

Optimalisasi Algoritma C4.5 untuk Prediksi Kerusakan Mesin ATM

Optimization of the C4.5 Algorithm for Predicting ATM Machine Damage

Dahri Yani Hakim Tanjung

Sistem Informasi, Fakultas Teknik dan Ilmu Komputer

Universitas Potensi Utama, Medan

e-mail: notashapire@gmail.com

Abstrak

Penelitian ini bertujuan untuk memberikan evaluasi terhadap data kerusakan pada mesin ATM menggunakan data mining. Data mining dengan algoritma C4.5 digunakan untuk memprediksi kerusakan mesin ATM yang diharapkan dapat membantu para teknisi dalam mengambil keputusan untuk melakukan perbaikan. Hasil penelitian ini diharapkan dapat membantu unit pengelola ATM dalam mengoptimalkan dan memonitoring mesin ATM, sehingga dapat memberikan pelayanan yang optimal kepada nasabah. Algoritma C4.5 merupakan algoritma yang mampu dalam membuat membentuk pohon keputusan, dimana pohon keputusan yang dihasilkan akan membentuk sebuah pengetahuan baru. Berdasarkan hasil pengujian pada pohon keputusan diperoleh kesesuaian data terhadap data kerusakan mesin ATM. Hasil penerapan metode C4.5 terhadap prediksi kerusakan mesin Hasil dari algoritma C4.5 berupa pohon keputusan yang memiliki aturan yang dijadikan sebagai pengetahuan dan informasi yang lebih mudah untuk dipahami

Kata kunci—Data Mining, C4.5, Prediksi, Kerusakan ATM

Abstract

This study aims to provide an evaluation of the damage data on ATM machines using data mining. Data mining with the C4.5 algorithm is used to predict the damage to the ATM machine which is expected to assist technicians in making decisions to make repairs. The results of this study are expected to help the ATM management unit in optimizing and monitoring ATM machines, so that they can provide optimal service to customers. The C4.5 algorithm is an algorithm that is able to help form a decision tree, where the resulting decision tree will form new knowledge. Based on the test results on the decision tree, it was obtained the suitability of the data to the ATM machine damage data. The results of applying the C4.5 method to prediction of engine failure The results of the C4.5 algorithm are in the form of a decision tree that has rules that are used as knowledge and information that is easier to understand

Keywords— Data Mining, C4.5, Prediction, Damage ATM Machine

1. PENDAHULUAN

ATM (Automated Teller Machines) merupakan salah satu fasilitas yang diberikan perbankan dalam bentuk electronic delivery channel. ATM merupakan sebuah perangkat elektronik yang menggantikan kerja teller secara otomatis yang memberikan kemudahan dalam bertransaksi kepada para nasabahnya dan dapat dijangkau kapanpun dan dimanapun [1]

Kemudahan dalam melakukan transaksi bisnis merupakan tuntutan masyarakat saat ini. Untuk memenuhi tuntutan tersebut, dunia usaha senantiasa berinovasi dan menciptakan sarana atau fasilitas usaha. Salah satunya adalah penarikan atau transfer dana melalui anjungan tunai mandiri atau automatic money transaction (ATM). ATM adalah perangkat elektronik yang memungkinkan

nasabah bank untuk mengambil uang dan memeriksa rekening tabungan mereka tanpa harus dilayani oleh teller manusia. Banyak ATM juga mengizinkan setoran atau cek, dan transfer uang [2]

Seiring dengan meningkatnya transaksi yang dilakukan menggunakan ATM, tingkat kerusakan mesin ATM juga bertambah yang dapat mengakibatkan kerugian karena nasabah tidak bisa bertransaksi dengan ATM. Hal ini menjadi pertimbangan untuk tindak lanjut penyebab terjadinya kerusakan mesin ATM agar dapat dioptimalkan sesuai dengan kebutuhannya

Pada penelitian ini metode pemecahan masalah yang digunakan adalah Data Mining. Dimana algoritma data mining yang digunakan adalah algoritma C4.5 yang diharapkan pengetahuan baru yang dihasilkan dapat membantu teknisi dalam mengambil keputusan mengenai perbaikan mesin ATM sehingga tetap dapat memberikan pelayanan yang terbaik kepada konsumen.

