

PERBANDINGAN ALGORITMA C4.5 DAN NAÏVE BAYES DALAM MEMPREDIKSI KETEPATAN WAKTU LULUS MAHASISWA FASTIKOM UNSIQ

¹⁾ Reza Tri Kurniawan, ²⁾ Adi Suwondo, ³⁾ Saifu Rohman, ⁴⁾ Nur Hasanah, ⁵⁾ Sukowiyono

^{1,2,3,4,5)} Universitas Sains Al-Qur'an

¹⁾ trireza970@gmail.com, ²⁾ adiduwondo@unsiq.ac.id, ³⁾ rohman_saifu@yahoo.com, ⁴⁾ nurh.unsiq@gmail.com, ⁵⁾ suko34497@gmail.com

INFO ARTIKEL

Riwayat Artikel :

Diterima : 28 November 2023

Disetujui : 30 November 2023

Kata Kunci :

C4.5, Naïve Bayes, Perbandingan

ABSTRAK

Fakultas Teknik dan Ilmu Komputer merupakan salah satu fakultas dari Universitas Sains Al-Quran yang memiliki mahasiswa cukup banyak. Banyaknya mahasiswa baru tidak sebanding dengan mahasiswa yang lulus. Hal ini berpengaruh terhadap fakultas itu sendiri untuk proses akreditasi. Observasi dan wawancara dengan mahasiswa Fakultas Teknik dan Ilmu Komputer mengungkapkan beberapa kendala atau faktor yang membuat mahasiswa Fakultas Teknik dan Ilmu Komputer terlambat lulus. Penelitian ini menggunakan dua metode yaitu Algoritma C4.5 dan Naïve Bayes untuk mengetahui bagaimana penerapan Algoritma C4.5 dan Naïve Bayes dalam memprediksi ketepatan waktu lulus mahasiswa Fakultas Teknik dan Ilmu Komputer, dan dilakukan pengujian dengan aplikasi RapidMiner dan hasilnya menunjukkan bahwa Algoritma Naïve Bayes memiliki akurasi yang tinggi sebesar 100% dan Algoritma C4.5 sebesar 71,21%. Dapat disimpulkan bahwa algoritma Naïve bayes memiliki kinerja yang lebih baik daripada Algoritma C4.5

ARTICLE INFO

Article History :

Received : Nov 28, 2023

Accepted : Nov 30, 2023

Keywords:

C4.5, Naïve Bayes, Comparison

ABSTRACT

The Faculty of Engineering and Computer Science is one of the faculties of the University of Al-Qur'an Science which has quite a number of students. The number of new students is not proportional to the students who graduated. This affects the faculty itself for the accreditation process. Observations and interviews with students of the Faculty of Engineering and Computer Science revealed several obstacles or factors that made students of the Faculty of Engineering and Computer Science graduate late. This study uses two methods, namely the C4.5 and Naïve Bayes Algorithms to find out how the C4.5 and Naïve Bayes Algorithms are applied in predicting the timeliness of graduating students from the Faculty of Engineering and Computer Science, and testing is carried out with the RapidMiner application and the results show that the Naïve Bayes Algorithm has high accuracy of 100% and C4.5 Algorithm of 71.21%. It can be concluded that the Naïve Bayes algorithm has better performance than the C4.5 Algorithm.

1. PENDAHULUAN

Seiring perkembangan zaman, teknologi penyimpanan data dan informasi semakin maju, perubahan merupakan salah satu hal yang tidak bisa dihindari. Lahirnya inovasi akan merangsang lahirnya inovasi baru lain untuk ditingkatkan. Teknologi membantu segala aspek kehidupan manusia memberikan kemudahan dan peningkatan kualitas hidup manusia. Hal ini meningkatkan jumlah perangkat informasi yang menghasilkan data yang cukup banyak.

