



Implementasi Algoritma C4.5 Menentukan Pola Berangkatan Jamaah Haji

Fricles A Sianturi¹, Mina Kumari², Erlina Laian³

^{1,2}Jl. Iskandar Muda No.1, STMIK Pelita Nusantara Medan, Indonesia

E-mail : sianturifricles@gmail.com, minakumari@gmail.com, erlinalaian23@gmail.com

Abstract- Determination of the Hajj Pilgrim Departure Pattern is currently very much needed by the Haj organizers, especially at the Office of the Ministry of Religion Lubuk Pakam because the schedule of pilgrims' departure is still not regular, causing accumulation of data on pilgrims who will be departed to the holy land, but with The number of pilgrims who will depart for the pilgrimage is made to determine the pattern of Hajj departure with the C4.5 Algorithm to make it easier for the hajj organizer staff to determine what age should be sent to the holy land.

Keywords- C4.5 Algorithm, Determination of Hajj Pilgrims Departure Patterns

Abstrak- Penentuan Pola Keberangkatan Jama'ah Haji saat ini sangat dibutuhkan oleh pihak penyelenggara Haji Khususnya di Kantor Kementerian Agama Lubuk Pakam karena jadwal keberangkatan jama'ah haji masih belum teratur sehingga, menyebabkan penumpukan data jama'ah yang akan diberangkatkan ke tanah suci, akan tetapi dengan banyaknya jama'ah haji yang akan berangkat maka dibuatlah penentuan pola keberangkatan haji dengan Algoritma C4.5 untuk mempermudah para staff penyelenggara haji untuk menentukan usia berapakah yang seharusnya di berangkatkan ke tanah suci.

Kata kunci- Algoritma C4.5, Penentuan Pola Keberangkatan Jama'ah Haji

I. PENDAHULUAN

Para pendaftar ibadah haji di Kantor Kementerian Agama Lubuk pakam, calon jamaah haji diikut sertakan untuk mendaftar secara langsung diruang pendaftaran dengan memasukan (input) data informasi calon jamaah haji yang bersangkutan, baik berupa nama, jenis kelamin, umur, alamat, bank, dan data lainnya yang kemudian dimasukan dengan menggunakan Microsoft word dan Microsoft excel. Hal inilah yang menyebabkan antrian menjadi terlalu lama, karena cara menentukanjadwal keberangkatan haji dan mengisi data-data para calon jama'ah haji di Kantor Kemetrian Agama Lubuk Pakam bisa dibilang masi sangat manual. Oleh karena itu perlu adanya penentuan pola jadwal keberangkatan jama'ah haji dengan algoritma C4.5pada Kementerian Agama Lubuk Pakam.

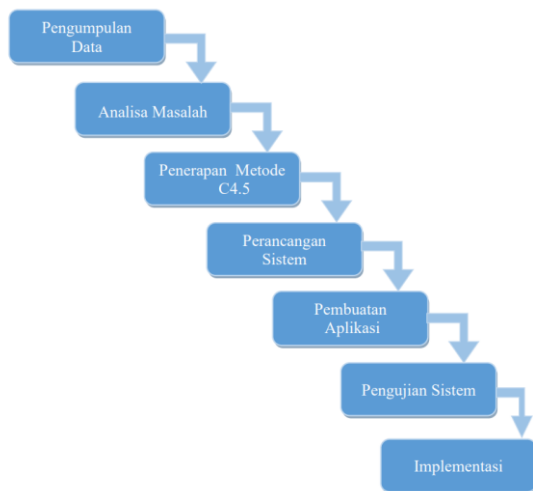
Kementerian Agama pada bidang penyelenggara haji (PHU) sebagai fasilitator dan organisator masyarakat tentu harus bersikap profesional dalam melayani para calon jamaah haji untuk mendampingi dan membantu segala sesuatu yang dibutuhkan dan dipersiapkan oleh para calon jamaah haji sebelum keberangkatan sampai kepulauan ibadah haji[1],[2]. Dalam hal ini yang perlu ditangani dan mendapat perhatian khusus adalah bidang pelayanan haji yang mengingat 12 semakin banyak peningkatan masyarakat mendaftar ibadah haji semakin meningkat pula antrian untuk mendaftar dan antrian

keberangkatan calon jama'ah haji di Kantor Kementerian Agama Lubuk Pakam[3],[4]

Metode C.45 yaitu sebuah algoritma yang digunakan untuk membangun decision tree (pohon keputusan) berdasarkan training data yang telah disediakan. Metode C4.5 dibuat oleh Ross Quinlan yang merupakan pengembangan dari ID3 yang juga dibuat oleh Quinlan (Quinlan, 1993) [5],[6]. Beberapa pengembangan yang dilakukan pada C4.5 adalah bisa mengatasi missing value, bisa mengatasi continue data, dan pruning [7].