Penelitian dimana dalam menentukan kemungkinan pengunduran diri mahasiswa di STMIK AMIKOM Yogyakarta, didapatkan hasil kinerja algoritma C4.5 lebih cepat dan akurat dibandingkan algoritma K-Nearest. Pada metode ini menggunakan nearest neighbor, proses pengujian memakan waktu lebih lama daripada menggunakan algoritma C4.5 karena setiap kasus baru akan dicocokkan dengan semua kasus lama [3]. Algoritma C4.5 memiliki tingkat akurasi yang tinggi dalam menghasilkan suatu keputusan, akurasinya mencapai 94% pada tahap pelatihan dan 93% pada tahap uji coba [4].

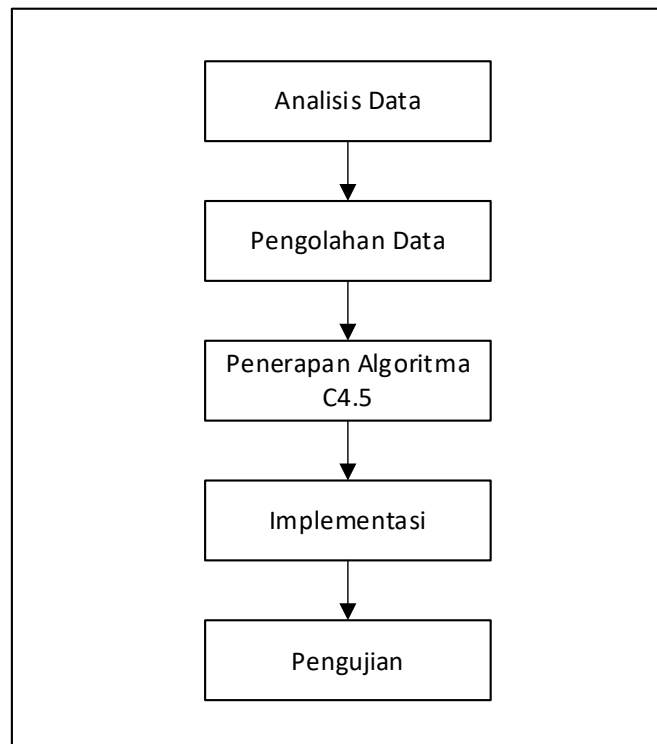
Algoritma ini menggunakan konsep data mining yang melakukan proses penggalian informasi pada keputusan sebelumnya, kemudian dibuat menjadi informasi yang membentuk pola pohon keputusan. Prosesnya diawali dengan pengolahan data latih, kemudian data tersebut dicari prioritas terbesar dan dijadikan root pada pohon. Proses ini diulangi dengan mencari keuntungan terbesar sampai perbandingan bersih terpenuhi dan pohon berhenti terbentuk [5]. Algoritma C4.5 digunakan untuk membentuk pohon keputusan dan dapat merepresentasikan dan memodelkan hasil eksplorasi data penting, sehingga pengetahuan atau informasi dari data tersebut lebih mudah dikenali[6]

Data mining dapat digunakan untuk memprediksi tingkat kelulusan menggunakan algoritma C4.5 untuk mendapatkan informasi yang berguna dan dapat membantu sekolah membuat keputusan di tahun mendatang[7]. Analisis keputusan pemberian kredit menggunakan algoritma klasifikasi C4.5 dan Naïve Bayes. Hasil yang diperoleh dari perbandingan kedua algoritma menunjukkan bahwa tingkat akurasi analisis lebih baik menggunakan algoritma klasifikasi C4.5 yaitu sebesar 88,90% sedangkan menggunakan algoritma klasifikasi Naïve Bayes sebesar 80,00% [8]. Data Mining adalah proses menemukan pola baru dari kumpulan data yang besar, salah satunya adalah metode klasifikasi yang digunakan untuk mengklasifikasikan setiap item dalam kumpulan data ke dalam kumpulan kelas[9]