Fakultas Teknik dan ilmu komputer (FASTIKOM) merupakan salah satu fakultas yang berada di UNSIQ memiliki lima program studi. Mahasiswa FASTIKOM salah satu terbanyak yang ada di UNSIQ. Sayangnya jumlah lulusan yang dihasilkan setiap periode tidak sebanding dengan jumlah mahasiswa yang masuk. Banyaknya mahasiswa yang lulus tidak tepat waktu menjadi masalah tersendiri bagi pihak fakultas (Asnawi, 2023).

Mahasiswa yang terlambat lulus biasanya dipengaruhi oleh beberapa faktor. Berdasarkan wawancara awal terhadap beberapa mahasiswa Fakultas Teknik dan Ilmu Komputer Ketika diwawancari mengatakan malas, tidak mempunyai judul, lingkungan, ketidakpatuhan dalam aturan, tidak tahu mulai dari mana, tidak nyaman dalam proses belajar mengajar, tidak serius dalam belajar, dan keuangan. Hal ini sejalan dengan hasil penelitian sebelumnya dari Etika dan Hasibuan (2016), Aslinawati dan Mintarti (2017) secara umum mengatakan bahwa hampir Sebagian mahasiswa Ketika diwawancarai mengatakan bahwa kesulitan-kesulitan yang mereka hadapi adalah judulnya ditolak, kuliah sambil kerja, keputusan untuk menikah, orang tuanya sedang sakit, aktif berorganisasi, tidak tau mulai dari mana.

Penelitian dengan Algoritma C4.5 dan Naïve bayes dilakukan oleh (Muhammad kamil dan Widya Cholil, 2020) yang berjudul perbandingan algoritma C4.5 dan Naïve Bayes pada lulusan tepat waktu mahasiswa (studi kasus: UIN Raden Fatah Palembang). Dari hasil penelitian tersebut menunjukkan bahwa algoritma C4.5 memiliki nilai akurasi yang lebih bagus dibandingkan dengan algoritma Naïve bayes dengan nilai akurasi algoritma C4.5 sebesar 69,54% sedangkan Naïve bayes memiliki nilai akurasi sebesar 68,38%.

Masalah yang diteliti bagaimana penerapan perbandingan algoritma C4.5 dan Naïve bayes pada prediksi ketepatan waktu lulus mahasiswa Fakultas Teknik dan Ilmu Komputer Universitas Sains Al-Qur'an (Khafidli, 2023).

Tujuan penelitian ini adalah membandingkan dua metode pada program studi kasus ketepatan waktu lulus mahasiswa FASTIKOM UNSIQ manakah metode yang lebih akurat antara algoritma C4.5 dan Naïve bayes

2. METODE

Penelitian ini menggunakan metode observasi dan wawancara, objek penelitian yang digunakan dalam penelitian ini adalah mahasiswa Fakultas Teknik dan Ilmu Komputer Universitas Sains Al-qur'an Angkatan 2013 dengan jumlah 132 mahasiswa yang akan digunakan sebagai data. pengambilan data dilakukan di Fakultas Teknik dan Ilmu Komputer Universitas Sains Al-Qur'an

3. HASIL DAN PEMBAHASAN

3.1 Algoritma C4.5

Evaluasi dan validasi hasil algoritma C4.5

Pada langkah awal dari proses evaluasi algoritma C4.5 ini adalah mempersiapkan data yang akan di proses, pada penelitian ini data yang akan diolah adalah Mahasiswa Fakultas Teknik dan Ilmu Informatika Universitas Sains Al-Qur'an Angkatan 2013

Hasil Perhitungan Entropy Dan Gain

Dari rumus perhitungan dari entropy dan gain, maka diperoleh hasil sebagai berikut.

