

ANALISIS KELAYAKAN MENERIMA PINJAMAN KREDIT DENGAN ALGORITMA C4.5 PADA PT BPR BUANA ARTA MULIA

David¹,
Erlin Elisa²

¹Mahasiswa Program Studi Sistem Informasi, Universitas Putera Batam

²Dosen Program Studi Sistem Informasi, Universitas Putera Batam
email: pb181510003@upbatam.ac.id

ABSTRACT

Credit is one of the easiest alternative solutions for those who need financial support, either to be used to meet consumptive needs, as business capital or for other purposes. With credit assistance, people's living standards will be better and other needs that could not be realized before can be fulfilled to the fullest, which makes many people choose credit as a solution for funding needs. PT BPR Buana Arta Mulia is a banking company which one of its operational activities is to provide services for providing funds or credit loans to prospective debtors, problems that are still often faced by banks so far are regarding arrears made by debtors and several cases related to bad loans. So far, the bank has a collection of data regarding the credit application history of prospective debtors that has not been used optimally so that the author intends to conduct an analysis of the data to find out the hidden rules in determining the feasibility of receiving credit loans with the aim of improving the quality of the credit analysis results as well as an evaluation material for determining creditworthiness in the future. In this study, data mining classification techniques will be used with the C4.5 algorithm model in carrying out the analysis process and to ensure the correctness of the decisions obtained through manual calculations, testing will be carried out using the WEKA 3.9.5 software. the results of the rules generated either from manual calculations or using WEKA are the same and from the results obtained indicate that there are two variables that have the most influence in determining creditworthiness, namely from collateral value and income variables.

Keywords: *Algorithm C4.5; Credit; Creditworthiness; Data Mining.*

PENDAHULUAN

Permintaan akan kredit terus meningkat setiap tahunnya hal tersebut terjadi karena beberapa faktor, yang pertama yaitu kondisi ekonomi masyarakat yang kurang baik sehingga tidak dapat memenuhi keperluan

kebutuhan tertentu secara maksimal khususnya pada tengah masa pandemi covid-19 saat ini, yang kedua dikarenakan oleh proses dari transaksi pengajuan kredit yang menjadi lebih mudah dan cepat dari sebelumnya berkat bantuan dari teknologi informasi yang terus berkembang dengan pesat.

Salah satu contoh dari implementasi teknologi informasi tentang kredit adalah aplikasi seperti peminjaman dana *online* yang sekarang banyak dapat ditemui, yang dimana proses peminjaman dapat dilakukan secara daring dan bisa diselesaikan dalam waktu yang sangat singkat. Dengan adanya bantuan dari kredit kualitas dari kehidupan masyarakat dapat ditingkatkan dan keperluan akan kebutuhan primer, komsumtif, modal kerja ataupun keperluan lainnya dapat terpenuhi secara maksimal. Pinjaman kredit dapat diperoleh melalui beberapa cara salah satunya yaitu melalui aplikasi jasa peminjaman *online* namun terdapat batasan pada jumlah dana yang dapat dipinjam, sehingga peminjaman kredit dalam jumlah besar lebih disarankan diajukan melalui bank.

PT BPR Buana Arta Mulia adalah sebuah perusahaan perbankan yang salah satu kegiatan operasionalnya yaitu memberikan pinjaman dana atau kredit kepada calon debitur. Perusahaan ini berlokasi di Jalan Ampera No.03, Kabupaten Karimun, Provinsi Kepulauan Riau dan didirikan pada tahun 2009 dan telah beroperasi selama sepuluh tahun lebih. Kendala yang masih banyak ditemukan oleh pihak bank adalah masalah tunggakan dari debitur dan sedikit kasus kredit macet, dan selama ini pihak bank memiliki sekelompok data tentang *history* keputusan kredit terdahulu yang belum dimanfaatkan secara optimal, maka penulis ingin melakukan proses analisis terhadap data yang telah disebutkan sebelumnya untuk memperoleh tentang suatu aturan yang tersembunyi dalam melakukan penentuan kelayakan kredit dengan tujuan untuk memperbaiki kualitas hasil dari proses analisa kredit dan juga sebagai masukkan informasi dalam penentuan kelayakan kredit selanjutnya. Dan teknik yang akan diimplementasikan dalam membantu melakukan proses analisis adalah teknik klasifikasi dari *data mining* dengan membuat sebuah *decision tree* memakai jenis algoritma C4.5 dan diharapkan penelitian ini dapat

membawa manfaat serta pengetahuan baru bagi semua orang.