Metode pemecahan masalah yang digunakan dalam penelitian ini adalah Data Mining. Data Mining ditawarkan menggunakan algoritma C4.5. Diharapkan dengan penelitian ini dapat membantu teknisi dalam mengambil keputusan mengenai perbaikan mesin ATM. Hasil penelitian ini diharapkan dapat membantu unit pengelola ATM dalam mengoptimalkan kerusakan mesin sehingga dapat memberikan pelayanan yang optimal kepada nasabah.

2. METODE PENELITIAN

Pada bab ini menjelaskan tentang metodologi penelitian yang digunakan dalam penyelesaian masalah. Metodologi penelitian dapat digambarkan melalui kerangka kerja penelitian. Adapun kerangka kerja dari penelitian ini dapat digambarkan pada gambar 1 berikut :



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian pada gambar 1 di atas maka dapat diuraikan sebagai berikut :

1. Analisis Data
Pada tahap ini adalah langkah awal untuk menentukan rumusan masalah dari penelitian. Permasalahan-permasalahan yang ada, selanjutnya dianalisa untuk mengetahui bagaimana cara penyelesaian terhadap masalah tersebut .
2. Pengolahan Data
Tahap selanjutnya adalah pengolahan data yang terdiri dari beberapa tahapan untuk menghasilkan pohon keputusan yaitu Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada.
3. Penerapan Algoritma C4.5
Setelah data dikumpulkan dilakukan analisa data untuk menyesuaikan kegiatan data yang akan diolah dengan algoritma C4.5
4. Implementasi
Pada tahap implementasi ini adalah bagaimana pengolahan datanya diterapkan dengan bantuan sebuah tools. Tools yang akan digunakan dalam implementasi penelitian ini adalah dengan menggunakan Software RapidMiner.
5. Pengujian
Pada tahap ini, dilakukan Pengujian untuk mengetahui apakah penelitian yang dilakukan telah sesuai dengan tujuan yang diharapkan. Dimana menyesuaikan hasil dari sistem dengan perhitungan manual dan menguji validasi dari algoritma dengan bantuan tools rapidminer

3. HASIL DAN PEMBAHASAN

Algoritma C4.5 merupakan algoritma yang sangat populer digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Xindong Wu dan Vipin Kumar dalam bukunya yang berjudul *The Top Ten Algorithms in Data Mining*. Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang dibuat oleh J. Rose Quinlan[10]. ID3 adalah model pohon keputusan yang menggunakan kriteria perolehan informasi untuk memilih atribut yang akan digunakan untuk memisahkan objek. Objek yang memiliki information gain tertinggi dibandingkan dengan atribut lainnya relatif terhadap himpunan y dalam suatu data, dipilih untuk melakukan solusi[11]

Langkah-langkah yang dilakukan dalam membangun sebuah pohon keputusan dengan algoritma C4.5 adalah sebagai berikut:

1. Memilih atribut sebagai akar.
2. Buat cabang untuk setiap nilai.
3. Membagi kasus dari tiap cabang
4. Ulangi langkah yang sama untuk setiap cabang hingga semua kasus memiliki kelas yang sama

Memilih atribut sebagai akar diambil berdasarkan nilai gain tertinggi dari atribut yang ada. Untuk menghitung gain digunakan persamaan yang dapat dilihat pada persamaan 1 di bawah ini:

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

Dimana :

