

# Klasifikasi Menggunakan Algoritma C4.5 dalam Memprediksi Kriteria Nasabah Pembiayaan Syariah BMT 071

M. Wahyu Santoso

Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Darwan Ali, Sampit, Indonesia  
Email: 'wahyu.santoso2900@gmail.com

Submitted 30-06-2020; Accepted 28-08-2020; Published 31-08-2020

## Abstrak

Seperti yang diketahui bahwa salah satu tugas dari bidang Pembiayaan adalah melakukan analisis data nasabah yang ada di suatu koperasi untuk mengetahui nasabah - nasabah mana saja yang berpotensi agar bisa melakukan kredit kembali. Adapun metode yang dapat digunakan untuk menganalisis data nasabah yang ada yaitu dengan cara mengklasifikasikan seluruh nasabah yang pernah melakukan atau menyelesaikan angsuran pembiayaannya ke dalam target pemasaran selanjutnya. Oleh karena itulah penelitian ini dilakukan untuk membantu menyelesaikan permasalahan diatas dengan cara mengklasifikasikan para nasabah lama tersebut menggunakan Algoritma C4.5 yang dapat berfungsi untuk memprediksi kriteria nasabah pinjaman dana yang berpotensi melakukan peminjaman dana kembali terhadap BMT 071. Bagian Pembiayaan BMT 071 yang berlokasi di Kalimantan Tengah, Kabupaten Kotawaringin Timur ini merupakan tempat yang dipilih oleh peneliti sebagai studi kasus, dengan asumsi Bagian Pembiayaan BMT 071 ini telah mengalami permasalahan seperti yang telah dijelaskan di atas. Oleh karena itu penulis menggunakan teknik *data mining* yang menggunakan teknik klasifikasi *Decision Tree* (Pohon Keputusan). Sedangkan untuk algoritma yang dipakai sebagai algoritma pembantu untuk membentuk pohon keputusannya adalah Algoritma C4.5. Untuk dataset yang di proses dalam penelitian ini adalah data angsuran nasabah pembiayaan BMT 071 pada bulan Agustus 2018 sampai dengan Maret 2020 dalam format Microsoft Excel yang berjumlah 313 Nasabah. Hasil penelitian ini adalah sebuah rancangan aplikasi yang dapat memudahkan bagian Pembiayaan di BMT 071 dalam memperoleh target pembiayaan di masa yang akan datang.

**Kata Kunci:** Pembiayaan, Koperasi, *Data mining*, Prediksi, *Decision Tree*, Algoritma C4.5, Koperasi Syariah

## Abstract

As it is known that one of the tasks of the Financing field is to analyze customer data in a cooperative to find out which customers have the potential to be able to do credit again. The method that can be used to analyze existing customer data is by classifying all customers who have made or completed their financing installments into subsequent marketing targets. That is why this research was conducted to help resolve the above problems by designing a *data mining* application that can serve to predict the criteria of loan fund customers who have the potential to borrow funds back to the BMT 071. Financing Section of BMT 071 located in Central Kalimantan, Regency Kotawaringin Timur is a place chosen by researchers as a case study, assuming the BMT 071 Financing Section has experienced problems as explained above. Therefore, the authors use *data mining* techniques that are applied to the application to be built is the Decision Tree classification technique. As for the algorithm used as a helper algorithm to form a decision tree is the C4.5 Algorithm. For the dataset processed in this study is the installment data for BMT 071 financing customers from August 2018 to March 2020 in Microsoft Excel format. The results of this study are a design application that can facilitate the Financing section of the BMT 071 in obtaining financing targets in the future.