KAJIAN TEORI

2.1 Data Mining

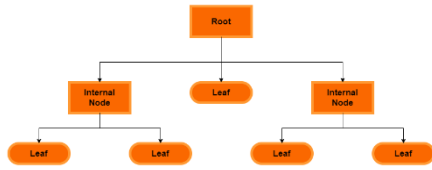
Data Mining adalah sebuah bentuk proses yang dilakukan untuk mengetahui pola aturan ataupun *relation* baru yang tersembunyi antar variabel dalam sebuah data yang disimpan pada basis data dengan menggunakan metode ataupun teknik tertentu untuk melakukan pemisahan data dari sekelompok data dalam jumlah besar (Iriadi & Nuraeni, 2016).

2.2 Klasifikasi

Klasifikasi merupakan salah satu teknik yang terdapat pada *data mining*, teknik ini dapat digunakan untuk melakukan pemisahan terhadap sebuah *object* kedalam beberapa *class* yang berbeda berdasarkan dari banyaknya *class* yang diinginkan, serta teknik ini dapat digunakan juga dalam membantu menemukan *pattern* yang bisa memisahkan kelas data dengan satu sama lain untuk penentuan *object* yang masuk dalam kategori kelompok tertentu, yaitu dengan cara melakukan perbandingan dengan karakteristik dari variabel kelompok yang sudah ditetapkan (Romli & Zy, 2020).

2.3 Pohon Keputusan

Pohon keputusan adalah sebuah struktur pohon yang terbentuk dari gabungan beberapa macam simpul yang pada umumnya dari yang paling rendah terdiri dari simpul daun yang merupakan perwakilan dari kelas cabang tertentu, pada tingkatan selanjutnya terdapat simpul internal yang merupakan simpul yang masing-masing mewakili sebuah *variable* atau *attribut* dan pada tingkatan tertinggi terdapat simpul akar yang merupakan *variable* yang paling berpengaruh diantara yang lainnya (Salim & Giap, 2017).



Gambar 1. Model Pohon Keputusan
Sumber : (Penulis, 2022)

2.5 Algoritma C4.5

Algoritma C4.5 merupakan sebuah model algoritma yang terdapat pada teknik klasifikasi *data mining* yang biasanya digunakan dalam melakukan pembentukan sebuah *decision tree* dengan melakukan proses tahapan klasifikasi terhadap data, dan hasil yang diperolehnya biasanya berupa informasi tentang *pattern* ataupun *rules* tersembunyi yang belum diketahui sebelumnya yang dimana dapat dimanfaatkan dalam membantu proses pengambilan keputusan.

Berikut dibawah adalah rumus yang akan digunakan dalam melakukan perhitungan untuk mencari nilai dari *gain* tertinggi dan *entropy*.

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

Rumus 1. Rumus Perhitungan Gain

Berikut adalah penjelasan simbol dari rumus diatas :

- S : himpunan kasus
- A : atribut atau deskripsi data
- n : jumlah partisi pada atribut A
- |S_i| : total kasus pada partisi yang ke-i
- |S| : total kasus yang terdapat pada S

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

Rumus 2. Rumus Perhitungan Entropy

Berikut adalah penjelasan simbol dari rumus diatas :

- S : himpunan kasus
- A : atribut atau deskripsi data

n : jumlah partisi pada S
 pi : proporsi dari S_i kepada S

2.4 Kredit

Kredit adalah kegiatan peminjaman dana kepada pihak calon debitur oleh pemberi pinjaman yang dilakukan berdasarkan dengan perjanjian yang telah disepakati bersama, yaitu dengan persyaratan bahwa peminjam harus bisa melakukan pelunasan atas semua utangnya serta seluruh bunga pinjaman dalam kurun waktu yang sudah disetujui sesuai perjanjian (Khasanah, 2019).