S : Kumpulan Kasus

A : Atribut

N : Jumlah Partisi Atribut A

|S_i| : Jumlah kasus di Partisi ke-i

|S| : Jumlah Kasus di S

Sedangkan perhitungan nilai entropi dapat dilihat pada persamaan yang dapat dilihat pada persamaan 2 sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (2)$$

Dimana :

S : Kumpulan Kasus

A : Atribut

N : Jumlah Partisi S

Pi : Proporsi dari Si ke S

Hasil perhitungan entropi untuk masing-masing variabel kemudian digunakan untuk mencari gain tertinggi dan gain tertinggi digunakan sebagai akar dari pohon keputusan yang akan dibentuk, kemudian proses akan berulang sampai semua kasus pada cabang memiliki kelas yang sama. Pohon keputusan yang terbentuk kemudian digariskan dalam suatu aturan.

A. Analisis Data

Data penelitian ini berdasarkan data yang diambil dari lokasi penelitian yang dapat dilihat pada Tabel 1.

Tabel 1. Data Kerusakan ATM

ATM Status	Card Reader	Receipt paper does not come out	Dispenser Failure	Damage Rate
Online	Function	Exit	Yes	High
Online	does not work	Exit	Yes	Low
Online	Function	Exit	No	High
Offline	Function	Not Exit	Yes	Low
Offline	Function	Not Exit	Yes	Low
Online	does not work	Exit	No	Low
Online	does not work	Exit	No	High
Offline	Function	Not Exit	Yes	Low
Offline	Function	Exit	Yes	Medium
Offline	does not work	Exit	Yes	Low
Online	Function	Exit	Yes	Medium
Online	does not work	Exit	No	High
Offline	Function	Not Exit	No	Low
Offline	Function	Not Exit	No	Low
Online	Function	Exit	No	Medium
Online	Function	Not Exit	Yes	Medium

B. Perhitungan Algoritma C4.5

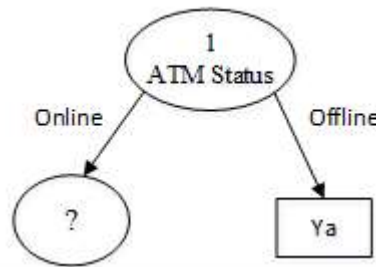
Data yang telah ditransformasi kemudian dianalisis untuk menghasilkan pohon keputusan dengan menggunakan algoritma C4.5. Perhitungan Entropy dan Gain node 1 yang dapat dilihat pada tabel 2 sebagai berikut:

Tabel 2 Hasil Perhitungan Entropi dan Gain

Node	Atribut	Score	Cases	S1	S2	S3	Entropy	Gain
1	Total		16	8	4	4	1,5	
	ATM Status							0,38
		Online	9	2	3	4	1,53	
		Offline	7	6	1	0	0,592	
	Card Reader							0,024
		function	10	5	3	2	1,485	
		Doesnt Function	6	3	1	2	1,459	
	Receipt paper does not come out							0,274
		Exit	10	3	3	4	1,571	
		No Exit	6	5	1	0	0,65	

Node	Atribut	Score	Cases	S1	S2	S3	Entropy	Gain
	Dispenser Failure							0,106
		Yes	9	3	5	1	1,352	
		No	7	1	3	3	1,449	

Berdasarkan perhitungan entropi dan gain untuk mencari akar nilai gain tertinggi adalah variabel Status ATM dengan 2 nilai variabel yaitu Online dan Offline. Selanjutnya adalah mendeskripsikan hasil perhitungan entropi dan gain dalam bentuk pohon keputusan. Pohon keputusan untuk mencari nilai akar dapat dilihat pada Gambar 1.