**Keywords:** Pembiayaan, Koperasi, *Data mining*, Prediksi, *Decision Tree*, Algoritma C4.5, Koperasi Syariah

## 1. PENDAHULUAN

Perlu diketahui BMT 071 merupakan sebuah koperasi syaria'ah yang salah satu tugasnya adalah memberikan pelayanan pinjaman / pembiayaan dana kepada para nasabah, di BMT 071 mempunyai beberapa bidang / bagian contohnya bagian Pembiayaan, tugas utama bagian Pembiayaan BMT 071 selama ini adalah mencari nasabah lama yang pernah melakukan pembiayaan di BMT 071 untuk menawarkan kembali Pembiayaan dan menghimpun dana Masyarakat Kotim, untuk saat ini jumlah Anggota Nasabah BMT 071 adalah 4710 orang. Adapun Metode yang digunakan untuk menentukan target Pembiayaan selama ini adalah menganalisa dan mengira-ngira semua nasabah lama yang berjumlah kurang lebih 313 nasabah yang telah melunasi angsuran pembiayaannya ke dalam target pemasaran selanjutnya, padahal seperti yang Kita ketahui belum tentu efektif dan semua nasabah akan tertarik melakukan pembiayaan lagi.

Koperasi merupakan salah satu lembaga keuangan yang melakukan transaksi kredit. Dalam rangka mengembangkan usahanya dalam bentuk pinjaman anggota, koperasi memiliki prinsip kehati-hatian sebagaimana yang diterapkan pada perbankan dengan melakukan analisis pemberian kredit kepada calon anggotanya [2].

Oleh karena itulah, untuk meningkatkan loyalitas dan kepercayaan para nasabah lama BMT 071, bagian Pembiayaan ini dituntut agar lebih cepat dan tepat dalam mengambil keputusan dalam menentukan strategi pemasaran selanjutnya, baik produk utama Pembiayaan maupun layanan jasa Gadai yang sering ditawarkan kepada nasabah khususnya dalam masalah pembiayaan. Hal ini dapat dilakukan apabila Tim Pembiayaan memperoleh informasi yang cukup untuk membantu proses pengambilan keputusan, salah satunya yaitu dengan cara melakukan prediksi kriteria nasabah lama untuk kepentingan target pemasaran.

Adapun teknik *Data mining* ini dapat digunakan oleh bagian Pembiayaan untuk menganalisis pasar sehingga seorang Admin Pembiayaan dapat memahami pasar yang sudah ada ataupun menemukan peluang-peluang yang baru untuk meningkatkan keuntungan yang didapatkan, termasuk memprediksi kriteria target pasar yang ada. Oleh karena itulah, penelitian ini difokuskan pada pemanfaatan *data mining* untuk mengklasifikasikan dan memprediksi kriteria nasabah pembiayaan BMT 071, sehingga dapat diketahui apakah nasabah yang tersebut merupakan nasabah lama yang dapat

berpotensi menjadi nasabah pembiayaan yang sangat berpotensi melakukan pembiayaan kembali atau tidak.

## 2. METODE PENELITIAN

### 2.1 Metode Pengumpulan Data

Dalam penelitian dan penyusunan serta pengumpulan data penulis menggunakan metode – metode sebagai berikut :

1. Wawancara  
Wawancara ini dilakukan dengan cara mengadakan wawancara secara langsung dengan salah satu karyawan yang bertanggung jawab atas pengelolaan data Pembiayaan pada BMT 071 Sampit yakni Pak Wahyu agar sistem yang dibuat sesuai dengan kebutuhan yang ada di perusahaan.
2. Pengamatan  
Dilakukan dengan mengamati langsung kegiatan pemasaran pembiayaan pada perusahaan BMT 071 Sampit agar dapat memahami sistem pembiayaan yang saat ini sedang berjalan.
3. Studi Pustaka  
Pengumpulan data dengan mempelajari metode yang berhubungan dengan Pembiayaan Syariah dan mempelajari aplikasi-aplikasi yang dipakai dalam penerapan *Data mining*, oleh karena itu penulis banyak mengutip, mempelajari dari buku-buku elektronik yang berhubungan langsung dengan masalah diatas.

### 2.2 Teori Pendukung

#### a. Konsep Mendasar *Data mining*

*Data mining* merupakan salah satu proses yang mengacu pada Teknik *mining* (menambang) yang digunakan untuk menganalisa data dari pandangan atau perspektif yang berbeda dan menyimpulkannya menjadi informasi- informasi penting dan juga dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya yang dikeluarkan, atau bahkan keduanya. *Data mining* secara teknis dapat disebut juga sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan *field* dari sebuah relasional *database* yang besar (dataset).

