Salah satu pelatihan yang diselenggarakan oleh e-guru.id dan yang diminati oleh para pengajar yaitu "Membuat video tutorial menggunakan camtasia" dengan jumlah peserta 261 orang. Dengan banyak peserta tersebut diperlukan evaluasi untuk mengetahui model prediksi kelulusan yang dapat memberikan informasi terbaru untuk sebuah kebijakan organisasi. Maka dalam penelitian ini mengolah dataset kelulusan peserta pelatihan menggunakan metode klasifikasi datamining dengan algoritma c4.5. Dan penelitian ini menghasilkan bahwa atribut tugas sangat berpengaruh terhadap kelulusan peserta disamping atribut keaktifan peserta dengan nilai akurasi 98,75 % dan performance klasifikasi sangat baik dengan nilai 0,990.

Kata kunci: Prediksi, Kelulusan Peserta, Decision Tree, Algoritma C4.5

I. PENDAHULUAN

Pelatihan online atau daring merupakan metode pembelajaran yang sangat efektif untuk mematuhi protokol Kesehatan pada masa pandemi covid-19 ini, untuk menghindari terjadinya kerumunan masa [1].

Metode pelatihan atau pembelajaran dapat menerapkan aplikasi forum, meeting, chat untuk menunjang tercapainya tujuan pembelajaran yang ditentukan [2].

Salah satu organisasi yang dapat memberikan fasilitas pelatihan online kepada guru, dosen, tenaga pendidik dan umum yaitu e-guru.id yang berdiri tahun 2019. Dan pelatihan yang di minati oleh pengajar dan tenaga pendidik yaitu "Pelatihan Membuat Video Tutorial Menggunakan Camtasia", dengan jumlah peserta 261 orang yang dilakukan 36 jam dalam 6 hari.

Untuk mengukur atau megevaluasi peserta pelatihan tersebut dilakukan standar kelulusan dengan kriteria jumlah kehadiran sebanyak 75% dan pengumpulan serta nilai tugas harus dikumpulkan, sehingga peserta akan dinyatakan status lulus dan tidak lulus.

Dengan banyaknya peserta pelatihan tersebut, maka diperlukan pengolahan dataset kelulusan peserta dengan menggunakan datamining untuk mendapatkan informasi baru dalam hal prediksi [3].

Algoritma decision tree C4.5 merupakan metode klasifikasi datamining yang banyak digunakan dan efektif untuk memprediksi dengan pola pohon keputusan [4][5][6]. Maka dalam penelitian ini menggunakan algoritma c4.5 untuk memprediksi kelulusan peserta pelatihan.

II. TINJAUAN PUSTAKA

1. Kelulusan

Untuk mencapai tujuan pembelajaran dalam proses pembelajaran baik secara luring maupun daring Capaian

Pembelajaran Lulusan dengan kriteria sikap, pengetahuan dan keterampilan [7].

Sikap dalam proses pelatihan bisa dilihat dari aktifitas kehadiran dalam pertemuan materi, dan diskusi. Dan pengetahuan dan keterampilan dalam pelatihan dapat dilihat hasil tugas yang dikerjakan sehingga dapat melihat seberapa besar dalam kemampuannya.

2. Algoritma C4.5

Algoritma C4.5 merupakan metode klasifikasi datamining dengan ciri menghasilkan pola dan rule pohon keputusan untuk dijadikan suatu prediksi [8].

Ada 6 tahapan dalam algoritma decision tree C4.5 dalam menghasilkan pohon keputusan yaitu [8]:

1. Menyiapkan dataset
2. Menentukan nilai *entropy*

$$Entropy(i) = \sum_{j=1}^m f(i,j) \cdot 2 f[(i,j)] \quad (1)$$

3. Menentukan nilai *gain*:

$$Gain = - \sum_{i=1}^p \frac{n_i}{n} \cdot IE(i) \quad (2)$$

4. Menentukan nilai *Split Information*:

$$SplitInformation = - \sum_{i=1}^c \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (3)$$