Defenisi data mining adalah pembelajaran berbasis induksi (induction basedlearning) adalah proses pembentukan defenisi-defenisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep yang akan dipelajari. Knowledge Discovery in Database (KDD) adalah penerapan metode saintifikpada data mining [8],[9].

II. METODE PENELITIAN



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian yang telah digambarkan di atas, maka dapat diuraikan pembahasan masing-masing tahap dalam penelitian adalah sebagai berikut [10],[11]:

1. Pengumpulan Data
Dalam memperoleh data-data, beberapa metode telah dilakukan guna mendapatkan data yang sesuai dengan kebutuhan penelitian.
2. Analisa Masalah
Analisa masalah yang diperoleh selama penelitian pada Kementerian Agama Lubuk Pakam dari hasil observasi serta wawancara sehingga masalah masalah dalam penelitian mendapat solusi. Penelitian membahas dan menyampaikan uraian hasil penelitian mengenai menentukan antrian jadwal keberangkatan calon jama'ah ibadah haji dengan menggunakan Algoritma C4.5.
3. Perancangan Sistem
Pada perancangan sistem tahapan yang menggambarkan rancangan bangun sistem yang dibuat dengan metode UML (Unified Modelling Language). Perancangan sistem ini terdiri dari perancangan Use Case Diagram, Activity Diagram, Sequence Diagram, Class Diagram dan Flowchart.
4. Pembuatan Aplikasi
Pada tahap pembuatan aplikasi melakukan perancangan sistem mengenai sistem aktivitas yang ada dibangun sistem dengan pemrograman berbasis 41 WEB dengan bahasa pemrograman HTML (HyperText Markup Language), CSS (Cascading Style Sheets), PHP (Hypertext Preprocessor), J-Query, dengan menggunakan aplikasi Sublime Text, dan menggunakan Database MySQL, PhpMyAdmin.
5. Pengujian Sistem
Pada tahap pengujian data mining dalam menentukan antrian jadwal keberangkatan ibadah haji menggunakan C4.5. Proses pengujian membuktikan bahwa aplikasi telah sesuai dengan rancangan awal dari sistem yang telah dirancang dan telah memenuhi kebutuhan yang diharapkan.

6. Implementasi

Pada tahap Implementasi program diterapkan sesuai dengan tahap dimana sistem siap untuk digunakan di Kementerian Agama Lubuk Pakam untuk membantu menentukan antrian jadwal keberangkatan ibadah haji.

III. HASIL DAN PEMBAHASAN

Berdasarkan pengamatan, permasalahan ini dapat diatasi dengan cara melakukan pendataan kepada calon jama'ah haji di Kementerian Agama Lubuk Pakam dan penentuan pola keberangkatan dengan algoritma C4.5 di penyelenggara haji untuk mengatasi antrian jadwal keberangkatan jama'ah haji .

Sumber data pada penelitian ini di dapat dari Penyelenggara Haji di Kantor Kementerian Agama Lubuk Pakam. Contoh data yang digunakan yaitu data nama calon pendaftar haji dan data jadwal keberangkatan jama'ah haji, data tersebut digunakan sebagai sample analisa dan juga pengujian yang dapat dilihat pada table.

Tabel 1. Data Jemaah Haji

N o	Nama	Sisa Dana Haji	Pendidik an	Usi a	Berka s
1	sulaiman	Rp.4,339,000	S1	24	50
2	Amnah	Rp.4,339,000	S1	40	80
3	Fatmawaty	RP.2,956,000	S1	30	80
4	Siti hajar siregar	RP.4,475,700	S1	35	70
5	Sulaiman, ST	RP.4,475,700	S1	35	85
...
...
...
50	Nuraini Siregar	Rp.2,956,000	S1	35	50

Keterangan I(Persiapan Berkas):