Secara umum kegunaan data mining dapat dibagi menjadi dua, yaitu deskriptif dan prediktif. Deskriptif berarti data mining digunakan untuk mencari pola-pola yang dapat dipahami manusia yang menjelaskan karakteristik data. Sedangkan prediktif berarti data mining digunakan untuk membentuk sebuah model pengetahuan yang akan digunakan untuk melakukan prediksi [1].

Saat pengaplikasiannya, *data mining* adalah salah satu bagian dari proses *KDD* (*Knowledge Discovery in Database*) yang mempunyai tugas untuk mengekstrak suatu pola atau model dari data dengan menggunakan suatu algoritma yang spesifik. Adapun beberapa proses *KDD* adalah sebagai berikut :

1. *Data Selection* : Pemilihan data dari sekumpulan data operasional yang perlu dilakukan sebelum tahap penggalian informasi ketika *KDD* dimulai.
2. *Preprocessing* : Saat sebelum proses *data mining* dilaksanakan, perlu dilakukan terlebih dahulu proses *cleaning*, yang bertujuan untuk membuang duplikasi data atau data yang sama, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan yang ada pada data, seperti kesalahan cetak atau penulisan (tipografi).
3. *Transformation* : Proses *coding* pada data yang telah dipilih sebelumnya, sehingga data tersebut sesuai untuk di proses *data mining*. Proses *coding* dalam *KDD* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam suatu database.
4. *Data mining* : Suatu proses pencarian pola atau informasi menarik dalam data terpilih dengan menggunakan teknik *mining* (menambang) atau metode tertentu.
5. *Evaluation / Interpretation*: Sebuah pola informasi yang telah dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak manapun. Tahap ini merupakan salah satu bagian dari proses *KDD* yang disebut dengan *interpretation*. Pada tahap ini mencakup pemeriksaan apakah sebuah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak.

#### b. Klasifikasi

Klasifikasi merupakan suatu proses mencari suatu himpunan model (fungsi) yang dapat digunakan untuk mendeskripsikan dan membedakan kelas-kelas data atau konsep-konsep, yang bertujuan agar dapat menggunakan model tersebut untuk memprediksi kelas dari suatu objek yang saat kelasnya belum diketahui.

#### c. Algoritma C4.5

Ada beberapa pengembangan yang dilakukan pada Algoritma C4.5 antara lain adalah dapat mengatasi missing value, dapat mengatasi kontinu data, dan pruning. Pada artikel jurnal di [4] menyebutkan bahwa Algoritma C4.5 merupakan salah satu teknik decision tree yang sering digunakan, yang menghasilkan beberapa aturan-aturan dan sebuah pohon keputusan dengan tujuan untuk meningkatkan keakuratan dari prediksi yang sedang dilakukan, di samping itu algoritma C4.5 merupakan algoritma penerus *Decision Tree* yang dapat dengan mudah untuk dimengerti.

Secara umum Algoritma C4.5 untuk membangun sebuah pohon keputusan adalah dengan beberapa cara seperti berikut ini :

1. Pilih suatu atribut sebagai akar
2. Buatlah cabang untuk setiap nilai
3. Pembagian kasus dalam cabang
4. Melakukan perulangan pada setiap cabang sampai semua kasus di setiap cabang memiliki kelas yang sama.

#### d. Entropy dan Informasi Gain

Setiap Entropy dan Informasi Gain pada sebuah obyek yang diklasifikasikan dalam pohon keputusan harus diletakkan terlebih dahulu nilai Entropy nya. Entropy merupakan suatu ukuran dari teori informasi yang dapat mengetahui karakteristik dari impurity dan homogeneity dari sekumpulan data. Dari nilai Entropy itulah kemudian dihitung nilai information gain dari masing-masing atribut. Entropy (S) merupakan jumlah bit yang dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. Entropy dapat juga dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas, semakin kecil nilai Entropy nya maka akan semakin besar peluang Entropy digunakan untuk mengekstrak suatu kelas. Besarnya suatu Entropy pada suatu ruang sampel S dapat didefinisikan sebagai berikut :

$$\text{Entropy}(S) = -p(+) \log_2 p(+) - p(-) \log_2 p(-) \quad (1)$$

Keterangan:

