

PENERAPAN ALGORITMA C 45 UNTUK KLASIFIKASI KEPATUHAN KONSUMEN

Basuki Rakhim Setya Permana

Jurusan Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Bina Bangsa
basuki.rakhim@binabangsa.ac.id

ABSTRAK

Mengingat pentingnya analisis resiko dalam dunia perkreditan, para ahli telah melakukan penelitian dengan berbagai metode yang berbeda untuk menganalisa tingkat keberhasilan pemberian kredit. Pada penelitian ini tahapan yang dilakukan dalam melakukan eksperimen ini, penulis menggunakan *model Cross-Standard Industry for Data Mining*. Model yang telah dikembangkan akan diuji keakuratannya dengan memasukkan sejumlah data uji (*test set*) ke dalam model. Untuk mengukur keakuratan model dengan baik, data uji seharusnya bukan data yang berasal dari data training. Data uji diambil dari Data Nasabah Tahun 2008 s.d 2012 untuk Paket Keahlian Teknik Komputer dan Jaringan Jalur Regular sebanyak 163 data. Dari hasil evaluasi baik secara *confusion matrix* maupun *empiris* ternyata terbukti akurat dalam penerapan rule pada ketepatan waktu pembayaran. Dengan hasil ini, maka kemungkinan terjadinya kesalahan dan kekeliruan dalam penentuan hasil ketepatan waktu pembayaran atas nasabah yang diterima ternyata tidak layak dapat dikurangi dan menekan jumlah jumlah data yang tidak akurat dalam penentuan hasil ketepatan waktu pembayaran, dengan demikian algoritma C4.5 dapat memberikan solusi untuk penentuan hasil ketepatan waktu pembayaran dan sebagai permasalahan penentuan nasabah PT. Wahana Ottomitra Multiartha.

Kata kunci: Algoritma C45, Kepatuhan Konsumen.

ABSTRACT

Given the importance of risk analysis in the world of credit, experts have conducted research with a variety of different methods to analyze the success rate of credit. In this study the stages carried out in conducting this experiment, the authors use the Cross-Standard Industry for Data Mining model. The model that has been developed will be tested for accuracy by entering a number of test data (test set) into the model. To measure the accuracy of the model properly, test data should not be data that comes from training data. Test data was taken from Customer Data for 2008 to 2012 for Computer Engineering and Network Routes Expertise Package of 163 data. From the evaluation results both in confusion matrix and empirical, it turns out to be accurate in applying the rule on the timely

payment. With this result, the possibility of errors and errors in determining the timeliness of payment results for customers received turned out to be inappropriate and can reduce the amount of inaccurate data in determining the timeliness of payment results, thus the C4.5 algorithm can provide solutions for determining the results of the timeliness of payment and as a matter of determining the customers of PT. Wahana Ottomitra Multiartha.

Keywords: Algorithm C45, Consumer Compliance

PENDAHULUAN

Latar Belakang

Perusahaan pembiayaan sangat bergantung kepada hasil penjualan yang dilakukan department marketing yang bekerjasama dengan dealer atau showroom motor mitranya, dan hasil dari tagihan dari pembiayaan yang diberikan kepada nasabahnya melalui pembiayaan kredit yang menjadi bagian department collection. Kedua department ini saling terkait satu sama lain, terutama dalam menjaga perputaran uang atau cashflow yang terjadi di cabang. Semakin banyak department marketing menarik atau mendapatkan nasabah baru, maka semakin banyak tagihan yang akan di dapat oleh department collection, sehingga department collection sangat dibutuhkan dalam hal menjaga keseimbangan perputaran uang atau cashflow cabang tersebut.

Sunaryanto dalam penelitiannya pada tahun 2009, menyimpulkan untuk memperkecil resiko kemacetan kredit, perlu dilakukan tindakan – tindakan lanjutan kepada debitur nasabah sebagai sarana untuk menurunkan tingkat kemacetan kredit dengan cara reminding call, rescheduling maupun reconditioning. Veronika Moertini dalam jurnal yang ditulis pada tahun 2003, melakukan perbandingan antara algoritma data mining C4.5 dengan algoritma ID3 untuk meklakukan klasifikasi nasabah [1].