- a. Jika Persiapan Berkas<60% maka keterangan "MINIMUM"
- b. Jika Persiapan Berkas>60% dan Persiapan Berkas<84% maka keterangan "MEDIUM"
- c. Jika Persiapan Berkas>84% maka keterangan "MAXIMAL"



- a. Jika USIA < 50 maka keterangan “Paruh Baya”
- b. Jika USIA > 31 dan USIA < 50 maka keterangan “Tua”
- c. Jika USIA > 30 maka keterangan “TUA”
- a. Jika Sisa Dana Haji < 3.000.000 maka keterangan “RENDAH”
- b. Jika Sisa Dana Haji > 3.000.000 dan Sisa Dana Haji < 4.000.0000 maka keterangan “TUA”
- c. Jika Sisa Dana Haji > 4.000.000 maka Keterangan “TINGGI”

Untuk mengetahui kategori *Keberangkatan Haji* pada data Jemaah Haji, dapat dilihat dari Tabel 4.4 *learning dataset II* (kategorikal *Keberangkatan Haji*) berikut ini:

Tabel 2. Learning Data Set II (Kategorial Keberangkatan Haji)

N	Nama	Sisa Dana Haji	Pendidikan	Usia	Berkas	Keberangkatan
1	sulaiman	Rp.4,339,000	S1	24	50	Tidak Prioritas
2	Amnah	Rp.4,339,000	S1	40	80	Tidak Prioritas
3	Fatmawaty	RP.2,956,000	S1	30	80	Tidak Prioritas
4	Siti hajar siregar	RP.4,475,700	S1	35	70	Prioritas
5	Sulaiman, ST	RP.4,475,700	S1	35	85	Prioritas
...
...
...
50	Nuraini Siregar	Rp.2,956,000	S1	35	50	Tidak Prioritas

Untuk menghitung entropy, ada beberapa proses yang harus didahului. Berikut adalah proses perhitungannya:

1. Proses Node Akar

Dimulai dari node akar, harus dihitung dulu entropy untuk node akar (semua data) terhadap komposisi kelas. Rumus menghitung entropynode akar adalah sebagai berikut:

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

Jadi Entropy (S)=

$$E(semua) = -((pi(ya|semua) \times \log_2 pi(ya|semua)) + (pi(tidak|semua) \times \log_2 pi(tidak|semua)))$$

$$\left(-\left(\frac{32}{50}\right) \times \log_2 \left(\frac{32}{50}\right)\right) + \left(-\left(\frac{18}{50}\right) \times \log_2 \left(\frac{18}{50}\right)\right) = 0.942683189$$

1. Hitung Entropy Atribut Persiapan Berkas

Persiapan Berkas= Maximum

$$Entropy (Maximum) = -(11/11) * \log_2(11/11) -(0/11) * \log_2(0/11) = 0$$

Persiapan berkas =medium

$$Entropy (Medium) = -(8/28)*\log_2(8/28)-(20/28)*\log_2(20/28) = 0.863120569$$

Persiapan berkas = minimum

$$Entropy (Minimum) = -(1/11) * \log_2 (1/11) -(10/11) * \log_2(10/11) = 0.439496987$$

Setelah Entropydiperoleh lalu menghitung nilai Gain atribut Persiapan Berkas, dengan persamaan dibawah ini. Untuk menghitung gain setiap atribut rumusnya adalah:

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

$$Gain (Persiapan Berkas) = 0.942 - \left(\left(\frac{11}{50}\right) \times 0 + \left(\frac{28}{50}\right) \times 0.863 + \left(\frac{11}{50}\right) \times 0.439\right)$$

$$Gain (Persiapan Berkas) = 0.362646334$$

2. Hitung Entropy Atribut USIA

USIA = Paruh Baya

$$Entropy (Paruh Baya) = -(14/16) * \log_2(14/16) -(2/16) * \log_2(2/16) = 0,543564443$$

USIA = TUA

$$Entropy (TUA) = -(16/19) * \log_2(16/19) -(3/19) * \log_2(3/19) = 0,629249224$$

USIA = MUDA

$$Entropy (MUDA) = -(2/15) * \log_2(2/15) -(13/15) * \log_2(13/15) = 0,566509507$$

Setelah entropy diperoleh lalu menghitung nilai gain atribut USIA, dengan persamaan dibawah ini.

$$Gain(USIA) = 0.942 - \left(\left(\frac{16}{50}\right) \times 0,543 + \left(\frac{19}{50}\right) \times 0,629 + \left(\frac{15}{50}\right) \times 0,566\right)$$

$$Gain (USIA) = 0.35967501$$

Setelah mendapat entropy dan gain dari keseluruhan kasus, lakukan analisis pada setiap atribut dan nilai-nilainya lalu hasilnya ditampilkan seperti yang ada pada Tabel 3.