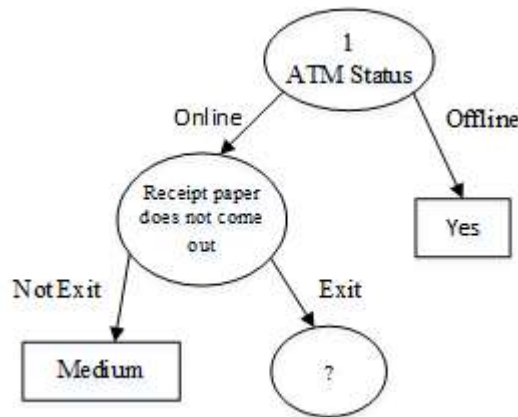
Gambar 3. Hasil perhitungan pohon keputusan

Langkah selanjutnya adalah menghitung cabang pada simpul 1.1 yang terdiri dari dua simpul.

Tabel 3 Hasil perhitungan Entropi dan Gain node 1.1

Node	Atribut	Score	Cases	S1	S2	S3	Entropy	Gain
1.1	Total		9	3	2	4	1,53	
	ATM Status - Online							
	Card Reader							1,41
		Function	4	0	2	2	1	
		Does not Function	5	2	1	2	1,52	
	Receipt paper does not come out							1,16
		Exit	8	2	2	4	1,5	
		Not Exit	1	0	1	0	0	
	Dispenser Failure							0,60
		Yes	4	1	2	1	1,5	
		Not Exit	5	1	1	3	1,37	

Dari Tabel 3 diperoleh nilai gain tertinggi dengan nilai 1,16, nilai gain tertinggi adalah cabang untuk atribut makalah. Resi tidak keluar. nilai memiliki dua cabang yaitu Exit dan Not Exit, kedua cabang sudah memiliki keputusan untuk nilai entropi 0, seperti yang ditunjukkan di bawah ini pada gambar 2



Gambar 4. Perhitungan simpul pohon keputusan pada 1.1

Sedang cabang sudah bercabang keputusan berarti proses berhenti. Langkah selanjutnya adalah membentuk simpul cabang 2,1 dari Keluar

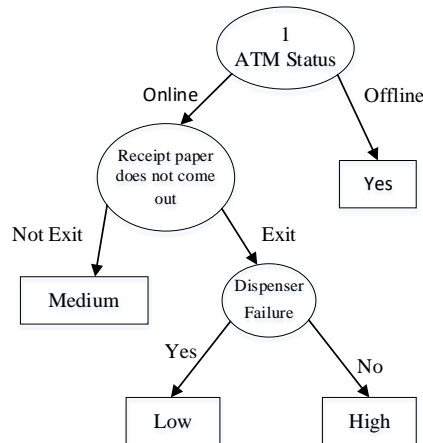
Tabel 4 Hasil perhitungan Entropi dan Gain node 2.1

Node	Atribut	Score	Cases	S1	S2	S3	Entropy	Gain
2.1	ATM Status - Online Anda Card Reader - Doesn't Function		5	2	1	2	1,52	
	Receipt paper does not come out							0
		Exit	5	2	1	2	1,52	
		Not Exit	0	0	0	0	0	
	Dispenser Failure							3,69
		Yes	1	1	0	0	0	
		No	4	1	1	2	0,50	

Dari tabel IV nilai gain tertinggi terdapat pada kegagalan dispenser dengan nilai 3,69. Yang terdiri dari dua cabang, yaitu ya dan tidak. Ya, untuk cabang ya dengan nilai entropi 0 dan cabang tidak masih memiliki nilai sehingga perhitungan selanjutnya masih dilakukan.

C. Analisis Hasil

Dari hasil perhitungan yang telah dilakukan dapat disimpulkan bahwa proses prediksi kerusakan mesin ATM menggunakan data mining dengan algoritma C4.5 dapat menghasilkan pohon keputusan yang digunakan untuk memprediksi kerusakan mesin ATM berdasarkan penyebab atau gejala kerusakan. Untuk hasil pohon keputusan dapat dilihat pada Gambar 3.