Dalam penelitian ini, penulis menggunakan referensi-referensi yang ada untuk menjelaskan data mining dan model algoritma C4.5 . Data mining disebut sebagai proses ekstraksi pengetahuan dari data yang besar. Sesuai fungsinya, data mining adalah proses pengambilan pengetahuan dari volume data yang besar yang disimpan dalam basis data, data warehouse, atau informasi yang disimpan dalam repositori [2]. Gartner Group dalam [3] menyebutkan bahwa data mining adalah proses menelusuri pengetahuan baru, pola dan tren yang dipilih dari jumlah data yang besar yang disimpan dalam repositori atau tempat penyimpanan dengan menggunakan teknik pengenalan pola serta statistik dan teknik matematika.

Algoritma C4.5 didesain oleh J. Ross Quinlan, dinamakan C4.5 karena merupakan keturunan dari pendekatan ID3 untuk membangun pohon keputusan. C4.5 merupakan algoritma yang cocok digunakan untuk masalah klasifikasi pada machine learning dan data mining [4]. C4.5 memetakan atribut dari kelas sehingga dapat digunakan untuk menemukan prediksi terhadap data yang belum muncul. C4.5 merupakan algoritma

yang cocok digunakan untuk masalah klasifikasi pada machine learning dan data mining [5]. C4.5 memetakan atribut dari kelas sehingga dapat digunakan untuk menemukan prediksi terhadap data yang belum muncul. Pohon keputusan sendiri merupakan pendekatan “divide and conquer” dalam mempelajari masalah dari sekumpulan data independen yang digambarkan dalam bagan pohon [4]. Pohon keputusan juga merupakan sekumpulan pertanyaan yang tersusun secara sistematis, dimana setiap pertanyaan yang ada menentukan percabangan berdasarkan nilai atribut dan berhenti pada daun dari pohon yang merupakan prediksi dari kelas variable.

METODE PENELITIAN

Ada beberapa tahap yang dilakukan dalam melakukan eksperimen ini, penulis menggunakan *model Cross-Standard Industry for Data Mining* (CRISPDM) [6] yang terdiri dari 6 tahap, yaitu:

a. Tahap Business Understanding.

Berdasarkan proses data di hasilkan 852 nasabah dari gabungan data, sehingga ini menjadi permasalahan dan merupakan imbas dari analisa analis divisi collection yang kurang akurat dalam mengukur kondisi ketepatan nasabah melakukan pembayaran dalam hal ini nilai. Agar akurat maka dikembangkan model klasifikasi algoritma C4.5 untuk dijadikan rekomendasi sebagai prediksi ketepatan pembayaran nasabah dengan tujuan analisa yang dilakukan lebih akurat.

b. Tahap data understanding.

Data PPDB diambil dari PT. Wahana Ottomitra Multiartha pada tahun 2014 hingga 2018, dimana dari lima tahun terakhir 852 data nasabah . Ada beberapa atribut yang digunakan didalam data debitur yaitu no order (no order, Nama, Tenor, Tanggal bayar dan Tanggal Jatuh Tempo) dari semua atribut yang ada di tabel, merupakan nilai kategorikal dan bukan nilai angka, misalnya seperti atribut Jumlah Tanggungan yang mempunyai nilai sedikit merupakan kategori. Tabel 1 di bawah ini ditampilkan nama atribut, kategori, dan nilai angka (rangennya). Berikut rule nilainya :

Tabel 1 Kategori Atribut

Atribut	Nilai Angka	Kategori
Waktu efektif	A	
	B	
	C	
	D	
	E	
Tanggal Bayar	1-7	Sangat patuh
	8-15	Patuh

16-21	Kurang Patuh
22-31	Tidak Patuh
Tanggal Jatuh Tempo	15

c. Tahap *data preparation*.

Tahapan ini merupakan data murni yang belum diolah menjadi data training. Untuk itu maka diperlukan teknik dalam *preprocessing* [1].

d. Tahap Data Preprocessing

i. Data Cleaning

Bekerja untuk membersihkan nilai yang kosong ,tidak konsisten atau mungkin tupel yang kosong (*missing values* dan *noisy*).

ii. Data Integration

Berfungsi menyatukan tempat penyimpanan (arsip) yang berbeda kedalam satu data.Dalam hal ini,ada dua arsip yang diambil sebagai *data warehouse* yaitu data nasabah dan data angsuran.

iii. Data Reduction

Jumlah atribut dan tupelo yang digunakan untuk data training mungkin terlalu besar,hanya beberapa atribut yang diperlukan sehingga atribut yang tidak diperlukan akan dihapus. Tupel dalam data set mungkin terjadi duplikasi atau terdapat tupel yang sama,sehingga untuk memper kecil jumlah tupel, tupel yang sama dijadikan akan dijadikan dalam satu tupel untuk mewakili tupel tersebut tetapi pada penelitian ini Penulis mencoba menyajikan data apa adanya sesuai dengan yang diperoleh dari sumber penelitian.

e. Tahap Modeling

Tahap ini juga dapat disebut tahap *learning* karena pada tahap ini data training diklasifikasikan oleh model dan kemudian menghasilkan sejumlah aturan.