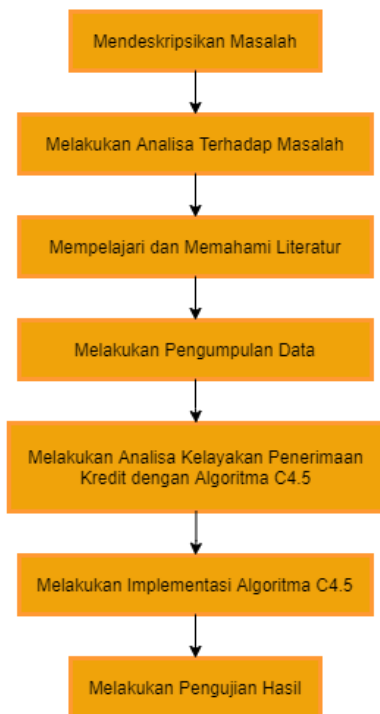
2.5 Kelayakan Kredit

Kelayakan kredit adalah kegiatan dimana dilakukanya sebuah proses pertimbangan atau estimasi terhadap kesanggupan dari pihak peminjam dalam memenuhi tanggung jawabnya untuk membayar dana yang telah dipinjamkan, dan penentuan seseorang apakah layak diberikan pinjaman pada umum diputuskan oleh seorang analis kredit (Dyah, Nur, & Heny, 2019).

METODE PENELITIAN

3.1 Desain Penelitian

Agar langkah penelitian yang dilakukan menjadi lebih terarah dan terstruktur maka akan dibuatkan sebuah desain penelitian yang akan menjelaskan seluruh tahapan langkah yang akan dilalui dalam penelitian yaitu sebagai berikut.



Gambar 2. Model Desain Penelitian

Sumber : (Penulis, 2022)

Berikut dibawah adalah deskripsi tentang tahapan langkah-langkah pada desain penelitian diatas.

1. Langkah awal yang perlu dilakukan pertama yaitu mendeskripsikan persoalan atau *problem* yang ditemukan pada objek penelitian yang sedang diteliti.
2. Langkah kedua yang perlu dilakukan yaitu melakukan kegiatan analisis terkait persoalan yang sudah dideskripsi pada langkah pertama.
3. Langkah ketiga yang perlu dilakukan adalah melakukan kegiatan pembelajaran teori ataupun literatur yang mempunyai kaitan dengan penelitian yang dilakukan.
4. Pada langkah keempat akan diadakan kegiatan pengumpulan terhadap data yang akan digunakan dalam mendukung proses penelitian.

5. Pada langkah kelima akan dilakukan pengolahan terhadap data yang telah dikumpulkan pada langkah keempat sesuai dengan format data yang cocok untuk pengujian menggunakan algoritma C4.5.
6. Pada langkah keenam akan dilakukan pengujian menggunakan algoritma C4.5 terhadap data yang telah diolah pada langkah sebelumnya, dengan mencari nilai dari *entropy*, *gain* serta pembentukan pohon keputusan.
7. Pada langkah terakhir yaitu melakukan pengujian hasil akhir menggunakan *software* WEKA versi 3.9.5.

3.2 Objek Penelitian

Pada penelitian yang dilakukan yang menjadi sebagai objek penelitian adalah PT BPR Buana Arta Mulia sebuah perusahaan perbankan yang berlokasi pada Jalan Ampera No. 03, Kabupaten Karimun, Provinsi Kepulauan Riau.

3.3. Populasi dan Sampel

Pada penelitian yang dilakukan yang menjadi sebagai populasi adalah data tentang pengajuan kredit dari pihak debitur, yang diambil dari PT BPR Buana Arta Mulia dan datanya merupakan data dari periode bulan Januari sampai dengan bulan September tahun 2021. Sedangkan untuk sampelnya dipilih dengan memakai teknik sampling jenuh dengan jumlah sebanyak 40 data.

3.4 Variabel Penelitian

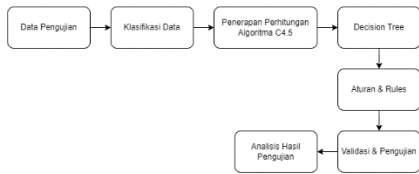
Berikut dibawah adalah semua variabel yang akan digunakan dalam penelitian sebagai faktor penilain kelayakan kredit.

1. Pendapatan
2. Jumlah Kredit
3. Tujuan Kredit
4. Nilai Jaminan

3.5 Model Penelitian

Pada penelitian ini akan digunakan model penelitian *data mining* dengan teknik *classification* dan untuk pembentukan pohon keputusannya akan

digunakan algoritma C4.5. tahap pertama yang nantinya akan dilakukan adalah mengklasifikasi data sesuai dengan golongan kelompok datanya selanjutnya akan dimulai pencarian terhadap nilai *entropy* serta *gain* nya untuk membuat *decision tree* nya dan proses tersebut akan di repetitif hingga node tidak mempunyai keturunan lagi dan disitu akan didapatkan aturan atau *rules* untuk penentuan kelayakan kredit, setelah itu akan dilakukan pengujian akhir untuk meyakinkan *result* yang diterima.



adalah dari variabel nilai jaminan dan pendapatan.