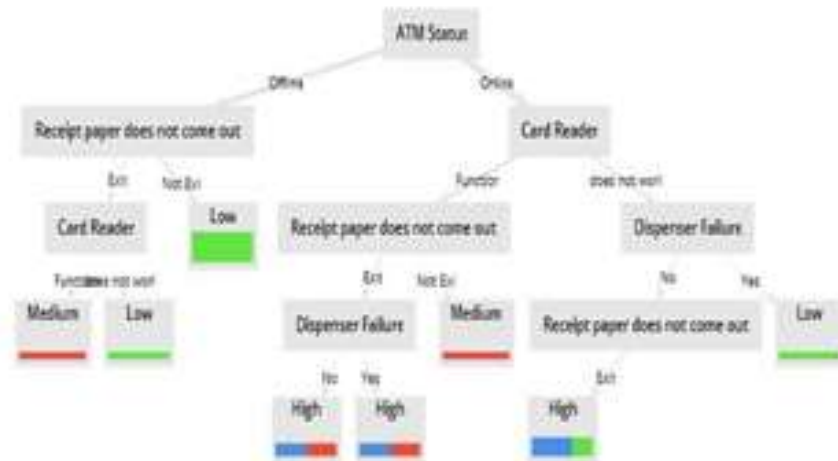


Gambar 5. Perhitungan simpul pohon keputusan pada 2.1

Penjelasan dari pohon keputusan di atas adalah sebagai berikut:

1. Jika Status ATM = Offline Maka Damage (tinggi) = yes
2. Jika Status ATM = Online dan Kertas Resi tidak keluar (Tidak Keluar) maka kerusakan = sedang
3. Jika Status ATM = Online Dan Kertas Resi tidak keluar = keluar dan dispenser gagal = ya maka kerusakan = rendah
4. Jika Status ATM = Online Dan Kertas Resi tidak keluar = keluar dan dispenser gagal = Tidak maka kerusakan = Tinggi

Kemudian data tersebut akan diuji menggunakan software Rapidminer. Dimana rapidminer merupakan software data mining yang dapat digunakan untuk mengakses beberapa metode data mining salah satunya adalah untuk mengambil keputusan. Dimana hasil pengujian yang dilakukan dengan alat rapidminer untuk mengetahui kerusakan mesin ATM dapat dilihat pada gambar. 4



Gambar 6. Hasil pengujian dengan Rapidminer

Dimana rule yang dihasilkan dari pohon keputusan diatas menggunakan rapid miner adalah sebagai berikut :

Status ATM = Offline

| Kertas struk tidak keluar = Keluar

|| Pembaca Kartu = Fungsi: Sedang {Tinggi=0, Rendah=0, Sedang=1}

|| Pembaca Kartu = tidak berfungsi: Rendah {Tinggi=0, Rendah=1, Sedang=0}

| Kertas struk tidak keluar = Tidak Keluar : Rendah {Tinggi=0, Rendah=5, Sedang=0} Status ATM = Online
 | Pembaca Kartu = Fungsi
 || Kertas struk tidak keluar = Keluar
 ||| Kegagalan Dispenser = Tidak: Tinggi {Tinggi=1, Rendah=0, Sedang=1}
 ||| Kegagalan Dispenser = Ya: Tinggi {Tinggi=1, Rendah=0, Sedang=1}
 || Kertas struk tidak keluar = Tidak Keluar : Sedang {Tinggi=0, Rendah=0, Sedang=1}
 | Pembaca Kartu = tidak berfungsi || Kegagalan Dispenser = Tidak
 ||| Kertas struk tidak keluar = Keluar : Tinggi {Tinggi=2, Rendah=1, Sedang=0}
 || Kegagalan Dispenser = Ya: Rendah {Tinggi=0, Rendah=1, Sedang=0}

Dari hasil pengujian yang dilakukan dapat disimpulkan bahwa dengan mendahului kerusakan mesin ATM berdasarkan tingkat kerusakan dapat mengoptimalkan kerja dengan meminimalkan kesalahan. Dan dapat dijadikan sebagai keputusan ketika terjadi kerusakan, hal-hal apa saja yang harus dilakukan dan dipersiapkan.