Model yang digunakan dalam tahap ini menggunakan algoritma C4.5. Seperti yang telah dijelaskan sebelumnya,ada beberapa tahap yang harus dilalui dalam membentuk pohon keputusan, tentunya algoritma C4.5 digunakan untuk membuat pohon keputusan, dibutuhkan data training yang diambil dari nasabah PT. Wahana Ottomitra Multiartha.

HASIL DAN PEMBAHASAN

a. Klasifikasi Hasil ketepatan waktu pembayaran dengan Algoritma C4.5

Proses klasifikasi hasil ketepatan waktu pembayaran dilakukan pada data Nasabah pada tahapan proses ketepatan waktu pembayaran. Proses pengumpulan data Nasabah oleh *Departement Collection* dimulai pada setelah berkas persyaratan tagihan diserahkan oleh nasabah kepada Persusahaan, berkas pendaftaran tersebut berisi formulir yang telah terisi, kontrak dan tenor pembayaran.

Secara gamblang proses klasifikasi sudah dilakukan penulis menggunakan model *Cross-Standard Industry for Data Mining* (CRISPDM) (Larose, 2006) terdiri dari 6 tahap, yaitu :

- a. Tahap *business understanding*
- b. Tahap *data understanding*
- c. Tahap *data preparation*.
- d. Tahap *data preprocessing*, (proses *data cleaning*, *data integration*, *data reduction*), menghitung hasil PPDB yang diterima dan tidak diterima
- e. Tahap *evaluation*
- f. Tahap *deployment*.

Pada bagian ini akan disampaikan proses evaluasi terhadap data hasil klasifikasi dengan menggunakan Algoritma C4.5 dengan rule yang terbentuk sebagai berikut :

Tabel 2. Rule Model Decision PPDB SMK

No	Nama Rule	Rule
1	R1	<i>If Test=A then Hasil=Sangat Patuh</i>
2	R2	<i>If Test=B then Hasil=Patuh</i>
3	R3	<i>If Test=C then Hasil=Kurang Patuh</i>
4	R4	<i>If Test=C then Hasil=Tidak Patuh</i>

Pengujian Model

Nilai *accuracy*, *precision*, dan *recall* dari data training dapat dihitung dengan menggunakan Rapid Miner. Setelah diuji coba dengan metode crossvalidation, didapatkan hasil pengukuran terhadap data training yaitu *accuracy*= 98.83%, *precision*= 97.61% dan *recall*= 98.45%. Model yang telah dikembangkan akan diuji keakuratannya dengan memasukkan sejumlah data uji (*test set*) ke dalam model. Untuk mengukur keakuratan model dengan baik, data uji seharusnya bukan data yang berasal dari data training (Han & Kamber, 2006). Data uji diambil dari Data Nasabah Tahun 2014 s.d 2018 untuk Paket Keahlian Teknik Komputer dan Jaringan Jalur Regular sebanyak 163 data.

Sampel akan diujikan ke dalam data training untuk mendapatkan hasil klasifikasi dari C4.5. Pada pengujian ini, ditambahkan metode penyeleksian atribut lain yaitu *gain informatin* dengan tujuan untuk melihat akurasi dari metode penyeleksian atribut dengan algoritma yang sama yaitu C4.5. Dari pengujian yang dilakukan didapatkan hasil :