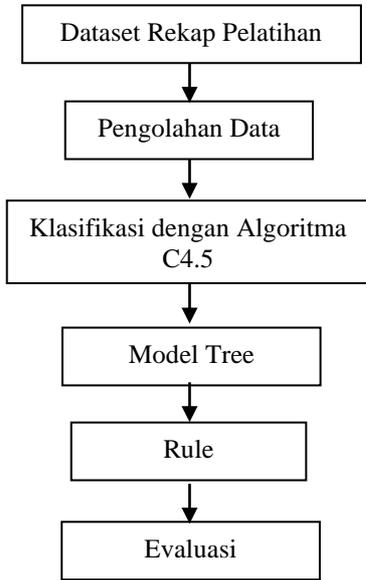
5. Selanjutnya menghitung gain ratio

$$Gainratio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)} \quad (4)$$

6. Mengulangi langkah kedua sehingga semua *record* terpartisi

III. METODE PENELITIAN

Untuk menghasil model prediksi kelulusan peserta pelatihan daring menggunakan algoritma decision tree C4.5 maka kerangka penelitiannya adalah sebagai berikut:



Gambar 1 Kerangka Pemikiran

Dataset yang digunakan dalam penelitian ini adalah rekap pelatihan “Membuat Video Tutorial Menggunakan Camtasia” yang berjumlah 261 orang dengan atribut 7 atribut yaitu:

1. Nama
2. Jenis Kelamin
3. Status Pekerjaan
4. Asal Institusi
5. jumlah kehadiran
6. Nilai tugas dan
7. keterangan kelulusan.