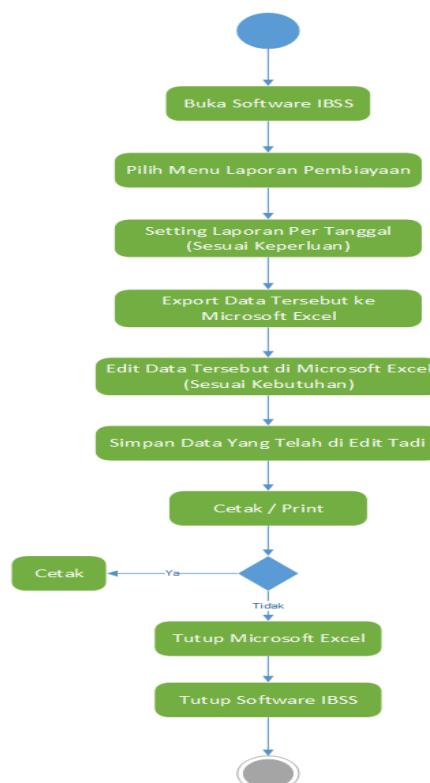
- S : Data sampel yang digunakan untuk Latihan.  
 $p(+)$  : Jumlah data yang bersolusi positif atau pendukung pada data sampel untuk kriteria tertentu.  
 $p(-)$  : Jumlah data yang bersolusi negatif atau pendukung pada data sampel untuk kriteria tertentu.

Information gain merupakan salah satu *attribute selection measure* yang biasanya digunakan untuk memilih test *attribute* tiap node pada pohon. Suatu atribut dengan informasi gain tertinggi dapat dipilih sebagai test atribut dari suatu node. Gain (S,A) merupakan sebuah perolehan informasi dari atribut A yang relative terhadap output data S. Perolehan informasi ini didapat dari output data atau variable dependent S yang dapat dikelompokkan berdasarkan atribut A, dan dinotasikan dengan gain (S,A).

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisa Data mining

Berdasarkan dari data yang telah dikumpulkan diketahui bahwa prosedur dalam pencarian nasabah yang dilakukan oleh Bagian Pembiayaan BMT 071, dapat dilihat pada Gambar 1. Berdasarkan dari hasil observasi yang telah dilakukan oleh peneliti, data nasabah pembiayaan BMT 071 tersimpan di dalam database navicat dari sebuah program *Integrated micro-Banking System Syariah (IBSS)* yang kemudian di *export* ke dalam *Microsoft Excel*, dimana tabel tersebut terdiri dari 17 *field* (tabel 1).