Tabel 3 : Hasil Pengujian Dengan Algoritma C4.5

Prestasi Pembayaran	Tanggal Bayar	Tanggal Jatuh Tempo	Kategori
patuh	b	15	lancar
sangat patuh	a	15	lancar
kurang patuh	c	15	macet
kurang patuh	c	15	macet
patuh	b	15	lancar

tidak patuh	d	15	macet
tidak patuh	d	15	macet
kurang patuh	d	15	macet
sangat patuh	a	15	lancar
kurang patuh	c	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
tidak patuh	d	15	macet
patuh	a	15	lancar
tidak patuh	d	15	macet
patuh	b	15	lancar
patuh	b	15	lancar
patuh	b	15	lancar
sangat patuh	a	15	lancar
patuh	c	15	macet
kurang patuh	c	15	macet
patuh	b	15	lancar
sangat patuh	a	15	lancar
tidak patuh	d	15	macet
tidak patuh	d	15	macet
patuh	b	15	lancar
tidak patuh	d	15	macet
patuh	b	15	lancar
kurang patuh	c	15	macet
tidak patuh	d	15	macet
kurang patuh	c	15	macet
kurang patuh	c	15	macet
tidak patuh	d	15	macet
kurang patuh	c	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
patuh	a	15	lancar
patuh	a	15	lancar
tidak patuh	d	15	macet
kurang patuh	c	15	macet
tidak patuh	d	15	macet
patuh	b	15	lancar
kurang patuh	d	15	macet

tidak patuh	d	15	macet
tidak patuh	d	15	macet
tidak patuh	d	15	macet
tidak patuh	d	15	macet
patuh	b	15	lancar
patuh	a	15	lancar
sangat patuh	a	15	lancar
kurang patuh	c	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
tidak patuh	d	15	macet
sangat patuh	a	15	lancar
patuh	a	15	lancar
kurang patuh	c	15	macet
kurang patuh	c	15	macet
kurang patuh	d	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
tidak patuh	d	15	macet
tidak patuh	d	15	macet
tidak patuh	d	15	macet
patuh	b	15	lancar
kurang patuh	c	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
tidak patuh	d	15	macet
tidak patuh	d	15	macet
tidak patuh	d	15	macet
patuh	b	15	lancar
patuh	b	15	lancar
patuh	b	15	lancar
tidak patuh	d	15	macet
tidak patuh	d	15	macet
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
sangat patuh	a	15	lancar
patuh	b	15	lancar
kurang patuh	d	15	macet

sangat patuh	a	15	lancar
kurang patuh	d	15	macet
patuh	a	15	lancar
kurang patuh	d	15	macet

Dari data pada tabel 2 dapat disajikan bahwa data Nasabah tahun 2014 s.d 2018 yang berjumlah 163 data sebagai berikut :

Tree

Kategori = LANCAR

/ *Tanggal bayar = A: SANGAT PATUH {PATUH=54, SANGAT PATUH=272, KURANG PATUH=0, TIDAK PATUH=0}*

/ *Tanggal bayar = B: PATUH {PATUH=155, SANGAT PATUH=0, KURANG PATUH=1, TIDAK PATUH=0}*

Kategori = MACET

/ *Tanggal bayar = C: KURANG PATUH {PATUH=1, SANGAT PATUH=0, KURANG PATUH=152, TIDAK PATUH=0}*

/ *Tanggal bayar = D: TIDAK PATUH {PATUH=0, SANGAT PATUH=0, KURANG PATUH=27, TIDAK PATUH=190}*

Dari hasil pengujian di atas akan dievaluasi tingkat akurasinya menggunakan 2 model yaitu menggunakan *confusion matrix* dan *ROC (Receiver Operating Characteristic)*. Evaluasi terhadap model pohon keputusan yang telah terbentuk, akan dilakukan pengukuran akurasi dan *area under curve (AUC)*. Akurasi diukur menggunakan *confusion matrix*, dan nilai *AUC* diukur dengan menggunakan *ROC Curve*. Proses evaluasi dilakukan dengan *cross validation*. Sehingga model yang terbentuk dapat langsung diuji dengan data yang secara acak dipisahkan dengan dengan *10 folds cross validation*.