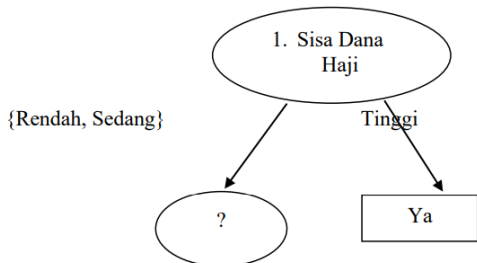
Tabel 3. Analisis Atribut, Nilai Entropy dan Gain

Node	Atribut	Nilai	Jlh Data	Ya	Tidak	Entropy	Gain
1	Persiapan Berkas	Maximum	11	11	0	0	0,362
		Medium	28	20	8	0,863	
		Minimum	11	1	10	1	
	USIA	Paruh Baya	16	14	2	0,566	0,359
		TUA	19	16	3	0,629	
		MUDA	15	2	13	0,566	
	Pendidikan	S1	47	31	16	0,925	
		D3	2	1	1	1	



Terakhir	SMK	1	0	1	0	0,032
	Tinggi	5	3	2	0,970	
	Sedang	8	6	2	0,811	
	Rendah	37	23	14	0,956	
Sisa Dana Haji						0,007

Pohon keputusan :



Gambar 2. Pohon Keputusan Node 1

Tabel 4. Data Persiapan Berkas Minimum

No	Nama	Sisa Dana Haji	Pendidikan	Usia	Berkas	Keberangkatan
1	Titikngatmi n Tarsih	Tinggi	S1	Muda	Minimum	Tidak
2	Sumarno	Sedang	S1	Muda	Minimum	Tidak
3	Amnah	Sedang	S1	Tua	Minimum	Ya
4	Fatmawaty	Rendah	S1	ParuhBaya	Minimum	Tidak
5	Sitihajar siregar	Rendah	S1	ParuhBaya	Minimum	Tidak
...
...
...
11	Siti Fatimah	Rendah	S1	Tua	Minimum	Tidak

Kemudian data di Tabel 4 dianalisis dan dihitung lagi entropy atribut Persiapan Berkas Minimum dan

entropy setiap atribut serta gainnya sehingga hasilnya seperti data pada Tabel 4.8. Setelah itu tentukan pilih atribut yang memiliki gain tertinggi untuk dibuatkan node berikutnya. Adapun cara perhitungan entropy dan gain sama seperti langkah sebelumnya.

Jadi Entropy Kasus (S) =

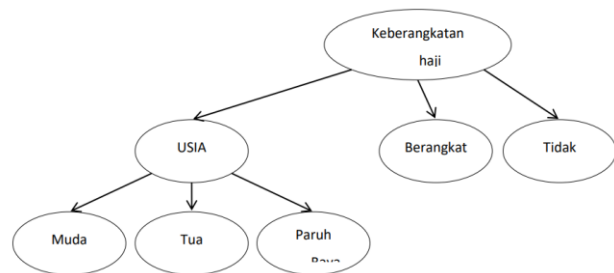
$$\left(-\left(\frac{1}{11}\right) \times \log_2\left(\frac{1}{11}\right)\right) + \left(-\left(\frac{10}{11}\right) \times \log_2\left(\frac{10}{11}\right)\right) = 0.439496$$

Dengan melakukan perhitungan pada setiap atribut didapatkan hasil sebagai berikut :

Tabel 5. Nilai Entropy dan Gain

NODE		JUMLAH KASUS	TIDAK	YA	ENTROPY	GAIN	
1.1	PERSIAPAN BERKAS MINIMUM	11	10	1	0.439496987		
	SISA DANA HAJI	RENDAH	8	8	0	0	0.257678805
		SEDANG	2	1	1	1	
		TINGGI	1	1	0	0	
PENDIDIKAN						0	
	SMK	0	0	0	0		
	D3	0	0	0	0		
USIA	S1	11	10	1	0.439496987		
	MUDA	6	6	0	0	0.189052669	
	TUA PARUH BAYA	3	2	1	0.918295834		
		2	2	0	0		