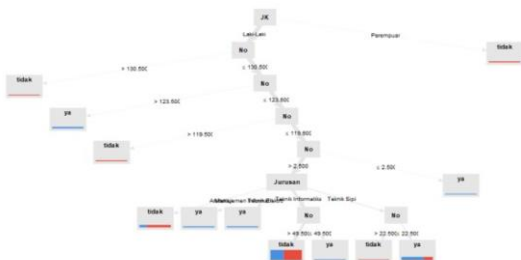
Tabel 1. Perhitungan Entropy dan Gain

	Jumlah kasus	Tidak tepat waktu	Tepat waktu	Entropy	Gain
	132	66	66	-0,1546	
Teknik sipil	25	9	16	0,942683	0,403226875
Manajemen informatika	4	0	4	0	

Arsitek ur	13	10	3	0,381472	
Teknik elektro	4	0	4	0	
Teknik informat ika	86	47	39	0,993749	
Wonoso bo	85	41	44	- 0,150687 474	- 0,081833 225
Luar wonoso bo	47	25	22	- 0,151113 173	
Laki- laki	128	62	66	- 3,001441	0
Peremp uan	4	4	0	0	
Sma	48	27	21	0,988699	0,515910 175
MA	21	13	8	-2,3223	
Smk	61	26	35	0,98424	
Paket c	2	0	2	0	

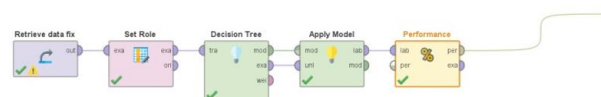
Penentuan Akar Decision Tree

Untuk menentukan akar dari pohon keputusan, maka diambil dari jumlah gain tertinggi. dari hasil perhitungan diatas, nilai gain tertinggi didapat dari atribut asal sekolah, yaitu 0,51591017. Maka atribut asal sekolah akan menjadi akar dari pohon keputusan.



Gambar 1. Pohon Keputusan

Dari gambar diatas dapat disimpulkan bahwa pengaruh terbesar dari tingkat ketepatan mahasiswa lulus adalah jenis kelamin laki-laki. Kemudian diikuti dengan jurusan mahasiswa. Validasi dan Akurasi



Gambar 2. Validasi

Dalam memvalidasi dataset, penulis menggunakan metode classification. Data dari pohon keputusan kemudian diolah dan dihitung hasilnya menggunakan variable Classification yang hasilnya seperti dibawah ini

accuracy: 71.21%

	true ya	true tidak	class precision
pred. ya	34	6	85.00%
pred. tidak	32	60	65.22%
class recall	51.52%	90.91%	

Gambar 3. Accuracy

Dari perhitungan accuracy didapatkan hasil bahwa prediksi ketepatan waktu lulus mahasiswa Fakultas Teknik dan Ilmu Komputer Universitas Sains Al-Qur'an sebesar 51.52% true tepat waktu class recall dan 90,91 untuk tidak tepat waktu untuk class recall, sedangkan yang tidak tepat waktu 65,22% pada class precision dan 85,00% untuk tepat waktu pada class precision.

3.2 Algoritma Naïve Bayes

Validasi Akurasi Naïve Bayes



Gambar 4. Validasi Naïve Bayes

Dalam memvalidasi dataset, penulis menggunakan metode classification (Asnawi, 2022). Data dari pohon keputusan kemudian diolah dan dihitung hasilnya menggunakan variable Classification yang hasilnya seperti dibawah ini.

recall: 100.00% (positive class: tidak)

	true ya	true tidak	class precision
pred. ya	66	0	100.00%
pred. tidak	0	66	100.00%
class recall	100.00%	100.00%	

Gambar 5. Accuracy Naïve Bayes

Dari perhitungan accuracy didapatkan hasil bahwa prediksi ketepatan waktu lulus mahasiswa Fakultas Teknik dan Ilmu Komputer Universitas Sains Al-Qur'an sebesar 100% untuk class recall tepat waktu dan 100% untuk class recall yang tidak tepat waktu, sedangkan yang tidak tepat waktu 100% class precision dan 100% tepat waktu pada class precision (Yunus, 2023).