Untuk melihat kualitas model yang dihasilkan, *ROC curve* akan dibuat dan nilai *AUC* dapat dijadikan ukuran untuk melihat model yang terbentuk. Selain itu *ROC Curve* akan digunakan untuk menemukan nilai *AUC* dimana nilai *AUC* dapat diagnosa sebagai berikut:

Nilai *AUC* memiliki tingkat diagnosa sebagai berikut:

Akurasi bernilai $0.9 - 1 = \text{Excellent classification}$

Akurasi bernilai $0.8 - 0.9 = \text{Good classification}$

Akurasi bernilai $0.7 - 0.8 = \text{Fair classification}$

Akurasi bernilai $0.6 - 0.7 = \text{Poor classification}$

Akurasi bernilai $0.5 - 0.6 = \text{Failure}$

HASIL DAN PEMBAHASAN

Dengan memiliki model rule untuk penentuan ketepatan waktu pembayaran di PT. Wahana Ottomitra Multiartha, maka PT. Wahana Ottomitra Multiartha dapat lebih mudah mengidentifikasi nasabah mana yang layak untuk diterima mana yang layak untuk diterima lagi dalam kontrak selanjutnya. Dari data nasabah, PT. Wahana Ottomitra

Multiartha dapat langsung menerapkan kriteria attribute ke dalam datanya, sehingga dapat diterapkan model pohon keputusan yang terbentuk. Dengan demikian Nasabah PT. Wahana Ottomitra Multiartha dapat menentukan Nasabah yang kemungkinan besar dapat diterima serta dapat mengevaluasi proses kontrak selanjutnya. Supaya bisa diterapkan rule yang terbentuk, maka model atau rule yang terbentuk harus baik dan akurat, dalam menilai tingkat akurasi sebuah model atau rule, model akan diterapkan pada data uji sehingga dapat diketahui tingkat keberhasilan prediksi model. Dalam penelitian ini, hasil pengujian model akan dibahas dengan menggunakan *confusion Matrix* dan *ROC Curve* untuk menunjukkan seberapa baik model yang terbentuk. Tanpa menggunakan boosting, algoritma C4.5 sendiri dengan menggunakan data sudah mampu menghasilkan model dengan tingkat akurasi yang cukup tinggi yaitu sebesar 94.48%.

Pengujian Model dengan 10 Folds Cross Validation Menggunakan Confusion Matrix Proses pengujian model dengan menggunakan *Cross Validation* adalah dengan membagi data yang akan diolah menjadi 2 bagian, yaitu data testing dari data nasabah tahun 2008 hingga tahun 2012 dengan menggunakan *10 folds Cross Validation*. Dataset akan dibagi menjadi 10 bagian sama rata, dimana setiap bagian akan memiliki komposisi yang sebanding dengan bagian lainnya. Proses pengujian akan dimulai dengan pembentukan model dengan data pada bagian pertama. Model yang terbentuk akan diujikan pada 9 bagian data sisanya. Setelah itu proses akurasi akan diperhitungkan dengan melihat seberapa banyak data yang terkласifikasi secara benar.

Evaluasi dengan model *confusion matrix* menggunakan tabel seperti matrix di bawah ini :

Tabel 4 :

Tabel Model Confusion Matrix nilai true positives, false negatives, false positives, dan true negatives didapat dari data uji (Han & Kamber, 2006)

		<i>Predicted class</i>	
		C1	C2
<i>Actual Class</i>	C1	<i>True Positive</i>	<i>False Negative</i>
	C2	<i>False Negative</i>	<i>True positive</i>

True positives merupakan tupel positif di data set yang diklasifikasikan positif. *True negatives* merupakan tupel negatif di data set yang diklasifikasikan negatif. *False positives* adalah tupel positif di data set yang diklasifikasikan negatif sedangkan *false negatives* merupakan jumlah tupel negatif yang diklasifikasikan positif. Dari hasil pengujian menggunakan tool aplikasi Rapidminer dan dapat kesamaan hasil dari perhitungan di atas bisa dilihat sebagaimana gambar 1 berikut:

Gambar 1.

Pengujian Model Dengan *Confusion Matrix* menggunakan Rapidminer

		true PATUH	true SANGAT PATUH	true KURANG PATUH	true TIDAK PATUH	class precision
		pred. PATUH	0	1	0	99.36%
		pred. SANGAT PATUH	272	0	0	83.44%
		pred. KURANG PATUH	1	0	152	99.35%
		pred. TIDAK PATUH	0	0	27	87.55%
		class recall	73.81%	100.00%	84.44%	100.00%

Dari hasil *performance confusion matrix* data testing set sebagaimana gambar 1 dari pendaftar sebanyak 163.