Dari Tabel 5, gain tertinggi ada pada atribut USIA, dan Nilai yang dijadikan daun atau leaf adalah Banyak dan Sedikit karena nilai entropi 0. Jika divisualisasikan maka pohon keputusan tampak seperti Gambar 3:



Gambar 3. Pohon Keputusan dengan Persiapan Berkas Minimum

IV. KESIMPULAN

Kesimpulan dari Penentuan Pola Keberangkatan Jama'ah Haji dengan Algoritma C4.5 pada Kementerian Agama Lubuk Pakamyang telah diselesaikan ini dapat diambil beberapa kesimpulan diantaranya adalah. Permasalahan yang terjadi berkenaan dengan prioritas keberangkatan Jemaah Haji di Kemenag Lubuk Pakam dapat di selesaikan dengan menggunakan Algoritma C.45., Penentuan Pola Keberangkatan Haji Di Kementerian Agama Lubuk Pakam dengan menyesuaikan data penilaian agar dapat digunakan dengan Metode C-45. , Aplikasi yang mengadopsi Metode C-45telah dapat digunakan Untuk Menentukan Antrian Jadwal



Keberangkatan Haji Di Kementerian Agama Lubuk Pakam.

V. REFERENSI

- [1] F. A. Sianturi, “Analisa Decision Tree Dalam Pengolahan Data Siswa,” *MEANS (Media Informasi Analisa dan Sistem)*, vol. 3, no. 2, pp. 166–172, 2018, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/
- [2] F. A. Sianturi, T. Informatika, and S. Utara, “Penerapan Algoritma Apriori Untuk Penentuan Tingkat,” *Mantik Penusa*, vol. 2, no. 1, pp. 50–57, 2018, [Online]. Available: <http://ejournal.pelitanusantara.ac.id/index.php/mantik/article/view/330>
- [3] M. Fahmi and F. A. Sianturi, “ANALISA ALGORITMA APRIORI PADA PEMESANAN KONSUMEN DI CAFÉ THE L . CO COFFE,” *SAINTEK (Jurnal Sains dan Teknologi)*, vol. 1, no. 1, pp. 52–57, 2019.
- [4] E. Utari and P. M. Hasugian, “Menentukan Pola Hubungan Penyebab Kecelakaan Lalu Lintas Dengan Algoritma Apriori,” vol. 3, no. 3, pp. 127–132, 2021.
- [5] K. R. Sitanggang and P. S. Hasugian, “Penerapan Data Mining Dalam Menganalisa Pola Peminjaman Buku di Perpustakaan SMP Negeri 2 Beringin Satu Atap Menggunakan Algoritma Apriori,” *JIKOMSI: Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 1, no. 1, pp. 1–5, 2018, [Online]. Available: <http://ejournal.sisfokomtek.org/index.php/jikom/article/download/16/8>
- [6] E. Febrivani and R. Winanjaya, “Penerapan Data Mining Asosiasi Pada Persediaan Obat,” vol. 3, no. 3, pp. 354–365, 2021.
- [7] E. Febrivani and R. Winanjaya, “Penerapan Data Mining Asosiasi Pada Persediaan Obat,” vol. 3, no. 3, pp. 354–365, 2021.
- [8] S. Dengan, C. Algoritma, and W. Hidayati, “Data Mining Penentuan Tenaga Perawat Di RSUD Sultan,” vol. 1, no. 2, pp. 1–7, 2018.
- [9] A. S. Sitio, “Implementasi Keamanan Data Keuangan di SMK Swasta Musda Perbaungan Menggunakan Metode RC4,” vol. 3, no. 3, pp. 60–66, 2021.
- [10] R. Situmorang and F. A. Sianturi, “Implementation of Data Mining to Predict Stocks of Goods Using the Apriori Algorithm at Mom’s Kitchen Bakery”, *J.Intell DSS, Intel DSS, jdss*, vol. 3, no. 3, pp. 22-30, Sep. 2020
- [11] M. Hartono and Fricles A Sianturi, “Penerapan Algoritma Apriori Dalam Menentukan Merek Pakaian Yang Paling Diminati Pada Darma Utama (DM Fashion)”, *JUMIN*, vol. 3, no. 1 Desember, pp. 71-78, Dec. 2021.