Berikut dibawah adalah daftar dari himpunan variabel yang digunakan

Gambar 3. Model Penelitian
Sumber : (Penulis, 2022)

HASIL DAN PEMBAHASAN

4.1 Hasil

Proses pengujian menggunakan *software* WEKA telah dilakukan pada data pengujian dan juga telah memberikan hasil keputusan dalam bentuk aturan atau *rules*, baik pengujian menggunakan *software* maupun perhitungan secara manual masing-masing menghasilkan nilai aturan yang sama yaitu menunjukkan bahwa variabel yang paling banyak memiliki pengaruh terhadap kelayakan kredit seseorang

dalam penelitian serta bentuk kategori nilai setelah dilakukanya proses klasifikasi.

Tabel 1. Himpunan Variabel

No.	Variabel	Kategori Nilai	Klasifikasi
1	Pendapatan	Tinggi	TG
		Sedang	SD
		Rendah	RD
2	Jumlah Kredit	Banyak	BYK
		Menengah	MNG
		Sedikit	SDK
3	Tujuan Kredit	Modal Usaha	MU
		Konsumsi	KS
4	Nilai Jaminan	Tinggi	TG
		Sedang	SD
		Rendah	RD

Sumber : (Penulis, 2022)

Tabel 2. Format Akhir Data Pengujian

No.	Pendapatan	Jumlah Kredit	Tujuan Kredit	Nilai Jaminan	Kelayakan Kredit
1	RD	BYK	MU	TG	Layak
2	RD	BYK	MU	RD	Tidak Layak
3	SD	MNG	KS	TG	Layak
4	SD	BYK	KS	RD	Layak
5	SD	BYK	MU	RD	Layak
6	TG	MNG	MU	SD	Layak
7	RD	MNG	MU	RD	Tidak Layak
8	RD	BYK	MU	RD	Tidak Layak
9	RD	MNG	KS	RD	Layak

10	SD	BYK	MU	TG	Layak
11	RD	BYK	MU	RD	Tidak Layak
12	RD	SDK	KS	SD	Layak
13	RD	BYK	KS	TG	Layak
14	TG	BYK	MU	TG	Layak
15	SD	BYK	MU	RD	Layak
16	SD	SDK	KS	SD	Layak
17	TG	BYK	KS	RD	Layak
18	RD	MNG	KS	RD	Layak
19	RD	MNG	KS	RD	Tidak Layak
20	RD	BYK	MU	RD	Tidak Layak
21	TG	MNG	MU	SD	Layak
22	TG	SDK	MU	RD	Layak
23	SD	SDK	KS	TG	Layak
24	SD	SDK	MU	SD	Layak
25	RD	SDK	KS	RD	Layak
26	RD	MNG	KS	RD	Tidak Layak
27	RD	BYK	MU	TG	Layak
28	SD	BYK	KS	TG	Layak
29	SD	BYK	MU	SD	Layak
30	TG	BYK	KS	SD	Layak
31	SD	SDK	KS	RD	Layak
32	RD	BYK	KS	RD	Tidak Layak
33	TG	MNG	MU	TG	Layak
34	RD	MNG	KS	RD	Tidak Layak
35	RD	BYK	KS	RD	Tidak Layak
36	RD	BYK	MU	RD	Tidak Layak
37	TG	SDK	MU	TG	Layak
38	TG	SDK	MU	SD	Layak
39	SD	MNG	KS	TG	Layak
40	RD	BYK	KS	RD	Tidak Layak

Sumber : (Penulis, 2022)

4.2 Pembahasan

Dari format data akhir pengujian diatas, maka akan dilakukanya proses *classification* terhadap data berdasarkan dengan *class* masing-masing menggunakan algoritma C4.5 dalam pembentukan pohon keputusan untuk menentukan kelayakan kredit sesuai dengan atribut atau variabel yang sudah dipilih, yang dimana terdiri dari pendapatan, jumlah kredit, tujuan kredit dan nilai jaminan. Cara melakukan penentuan variabel sebagai simpul akar dapat dilakukan dengan mencari hasil

dari *gain* tertinggi dari seluruh variabel yang tersedia dan perhitungan pencarian dari hasil nilai *gain* akan dilakukan secara repetitif supaya simpul internal dan simpul daun dapat ditemukan hingga tidak lagi mempunyai keturunan, sehingga langkah awal yang harus dilakukan untuk membentuk pohon keputusan adalah mengadakan proses perhitungan secara manual dan sebelum mencari nilai *gain* akan dicari terlebih dahulu nilai *entropy* dari masing-masing kategori nilai variabel berikut dibawah adalah cara perhitunganya.