Tabel 1. Dataset Rekap Kelulusan

No	Nama	Jenis Kelamin	Pekerjaan	Asal Institusi	1	2	3	4	5	6	Tugas	Keterangan
1	A. Kowadi, S.Pd.	Laki-laki	Guru	SMPIT PERMATA BUNDA D	1	1	1	1	1	0	0	Tidak Lulus
2	A.M. Khaled, S.Pd. Gr	Laki-laki	Guru	SMK Negeri 1 Sumatera Bar	1	1	1	1	1	1	1	Lulus
3	Abdurrahman Arrauf, S.Pd. I	Laki-laki	Guru	SD Negeri Sjar 09 Kota Batu	1	1	1	1	1	1	1	Lulus
4	Achma Puranib, S.Si	Perempuan	Guru	SMA PGRI 2 Jombang	1	1	1	1	1	1	1	Lulus
5	Ade Fitria Susanti, S.Pd	Perempuan	Guru	MTsN 8 Muaro Jambi	1	1	1	1	1	1	0	Tidak Lulus
6	Adiatun Nimah, S.Pd	Perempuan	Guru	MIN 13 HSU	1	1	1	1	1	1	1	Lulus
7	Adrawati, S.Pd M.Pd	Perempuan	Guru	SMP Negeri 2 Mandau	1	1	1	1	1	1	1	Lulus
8	AGIAL ROHMANIA, S.Pd	Perempuan	Guru	MTsN 2 Tanjung Jabung Bar	1	1	1	1	1	1	1	Lulus
9	Agnes Periwati Heria, S. Pd.	Perempuan	Guru	SMK Islam Hang Tuah Batan	1	1	1	1	1	1	0	Tidak Lulus
10	Aglovia Firmayasa SE.MM	Perempuan	Dosen	Universitas Bina Sarana Infb	1	1	1	1	1	1	0	Tidak Lulus
11	Agus Susyowadi, S.Pd. M. Pd	Laki-laki	Guru	SD Negeri 2 Baturahno Darso	1	1	1	1	1	1	1	Lulus
12	AGUSTINA, S.Pd	Perempuan	Guru	SMP NEGERI 1 MANGGAP	1	1	1	1	1	1	0	Tidak Lulus
13	Ahmad Yani Ibnu Shidiq, S.Si	Laki-laki	Guru	SMP Negeri 3 Cirebon	1	1	1	1	1	1	1	Lulus
14	Alfath Rosyada Rokhim, M.Pd	Laki-laki	Guru	SMP Negeri 2 Sedati	1	1	1	1	1	1	0	Tidak Lulus
15	ALFIANI, S.Pd	Perempuan	Guru	MAN 1 ACEH UTARA	1	1	1	1	1	1	0	Tidak Lulus
16	Afdan Mardani, S.Pd	Perempuan	Guru	SMK Negeri 1 Jatuhur	1	1	1	1	1	1	1	Lulus
17	Alin Panji Maulani, S.Pd	Perempuan	Guru	SMK Al Jabbar Cileug	1	1	1	1	1	1	1	Lulus
18	Alphonsa Yulita Wole, S.Pd	Perempuan	Guru	SMPK Alvarez Paga	1	1	1	1	1	1	0	Tidak Lulus
19	Amimnu Annalyan	Perempuan	Guru	SDA Negeri 2 Bantul	1	1	1	1	1	1	1	Lulus
20	Andriyani, S. Pd	Perempuan	Guru	TK Negeri 3 Bonang	1	1	1	1	1	1	1	Lulus
21	Anis Mardiasuti, S.Pd	Perempuan	Guru	SD Negeri 2 Tamansariyan	1	1	1	1	1	1	0	Tidak Lulus
22	Anisah, S.Pd	Perempuan	Guru	SD Negeri 1 Barang 1 Malang	1	1	1	1	1	1	1	Lulus
23	Anita Darana, S.Pd	Perempuan	Guru	SDN 01 TAMPEH DHARMA S	1	1	1	1	1	1	1	Lulus
24	Anita Prasasa, S.Pd, MM	Perempuan	Guru	SMK Negeri 1 Probolinggo	1	1	1	1	1	1	1	Lulus
25	ANIWATI, S.Si	Perempuan	Guru	SMA N 1 GANTUNG	1	1	1	1	1	1	1	Lulus
26	Anna Eka Salmiah, S.Pd	Perempuan	Guru	SMP PGRI 3 Bogor	1	1	1	1	1	1	0	Tidak Lulus
27	Aprili, S.Pd	Perempuan	Guru	TK Nabelah Batem	1	1	1	1	1	1	1	Lulus
28	Aprilia Hasriana, S.Pd	Laki-laki	Guru	SDN 0110132 Wianggo	1	1	1	1	1	1	1	Tidak Lulus
29	Arif Sihotang, S.Pd	Perempuan	Guru	SMP Negeri 1 Satu Atap Topar	1	1	1	1	1	1	1	Lulus
30	Ary Febryanto, S.Pd.	Laki-laki	Guru	SD Negeri 1 Banyuworo	1	1	1	1	1	1	0	Tidak Lulus
31	Arisa Manti, S.T	Laki-laki	Guru	SMK Negeri 1 Kelasa	1	1	1	1	1	1	1	Lulus
32	ARIP RAHILAO DARMAWAN	Laki-laki	Guru	MAN 1 PESIRIR BARAT	1	1	1	1	1	1	1	Lulus
33	Ariwi Hidayati, S.Pd.	Perempuan	Guru	SMA MAARIF NU PANDAAN	1	1	1	1	1	1	1	Lulus
34	ARIP SARI PUDIN, S.Pd	Laki-laki	Guru	SDN 2 KERTAJAYA	1	1	1	1	1	1	1	Lulus
35	Arian Orlendo Lukas, S.Pd.	Laki-laki	Guru	SMP Negeri 2 SO Jember	1	1	1	1	1	1	0	Tidak Lulus
36	Asmah Dewita, W.Pd.	Perempuan	Guru	MTsN 1 Bukittinggi Sumatera	1	1	1	1	1	1	1	Lulus
37	ATIK OURA ATUL AINI, S.Pd	Perempuan	Guru	SMK KOSGORO NGANJUK	1	1	1	1	1	1	1	Lulus
38	B. Budi Binanto, S.Pd	Laki-laki	Guru	SMP PL Domenico Savio Ser	1	1	1	1	1	1	1	Lulus
39	Budi Lestari, S.Pd	Laki-laki	Guru	SD Negeri Karangbesuki 3 M	1	1	1	1	1	1	0	Tidak Lulus

Dan seterusnya

Untuk tahap pengolahan data dilakukan pengambilan atribut terpenting dan berpengaruh terhadap pembuatan model pohon keputusan.

Klasifikasi datamining menggunakan algoritma C4.5 terhadap dataset tersebut diatas menggunakan program atau aplikasi framework yaitu Rapid Miner yang cocok untuk penelitian datamining [9][10][11].

Setelah dilakukan eksperimen menggunakan program Rapidminer akan menghasil model pola pohon keputusan, dengan menggunakan algoritma c4.5 sehingga akan menghasilkan rule.

Dan tahap terakhir yaitu mengevaluasi hasil eksperimen yang dihasilkan dengan menggunakan Confusion Matrix dan ROC Curva, sehingga menghasilkan nilai akurasi [12].