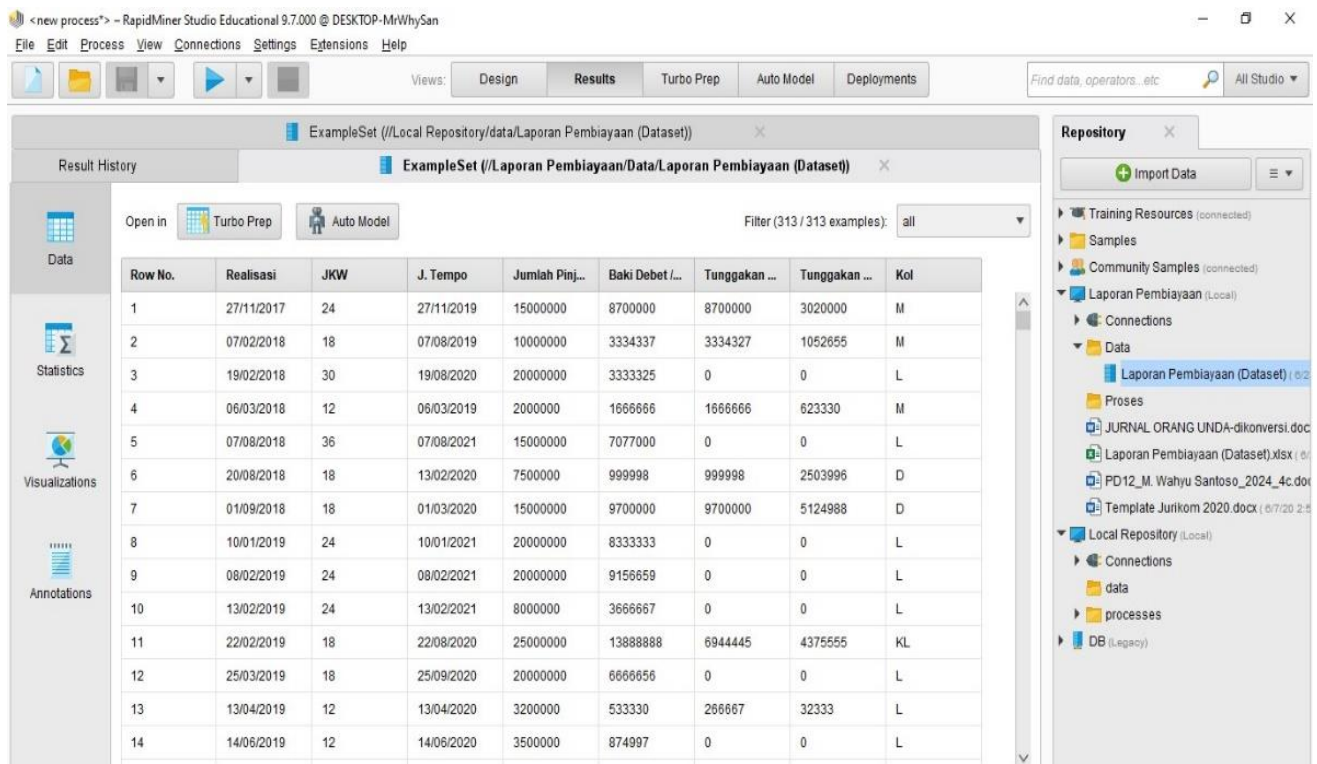
Gambar 1. Diagram Activity Prosedur Pencarian Nasabah

**Tabel 1.** Struktur tabel LAPORAN PEMBIAYAAN

Field
NO
NO REKENING
NAMA PEMINJAM
ALAMAT
TANGGAL REALISASI
JANGKA WAKTU
TANGGAL JATUH TEMPO
TYPE PINJAMAN
MARGIN
JUMLAH PINJAMAN
BAKI DEBET / SALDO POKOK
TUNGGAKAN POKOK
FT
TUNGGAKAN MARGIN
FT
JT
KOLEKTIF

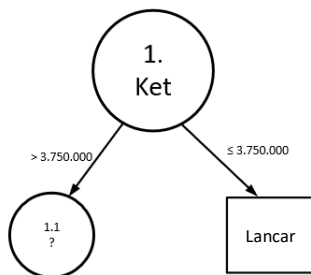
Pada teknik Klasifikasi ini, semua *output* dari setiap data yang dijadikan parameter target atau class harus berupa bilangan bulat atau diskrit. Berdasarkan Tabel 1, *field* yang dijadikan parameter target adalah *field* “KOLEKTIF” yang berisi nilai parameter Lancar dan Macet. Pada nilai parameter Lancar berarti nasabah pembiayaan tersebut sudah mempunyai riwayat bagus pada angsuran pembiayaan di BMT 071 sebelumnya, sedangkan Macet berarti kebalikannya yaitu mempunyai riwayat kurang bagus pada angsuran pembiayaan di BMT 071 sebelumnya. Berdasarkan *records* dari data pada *field* “Kolektif” tersebut, maka teknik Klasifikasi bisa diterapkan untuk melakukan *data mining* pada data tersebut. Adapun kolom yang bisa digunakan sebagai parameter - parameter pembentuk pohon keputusannya adalah sebagai berikut :

- a. TANGGAL REALISASI
- b. JANGKA WAKTU
- c. TANGGAL JATUH TEMPO
- d. JUMLAH PINJAMAN
- e. BAKI DEBET / SALDO POKOK
- f. TUNGGAKAN POKOK
- g. TUNGGAKAN MARGIN
- h. KOLEKTIF