Tabel 3. Hasil Proses Perhitungan Simpul 1

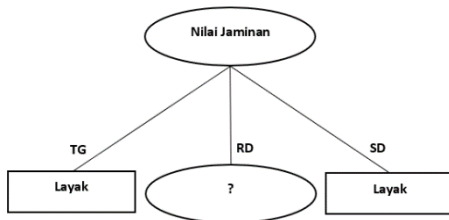
	Jumlah Kasus	Layak	Tidak Layak	Entropy	Gain
Total	40	28	12	0.8813	

Pendapatan

Tinggi	TG	9	9	0	0	
Sedang	SD	12	12	0	0	0.4303
Rendah	RD	19	7	12	0.9495	
Jumlah Kredit						
Banyak	BYK	20	12	8	0.9710	
Menengah	MNG	9	5	4	0.9911	0.1728
Sedikit	SDK	11	11	0	0	
Tujuan Kredit						
Modal Usaha	MU	20	14	6	0.8813	
Konsumsi	KS	20	14	6	0.8813	0
Nilai Jaminan						
Tinggi	TG	11	11	0	0	
Sedang	SD	11	11	0	0	0.4681
Rendah	RD	18	6	12	0.9183	

Sumber : (Penulis, 2022)

Dari hasil perhitungan tabel diatas, bisa dilihat bahwa variabel yang memiliki nilai dari *gain* tertinggi yaitu variabel nilai jaminan dengan nilai sebesar 0,4681, maka yang akan dijadikan sebagai simpul akar dari pohon keputusan sementara adalah variabel nilai jaminan sebagai berikut.



Gambar 4. Decision Tree Node 1
 Sumber : (Penulis, 2021)

Dari gambar pohon keputusan yang terbentuk diatas dapat dilihat bahwa simpul akar dari nilai jaminan mempunyai tiga cabang, dan ada dua cabang dengan hasil *entropy* yang bernilai nol yaitu kategori nilai tinggi (TG) dan sedang (SD) yang membuatnya tidak memiliki turunan sehingga akan dijadikan sebagai simpul daun, sedangkan untuk kategori nilai rendah (RD) masih terdapat keturunan pada simpul selanjutnya karena nilai hasil dari *entropy* belum berada pada angka nol dan masih mempunyai nilai sebesar 0,9183. Selanjutnya akan dilakukan pencarian dari simpul kedua dengan cara perhitungan yang sama pada simpul akar yaitu diawali dengan melakukan pencarian dari nilai *entropy* serta *gain* dengan variabel lainnya yang tersisa selain variabel nilai jaminan.

Tabel 4. Hasil Proses Perhitungan Simpul 2

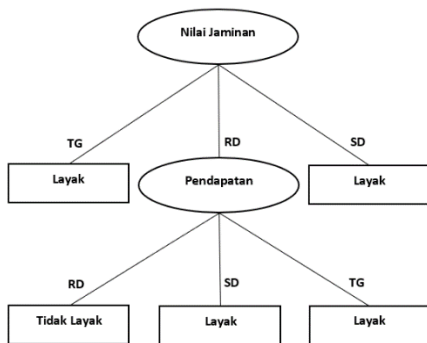
	Jumlah Kasus	Layak	Tidak Layak	Entropy	Gain
Total	18	6	12	0.9183	
Pendapatan					
Tinggi	TG	2	2	0	
Sedang	SD	4	4	0	0.9183
Rendah	RD	12	0	12	0

Jumlah Kredit						
Banyak	BYK	12	4	8	0.9183	
Menengah	MNG	4	0	4	0	0.3061
Sedikit	SDK	2	2	0	0	
Tujuan Kredit						
Modal Usaha	MU	9	3	6	0.9183	
Konsumsi	KS	9	3	6	0.9183	0

Sumber : (Penulis, 2022)