3.3. Perbandingan Algoritma C4.5 dan Naïve Bayes

Tabel 2. Hasil Perbandingan

	Naïve Bayes	C45
--	-------------	-----

Accuracy	100%	71,21%
Precision	100%	51,51%
Recall	100%	85%

Berdasarkan table diatas dapat disimpulkan algoritma Naïve bayes memiliki nilai accuracy 100%, precision 100% dan recall 100%. Sedangkan algoritma C4.5 memiliki nilai accuracy 71,21%, precision 51,51% dan recall 85%. Dengan itu algoritma C4.5 bisa dianggap sebagai pilihan yang baik dalam mengklasifikasi data dengan tingkat akurasi yang lebih tinggi dan kemampuan yang lebih baik.

4. PENUTUP

4.1. Kesimpulan

Penelitian terhadap mahasiswa fakultas Teknik dan Ilmu Komputer Angkatan 2013 menggunakan algoritma C4.5 dan Naïve Bayes dengan aplikasi RapidMiner menyimpulkan bahwa algoritma C4.5 menetapkan atribut sekolah sebagai akar pohon keputusan dengan gain tertinggi (0,51591017), sementara program studi teknik informatika memiliki peluang lulus tepat waktu lebih besar daripada program studi lainnya. Algoritma C4.5 juga mencapai nilai akurasi, presisi, dan recall sebesar 71,21%, 51,51%, dan 85%. Sebaliknya, algoritma Naïve Bayes menunjukkan kinerja lebih baik daripada C4.5, dengan nilai akurasi, presisi, dan recall mencapai 100%. Kelas paket C memiliki probabilitas tertinggi dengan nilai 1,51 dalam algoritma Naïve Bayes.

4.2. Saran

Berdasarkan hasil penelitian, peneliti merekomendasikan beberapa saran. Pertama, instansi perlu membuat kebijakan di Fakultas Teknik dan Ilmu Komputer untuk memastikan mahasiswa dapat lulus tepat waktu. Kedua, bagi peneliti selanjutnya, disarankan untuk menjajaki metode lain sebagai pembanding untuk penelitian ini.

5. DAFTAR PUSTAKA

Asnawi, M. F. (2023). Kepuasan Pengguna Sistem Informasi Evaluasi Dosen FASTIKOM UNSIQ. *Journal of Economic, Management, Accounting and Technology*, 6(2), 262-269.

Aslinawati, E.N. & Mintarti S.U. 2017. Keterlambatan Penyelesaian Skripsi Mahasiswa Angkatan 2012 (Studi Kasus di Jurusan Ekonomi Pembangunan Fakultas Ekonomi Universitas Negeri Malang). *Jurnal Pendidikan Ekonomi*. Vol.10, No.1,23-33.

Etika & Hasibuan (2016). Deskripsi Masalah Mahasiswa Yang Sedang Menyelesaikan Skripsi. *Jurnal Kopasta*. 40–45.

Muhammad Kamil, Widya cholil (2020). Perbandingan Algoritma C4.5 dan Naive Bayes Pada Lulusan Tepat Waktu Mahasiswa. *JURNAL INFORMATIKA*, Vol.7, No.2,97-106

Khafidli, M. M. I., Asnawi, M. F., Suwondo, A., & Utomo, D. P. (2023). DATA MINING TO PREDICTE THE TIME OF KHATAMAN TAHFID STUDENTS TAHFID PPTQ AL-ASY'ARIYYAH USING C4. 5 ALGORITHM. *Clean Energy and Smart Technology*, 1(2), 63-69.

Asnawi, M. F., & Aji, A. P. (2022). SISTEM PENGELOLAAN PEMBIAYAAN KEGIATAN OSIS SMP NEGERI 2 BATUR MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER DALAM PENGKLASIFIKASIAN DATA LAPORAN. *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, 1(3), 26-33.

Yunus, Y., & Nurcahyo, G. W. (2023). Perbandingan algoritma c4. 5 dan naive bayes dalam prediksi kelulusan mahasiswa. *Jurnal CoSciTech (Computer Science and Information Technology)*, 4(1), 193-199.