IV. HASIL DAN PEMBAHASAN

Dataset rekap kelulusan pelatihan yang terdiri dari 7 atribut menjadi 5 atribut yaitu:

1. Jenis Kelamin
2. Status pekerjaan
3. Jumlah kehadiran
4. Nilai Tugas
5. Keterangan

Jenis kelamin memiliki nilai binominal yaitu; Laki-laki dan Perempuan, Status pekerjaan terdiri dari Guru, Dosen, Tenaga Pendidik (tendik), Jumlah kehadiran yaitu 1,2,3,4,5,6, nilai tugas yaitu 1 (untuk tugas yang memenuhi syarat), dan 0 (tugas yang tidak memenuhi syarat). Dan Atribut Keterangan yaitu Lulus dan Tidak Lulus.

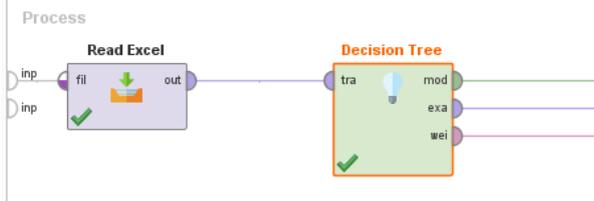
Keputusan lulus dan tidak lulus yaitu jika peserta kehadirannya minimal 75% dan tugas memenuhi syarat (1) maka peserta pelatihan dinyatakan lulus. Dan sebaliknya jika peserta dengan kehadiran 75% tapi tugas tidak memnuhi syarat atau tugas memenuhi syarat tapi kehadiran kurang dari 75% maka peserta dianggap tidak lulus.

Tabel 2 Dataset setelah pengolahan

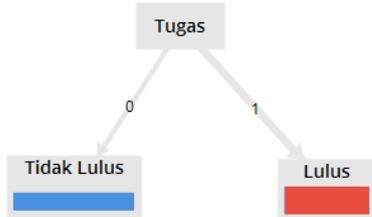
Row No.	Keterangan	Jenis Kelamin	Pekerjaan	Jumlah keh...	Tugas
1	Tidak Lulus	Laki-laki	Guru	5	0
2	Lulus	Laki-laki	Guru	6	1
3	Lulus	Laki-laki	Guru	6	1
4	Lulus	Perempuan	Guru	6	1
5	Tidak Lulus	Perempuan	Guru	6	0
6	Lulus	Perempuan	Guru	6	1
7	Lulus	Perempuan	Guru	6	1
8	Lulus	Perempuan	Guru	6	1
9	Tidak Lulus	Perempuan	Guru	6	0
10	Tidak Lulus	Perempuan	Dosen	6	0
11	Lulus	Laki-laki	Guru	6	1
12	Tidak Lulus	Perempuan	Guru	6	0
13	Lulus	Laki-laki	Guru	6	1
14	Tidak Lulus	Laki-laki	Guru	6	0

Dan seterusnya

Selanjutnya dilakukan eksperimen menggunakan algoritma C4.5 dengan maksimal depth 10, maka menghasilkan pohon keputusan sebagai berikut:



Gambar 2. Proses Klasifikasi Decision Tree Algoritma C4.5

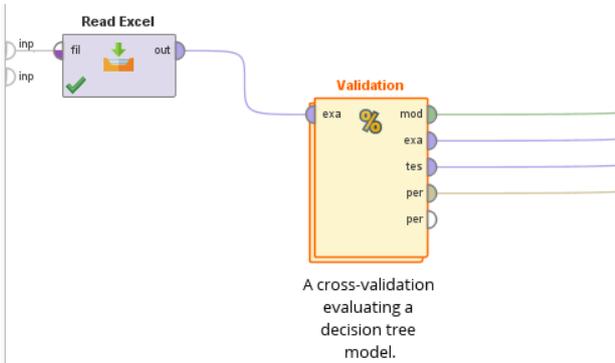


Gambar 3. Hasil Model Pohon Keputusan

Melihat gambar tersebut diatas, tujuan utama dari analisis data dengan algoritma c4.5 adalah menghasilkan pola atau rule [13], dalam hal ini yang didapat ada 2 yaitu:

1. Jika Tugas = 0 Maka peserta Tidak Lulus
2. Jika Tugas = 1 Maka peserta Lulus

Selanjutnya tahap terakhir untuk mengevaluasi model pohon keputusan terhadap kelulusan peserta pelatihan online dengan menggunakan Confusion Matrix dan ROC Curva.