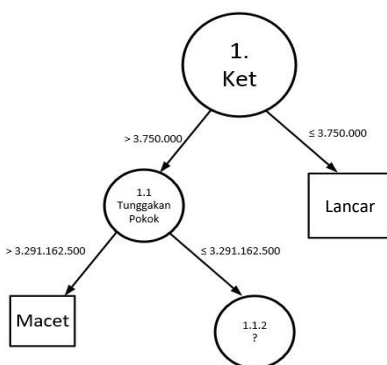
**Gambar 2.** Tampilan Parameter Keputusan Pada Software Rapidminer

Contoh Node 1 yang terbentuk dengan menggunakan algoritma C4.5 dapat dilihat pada gambar 3, pada node 1 itu dapat kita ketahui bahwa parameter Jumlah Pinjaman menjadi akar dari node pohon keputusan, karena atribut tersebut memiliki 2 nilai, maka atribut tersebut memiliki 2 cabang yaitu nilai Jumlah Pinjaman yang lebih besar dari 3.750.000 dan nilai lebih kecil sama dengan 3.750.000.



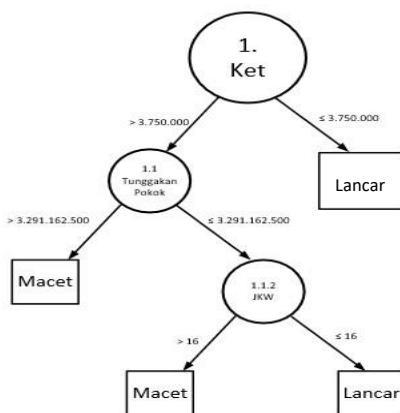
**Gambar 3.** Node 1

Kemudian untuk melanjutkan perhitungan untuk menemukan node 1.1 nya yang terbentuk dengan menggunakan algoritma C4.5 tadi, dapat dilihat pada gambar 4, pada node 1.1 itu dapat kita ketahui bahwa parameter Tunggalan Pokok adalah atribut yang mengisi tanda tanya pada node 1 tadi, karena atribut tersebut memiliki 2 nilai, maka atribut tersebut memiliki 2 cabang yaitu nilai Tunggalan Pokok yang lebih besar dari 3.291.162.500 dan nilai lebih kecil sama dengan 3.291.162.500.



**Gambar 4.** Node 1.1

Kemudian untuk melanjutkan perhitungan agar dapat menemukan node 1.1.2 nya, atau pohon keputusan akhir yang terbentuk dengan menggunakan algoritma C4.5 tadi, dapat dilihat pada gambar 5, pada node 1.1.2 itu dapat kita ketahui bahwa parameter JKW (Jangka Waktu) adalah atribut yang mengisi tanda tanya pada node 1.1 sebelumnya tadi, karena atribut tersebut memiliki 2 nilai, maka atribut tersebut memiliki 2 cabang, yaitu Jangka Waktu yang lebih besar dari 16 Bulan dan Jangka Waktu yang lebih kecil sama dengan 16 Bulan, dan Pohon Keputusan pun selesai dan tidak perlu dilakukan perhitungan kembali.



**Gambar 5.** Node 1.1.2

Kemudian hasil akhir dari pohon keputusan yang terbentuk dengan menggunakan algoritma C4.5 di atas tadi (Gambar 5), setelah pohon keputusan tersebut telah terbentuk secara utuh dan tidak ada perhitungan ulang lagi, maka dibuatlah aturan prediksi berdasarkan hasil pembentukan dari pohon keputusan tersebut. Pada pembentukan Pohon Keputusan tersebut dapat di peroleh aturan (*rules*) prediksi yang dapat kita simpulkan sebagai berikut:

1. Jika Jumlah Pinjaman yang di ajukan atau di berikan lebih dari Rp. 3.750.000, maka dapat di prediksi tidak berpotensi.
2. Jika Jumlah Pinjaman yang di ajukan atau di berikan lebih dari Rp. 3.750.000, maka dapat di prediksi masih berpotensi.
3. Jika Jumlah Tunggalan Pokok nya yang masih belum di bayarkan lebih dari Rp. 3.291.162.500, maka dapat di prediksi tidak berpotensi.
4. Jika Jumlah Tunggalan Pokok nya sudah di bayarkan kurang dari dari Rp. 3.291.162.500, maka dapat di prediksi berpotensi.
5. Jika Jumlah Jangka Waktu yang diberikan lebih dari 16 Bulan, maka dapat di prediksi tidak berpotensi.
6. Jika Jumlah Jangka Waktu yang diberikan kurang dari 16 Bulan, maka dapat di prediksi berpotensi.