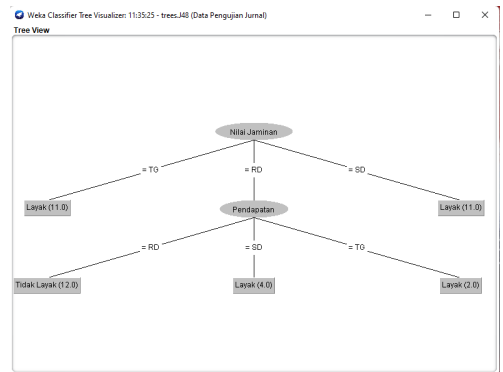
Dari hasil perhitungan tabel diatas, dapat diketahui bahwa variabel yang memiliki nilai dari *gain* tertinggi yaitu variabel pendapatan dengan nilai sebesar 0,9183, sehingga akan dijadikan sebagai simpul internal kedua dari pohon keputusan, dan selain itu dapat

diketahui juga bahwa *entropy* dari semua kategori nilai variabel pendapatan telah bernilai nol dan tidak diperlukan untuk pencarian lebih lanjut, supaya lebih jelas dibawah adalah *decision tree* yang dihasilkan.



Gambar 5. Decision Tree Node 2
Sumber : (Penulis, 2022)

Proses perhitungan secara manual telah dilakukan, untuk memastikan hasil keputusan yang didapatkan adalah benar akan dilakukan pengujian menggunakan bantuan *software WEKA*, berikut dibawah adalah hasil pohon keputusan yang terbentuk.



Gambar 6. Decision Tree Node 2 WEKA
Sumber : (Penulis, 2022)

Berdasarkan dengan *decision tree* yang telah terbentuk baik dari perhitungan secara manual maupun menggunakan WEKA, maka dapat dijelaskan bahwa aturan atau *rules* yang diperoleh adalah sebagai berikut.

1. IF Nilai Jaminan = (TG), THEN Hasil Keputusan = Layak
2. IF Nilai Jaminan = (SD), THEN Hasil Keputusan = Layak
3. IF Nilai Jaminan = (RD), AND Pendapatan = (TG) THEN Hasil Keputusan = Layak

4. IF Nilai Jaminan = (RD), AND Pendapatan = (SD) THEN Hasil Keputusan = Layak
5. IF Nilai Jaminan = (RD), AND Pendapatan = (RD) THEN Hasil Keputusan = Tidak Layak

SIMPULAN

Berikut dibawah adalah beberapa kesimpulan yang dapat ditarik dari hasil analisa dan pengujian yang telah dilakukan tentang kelayakan penerimaan pinjaman kredit.

1. Perhitungan secara manual maupun menggunakan aplikasi WEKA masing-masing menghasilkan aturan keputusan yang sama yaitu dengan menunjukkan bahwa variabel nilai jaminan dan pendapatan sebagai faktor pengaruh paling besar dalam penentuan kelayakan kredit, yang pada akhirnya membuktikan bahwa klasifikasi dengan menggunakan algoritma C4.5 cukup akurat dan dapat digunakan untuk pengidentifikasian kelayakan penerimaan kredit.
2. Aturan *rules* yang dihasilkan sangat bermanfaat dan membantu dalam meningkatkan kualitas serta akurasi dari hasil analisa yang dilakukan untuk kedepanya.

DAFTAR PUSTAKA

Dyah, W., Nur, L., & Heny, S. (2019). Metode Algoritma Decision Tree C4.5 Untuk Analisis Kelayakan

Kredit Nasabah Pada BSM KCP Kemang Pratama. *Jurnal Sains Dan Manajemen*, 7(2), 36–42.

Iriadi, N., & Nuraeni, N. (2016). Kajian Penerapan Metode Klasifikasi Data Mining Algoritma C4.5 Untuk Prediksi Kelayakan Kredit Pada Bank Mayapada Jakarta. *Jurnal Teknik Komputer AMIK BSI*, 2(1), 132–137.

Khasanah, S. N. (2019). *Komparasi Algoritma C4.5 dan Naive Bayes Untuk Menganalisis Kelayakan Menerima Pinjaman Kredit*. 3(3), 94–98.

	<p>Biodata, Penulis pertama, David, merupakan mahasiswa Prodi Sistem Informasi UniversitasPutera Batam.</p>
	<p>Biodata, Penulis kedua, Erlin Elisa, S.Kom., M.Kom. merupakan Dosen Prodi Sistem Informasi Universitas Putera Batam. Penulis banyak berkecimpung di bidang sistem informasi.</p>

Romli, I., & Zy, A. T. (2020). *Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5*. 4, 694–702.

Salim, T., & Giap, Y. C. (2017). Data Mining Identifikasi Website Phising Menggunakan Algoritma C4.5. *Jurnal TAM (Technology Acceptance Model)*, 8, 130–135.