4. KESIMPULAN

Kesimpulan yang dapat diberikan dari pembahasan ini adalah sebagai berikut :

1. Dengan menggunakan algoritma C4.5 dapat memprediksi kerusakan mesin ATM.
2. Model pohon keputusan yang dihasilkan adalah menjadikan variabel Status ATM sebagai root, Maka Resi tidak keluar sebagai cabang di Level 1
3. Algoritma C4.5 dapat diimplementasikan untuk mengoptimalkan kerusakan mesin ATM dengan mengolah data untuk dijadikan data latih dalam pembentukan pohon keputusan. Dari pohon keputusan ini Anda akan mendapatkan aturan yang dapat digunakan untuk memprediksi kerusakan mesin.

5. SARAN

Adapun saran yang dapat diberikan untuk pengembangan penelitian ini di masa yang akan datang adalah sebagai berikut :

1. Penulis mengharapkan dengan adanya sistem ini dapat mempermudah pekerjaan dalam mengoptimalkan kerusakan mesin.
2. Diharapkan hasil penelitian ini dapat dijadikan sebagai acuan perbandingan dengan metode yang berbeda.
3. Masih perlu pemahaman yang lebih detail untuk mengembangkan data mining sehingga dapat mempermudah perusahaan dalam melakukan perbaikan terhadap kerusakan mesin.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Potensi Utama yang telah banyak memberikan dukungan dalam penelitian ini.

DAFTAR PUSTAKA

- [1] Damaris, R. N., Sitanggang, M., & Simanjuntak, R. P. (2014). SISTEM PENGENDALIAN INTERN ATAS TRANSAKSI PENERIMAAN DAN PENGELUARAN KAS ANJUNGAN TUNAI MANDIRI (ATM). *Buletin Ekonomi*, 18(2), 54-63.
- [2] Simanjuntak, A. (2007). Tinjauan Yuridis Para Pihak dalam Transaksi Pengambilan atau Transfer Dana Melalui Mesin Anjungan Tunai Mandiri (ATM). *Jurnal Manajemen dan Kewirausahaan*, 9(2), 128-134.
- [3] Kusri, S. H., Wardoyo, R., & Harjoko, A. (2009). Perbandingan metode nearest neighbor dan algoritma c4. 5 untuk menganalisis kemungkinan pengunduran diri calon mahasiswa di stmik amikom yogyakarta. *Jurnal Dasi*, 10(1), 114-132
- [4] Rajesh, K. and S. A. (2012). Analysis of SEER Dataset for Breast Cancer Diagnosis using C4.5 Classification Algorithm. *International Journal of Advanced Research in Computer and Communication Engineering*, 1(2), 72– 77. www.ijarce.com
- [5] Julianto, W., Yunitarini, R., & Sophan, M. K. (2014). Algoritma C4. 5 Untuk Penilaian Kinerja Karyawan. *Scan*, vol, 9, 33-39
- [6] Buulolo, E., Silalahi, N., & Rahim, R. (2017). C4. 5 Algorithm To Predict the Impact of the Earthquake. *Entropy (S)*, 1, 1
- [7] Wiza, F., & Febriadi, B. (2019). Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students. *IJISTECH (International Journal Of Information System & Technology)*, 2(2), 43. <https://doi.org/10.30645/ijistech.v2i2.21>
- [8] Masripah, S. (2016). Komparasi Algoritma Klasifikasi Data Mining untuk Evaluasi Pemberian Kredit. *Bina Insani ICT Journal*, 3(1), 187-193
- [9] Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 2(2), 213–219.
- [10] Kacung, S., & Santoso, B. (2018). Sistem Deteksi Dini Untuk Meningkatkan Performance Kelulusan Mahasiswa Dengan Id3. 212–217. <https://doi.org/10.31227/osf.io/h7xrn>