#### 4. KESIMPULAN

Setelah Peneliti melakukan analisis terhadap para nasabah lama yang telah melakukan pembiayaan atau memiliki riwayat pembiayaan di BMT 071, telah disimpulkan bahwa mengklasifikasikan para nasabah lama tersebut dengan menggunakan algoritma C4,5 untuk menentukan target pemasaran pembiayaan dapat mempermudah bagian Pembiayaan BMT 071, sehingga diharapkan biaya operasional Bagian Pembiayaan BMT 071 dapat ditekan seminimal mungkin agar para staff Bagian Pembiayaan semakin efektif.

#### REFERENCES

- [1] Suyanto. "Data Mining Untuk Klasifikasi Dan Klasterisasi Data (1st ed.). Bandung:INFORMATIKA". 2017.
- [2] Pandie, E. S. Y. "Implementasi Algoritma Data Mining Naive Bayes Pada Koperasi. JICON, 6(1), 15–20". 2018.
- [3] Dela Sekardiana, Teguh Budi Santoso "Implementasi *Data Mining* Untuk Penentuan Kelayakan Pemberian Kredit Nasabah Menggunakan Algoritma C4.5 (Studi Kasus : Koperia / Koperasi Warga Komplek Gandaria)." 2017.
- [4] Abdul Haris Romdhoni "Analisis Komparasi Manajemen Risiko Pada Koperasi Syariah Di Kabupaten Boyolali" 2016.
- [5] Nurajjah, Dwiza Riana "Algoritma Naive Bayes, Decision Tree, Dan Svm Untuk Klasifikasi Persetujuan Pembiayaan Nasabah Koperasi Syariah." 2019.
- [6] Fatayat. Dan Risanto. Joko. "Proses *Data Mining* Dalam Meningkatkan Sistem Pembelajaran Pada Pendidikan Sekolah Menengah Pertama." 2013 .
- [7] Qonita Tanjung "Sistem Pendukung Keputusan Kelayakan Kredit Motor Menggunakan Metode Naive Bayes Pada Nsc Finance Cikampek." 2018.
- [8] Ginting, Selvia Lorena Br., Zarman,Wendi, Dan Hamidah, Ida. "Analisis Dan Penerapan Algoritma C4.5 Dalam *Data Mining* Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik." Prosiding Seminar Nasional Aplikasi Sains & Teknologi (Snasti). 1979911x. 2014
- [9] Emerensye S. Y. Pandie "Implementasi Algoritma *Data Mining* Naive Bayes Pada Koperasi." 2018.
- [10] Puji Santoso, Rudy Setiawan "Penerapan Metode Klasifikasi Decision Tree Dan Algoritma C4.5 Dalam Memprediksi Kriteria Nasabah Kredit Mega Auto Finance." 2020.
- [11] Noor Rahmat Effendy, Setyo Wahyu Saputro "Kajian Penerapan Metode Klasifikasi *Data Mining* Algoritma C4.5 Untuk Prediksi Kelayakan Kredit Pada Bank Mayapada Jakarta." 2018.
- [12] Siti Masripah "Evaluasi Penentuan Kelayakan Pemberian Kredit Koperasi Syariah Menggunakan Algoritma Klasifikasi C4.5." 2015.
- [13] Siti Lestari1, Akmaludin2, Mohammad Badrul "Implementasi Klasifikasi Naive Bayes Untuk Prediksi Kelayakan Pemberian Pinjaman Pada Koperasi Anugerah Bintang Cemerlang." 2020
- [14] Sumarni Adi "Sistem Informasi Untuk Prediksi Keamanan Pembiayaan Nasabah Bank Syariah Xyz." 2017.