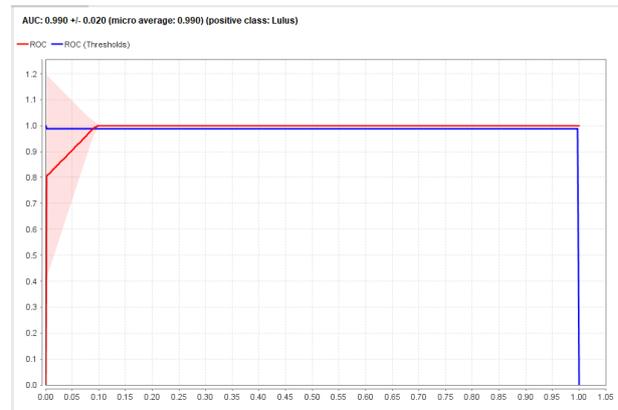
Gambar 4. Proses Evaluasi Model Decision Tree C4.5

Dari hasil evaluasi dengan menggunakan cross validation menghasilkan nilai akurasi 98,75% dengan confusion matrix sebagai berikut:

Tabel 3. Model Confusion Matrix

	True Tidak Lulus	True Lulus	Class Precisiom
Pred. Tidak Lulus	101	0	100%
Pred Lulus	2	158	98,75%
Class recall	98,06%	100%	

Dari tabel 3 terdapat grafik ROC dengan AUC (Area Under Curve) sebesar 0,990, sehingga klasifikasi tergolong sangat baik yaitu berada di sekitar 0,90 – 1.00 [8].



Gambar 5. Nilai AUC dalam Grafik ROC

V. KESIMPULAN

Dari hasil penelitian model prediksi kelulusan peserta pelatihan online dengan menggunakan algoritma C4.5 ini mendapat pola pohon keputusan dengan atribut tugas yang berpengaruh terhadap kelulusan peserta dengan nilai akurasi 98,75 % dan memiliki performance klasifikasi sangat baik dengan nilai 0,990.

VI. REFERENSI

- [1] Hadion Wijoyo, *Efektivitas Proses Pembelajaran Di Masa Pandemi*. Sumatra Barat: ICM Publisher, 2021.
- [2] M. Ngabidin, *Pemelajaran di Masa Pandemi, Inovasi Tiada Henti*. Yogyakarta: Deepublish, 2021.
- [3] D. T. Larose, *Data Mining Methods and Models*. 2006.
- [4] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," *J. Ultim.*, vol. 6, no. 1, pp. 15–20, 2014, doi: 10.31937/ti.v6i1.327.
- [5] M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Eeccis*, vol. 7, no. 1, pp. 59–64, 2013, doi: 10.1038/hdy.2009.180.
- [6] D. Himawan, "Aplikasi Data Mining Menggunakan Algoritma ID3 Untuk Mengklasifikasi Kelulusan Mahasiswa Pada Universitas Dian Nuswantoro Semarang," *Fak. Ilmu Komput.*, 2011.
- [7] Meda Yuliani, *Pemelajaran Daring Untuk Pendidik Teori dan Penerapan*. Sumatra Utara: Yayasan Kita Menulis, 2020.
- [8] F. Gorunescu, *Data Mining: Concepts, Models and Techniques (Intelligent Systems Reference Library)*. 2011.
- [9] Soni and Ganatra, "Categorazation of several

clustering algorithms from different perspective,” *J. Adv. Res. Comput. Sci.*, pp. 63–68, 2012.

- [10] F. Febrian, “Algoritma Klasifikasi Data Mining Pada Akseptasi Data Fakultatif,” *Program*, 2011.
- [11] B. Rifai, “Algoritma Neural Network Untuk Prediksi Penyakit Jantung,” *None*, vol. 9, no. 1, pp. 1–9, 2013.
- [12] C. Vercellis, *Business Intelligence: Data Mining and Optimization for Decision Making*. 2009.
- [13] I. H. Witten, “Data Mining Data Mining Complications : Overfitting Statistical modeling One attribute does all the work ?,” 2007.