KESIMPULAN DAN SARAN

Dari hasil evaluasi baik secara *confusion matrix* maupun *empiris* ternyata terbukti akurat dalam penerapan rule pada ketepatan waktu pembayaran. Dengan hasil ini, maka kemungkinan terjadinya kesalahan dan kekeliruan dalam penentuan hasil ketepatan waktu pembayaran atas nasabah yang diterima ternyata tidak layak dapat dikurangi dan menekan jumlah jumlah data yang tidak akurat dalam penentuan hasil ketepatan waktu pembayaran, dengan demikian algoritma C4.5 dapat memberikan solusi untuk penentuan hasil ketepatan waktu pembayaran dan sebagai permasalahan penentuan nasabah PT. Wahana Ottomitra Multiartha.

Aspek Manajerial

Dalam aspek manajerial, agar PT. Wahana Ottomitra Multiartha sebagai pelaku dalam penerapan algoritma klasifikasi C4.5 setelah penelitian ini tentunya PT. Wahana Ottomitra Multiartha perlu untuk mempersiapkan sumber daya manusia dan faktor pendukung lain seperti hardware dan software dan adanya kemauan untuk menerapkan hasil penelitian ini. Model Algoritma C4.5 sebagaimana dalam penelitian ini dalam penentuan nasabah patuh dan tidak patuh, tidak akan memiliki manfaat bila tidak didukung oleh manajemen yang baik salah satunya berupa kemauan serta kemampuan yang baik dalam penerapannya oleh semua pihak yang terlibat. Tim *Department Collection* sebagai pelaksana kecil yang ditunjuk oleh Manager harus mampu menyediakan dan menggunakan Hardware dan software yang baik.

Aspek Teknis Sistem

Seberapapun handalnya sebuah sistem dan metode perlu adanya pemeliharaan baik dari sisi hardware, software bahkan sumber daya manusia sekalipun. Setelah rule ketepatan waktu terbentuk dan teruji bisa diterapkan pada penentuan nasabah patuh dan tidak patuh maka perlu diadaptasikan dengan konsisten pada Sistem dan Aplikasi penentuan nasabah patuh dan tidak patuh bahkan perlu dilakukan *tracking* dan *tracing* sesuai dengan rule yang telah terbentuk sehingga prosedur dan penentuan selalu mengacu pada rule yang ada. Berikutnya perlu ditambah dengan monitoring dan evaluasi terhadap konsistensi pelaku sistem dan

aplikasi penentuan nasabah patuh dan tidak patuh dalam penerapan rule yang ada sehingga akan didapat akuntabilitas sistem dan kredibilitas PT. Wahana Ottomitra Multiartha.

Aspek Penelitian Lanjutan

Untuk lebih meningkatkan optimasi dan agar lebih dapat meningkatkan akurasi, dalam penelitian berikutnya diharapkan peneliti berikutnya melakukan pruning terhadap algoritma C4.5 sehingga pohon yang terbentuk tidak terlalu besar bahkan mungkin untuk jumlah data yang besar sekali pun. Ini dilakukan untuk mengefisienkan kinerja dari algoritma C4.5 tanpa mengurangi keakuratannya. Algoritma model akurasi untuk peningkatan akurasi yang lain seperti Algoritma *Particle Swarm Optimization* (PSO) yang mampu memberikan bobot attribut berupa nilai desimal.

DAFTAR PUSTAKA

- [1] V. S. Moertini, B. Sitohang, and O. S. Santosa, “Integrasi Algoritma Pohon Keputusan C4.5 Yang Dikembangkan Ke Dalam Object-Relational Dbms,” *JUTI J. Ilm. Teknol. Inf.*, vol. 6, no. 2, p. 39, 2007.
- [2] N. D. Prayoga, “Penerapan Algoritma C.45 Dalam Memprediksi Kelulusan Tepat Waktu Pada Perguruan Tinggi (Studi Kasus : Stmik Royal Kisaran),” 2018.
- [3] M. F. Arifin and D. Fitrianah, “Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada,” no. January 2018, 2018.
- [4] W. Ahmed and A. L. I. Saeed, “A Comparative Study on Machine Learning Tools Using WEKA and Rapid Miner with Classifier Algorithms C4 . 5 and Decision Stump for Network Intrusion Detection,” vol. VII, no. 2, pp. 852–861, 2019.
- [5] Sucipto, Kusrini, and E. L. Taufiq, *Classification method of multi-class on C4.5 algorithm for fish diseases*, no. October. 2017.
- [6] D. B. Febriyanto, L. Handoko, and H. Aisyah, “Implementasi Algoritma C4 . 5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop,” vol. 5, no. 6, pp. 569–575, 2018.