

# Penerapan Data Mining untuk Menentukan Besar Pinjaman pada Koperasi Simpan Pinjam dengan Algoritma C4.5

## (Studi Kasus: Koperasi Wanita XYZ)

Indi Nadya Syahfitri<sup>1</sup>

**Abstrak**— Koperasi Wanita XYZ merupakan salah satu koperasi yang bergerak di bidang simpan pinjam dan menawarkan permohonan kredit. Selama ini untuk menentukan pemberian besar pinjaman kepada anggota dilakukan secara manual oleh pengurus Koperasi, dengan berbagai macam pertimbangan. Koperasi Wanita XYZ menerapkan sistem tanggung renteng yaitu tanggung jawab bersama dalam kelompok atas segala kewajiban terhadap koperasi. Diangkat dari masalah tersebut penulis menerapkan teknik data mining yang dapat membantu pengambilan keputusan besar pinjaman bagi anggota. Algoritma data mining yang digunakan dalam penelitian ini adalah algoritma C4.5. Atribut yang dipakai dalam penelitian ini adalah jumlah anggota kelompok, usia, lama keanggotaan, status, pekerjaan, jumlah permintaan, keperluan, lama angsuran, pinjaman sebelumnya, pendapatan, dan jumlah realisasi. Proses Pengolahan data dilakukan dengan menggunakan software WEKA. Tujuan penelitian adalah melakukan penerapan data mining untuk menentukan besar pinjaman yang akan diberikan oleh Koperasi Wanita XYZ kepada anggota menggunakan algoritma C4.5 yang kemudian akan menghasilkan sebuah pohon keputusan. Hasil penelitian dari permasalahan yang dikembangkan, dapat disimpulkan bahwa pengimplementasian data mining terhadap data simpan pinjam koperasi memiliki tingkat akurasi yang baik yaitu sebesar 94,67%.

**Kata Kunci:** Data Mining, WEKA, Klasifikasi, Pohon Keputusan, C4.5, Koperasi Simpan Pinjam

**Abstract**— Koperasi Wanita XYZ is one of the cooperatives engaged in savings, loans and offers credit applications. During this time to determine the large lending to members is done manually by the board of Koperasi, with a variety of considerations. Koperasi Wanita XYZ implements a joint responsibility system that is joint responsibility in the group for all obligations on cooperative. Taken from the problem, the authors apply data mining techniques that can help decision making large loans for members. The data mining algorithm that used in this research is C4.5 algorithm. The used attributes on this study are number of group members, age, membership duration, status, occupation, number of requests, needs, installments, previous loans, income, and amount of

realization. Processing data is done by using WEKA software. The purpose of the research is to apply the data mining to determine the amount of loan that will be given by Koperasi Wanita XYZ to its member using C4.5 algorithm which then produces a decision tree. A Result from the research problems that has been developed, can be concluded that the implementation of data mining of cooperative savings and loan data has a good accuracy of 94.67%.

**Keywords:** Data mining, WEKA, Classification, Decision Tree, C4.5, koperasi, Credit Union

### I. PENDAHULUAN

Melimpahnya data menjadi tantangan dalam dunia teknologi informasi untuk tidak hanya mendapatkan informasi saja, tetapi juga menemukan pola untuk menghasilkan pengetahuan. Proses menemukan pola dan informasi dari sebuah data merupakan tujuan dari data mining [1].

Data mining sebagai suatu proses eksplorasi dan analisis secara otomatis maupun semiotomatis terhadap data dalam jumlah besar dengan tujuan menemukan pola atau aturan yang berarti[2]. Saat ini data mining sudah banyak diimplementasikan di berbagai bidang, diantaranya dalam bidang bisnis, perbankan, dan koperasi simpan pinjam.

Koperasi yang bergerak dalam bidang usaha simpan pinjam merupakan industri yang penuh dengan resiko. Koperasi perlu melakukan manajemen resiko dengan mengidentifikasi, mengukur, dan mengendalikan besar pinjaman yang akan diberikan kepada anggota sehingga tidak timbul resiko kredit yang tidak diinginkan. Proses mengidentifikasi dan mengukur potensi resiko kredit dapat dibantu oleh teknik data mining, sehingga koperasi dapat menentukan besarnya pinjaman yang akan diberikan kepada anggota koperasi dengan tepat dan dapat mengurangi resiko kredit [3].

Pada penelitian ini dilakukan studi kasus pada Koperasi Wanita XYZ yang bergerak di bidang simpan pinjam dan menawarkan permohonan kredit. Selama ini untuk menentukan pemberian besar pinjaman kepada anggota dilakukan secara manual oleh pengurus Koperasi Wanita XYZ. Di Koperasi Wanita XYZ, pengajuan pinjaman yang dilakukan oleh anggota belum tentu sepenuhnya disetujui oleh pengurus koperasi. Besarnya pinjaman yang diberikan oleh pengurus koperasi kepada anggota didasarkan pada berbagai pertimbangan, seperti penghasilan anggota, lama

<sup>1</sup> Mahasiswa, Program Studi Sistem Informasi Fakultas Industri Kreatif Universitas Ciputra, UC Town Citraland, Surabaya 60219 INDONESIA (telp: 085655105516; e-mail: inadya@student.ciputra.ac.id)

menjadi anggota, pekerjaan anggota, pinjaman sebelumnya, dan faktor lainnya.

Koperasi Wanita XYZ menjalankan sistem tanggung renteng. Sistem Tanggung renteng adalah tanggung jawab bersama di satu kelompok atas segala kewajiban terhadap Koperasi Wanita XYZ dengan dasar keterbukaan dan musyawarah kelompok. Dari hal tersebut penentuan pemberian besar pinjaman oleh pengurus koperasi juga didasarkan pada jumlah anggota kelompok. Semakin banyak jumlah kelompok, semakin besar kemungkinan pengurus koperasi memberikan besar pinjaman sesuai dengan yang diajukan oleh anggota, karena semakin banyak jumlah anggota semakin kecil resiko yang muncul.

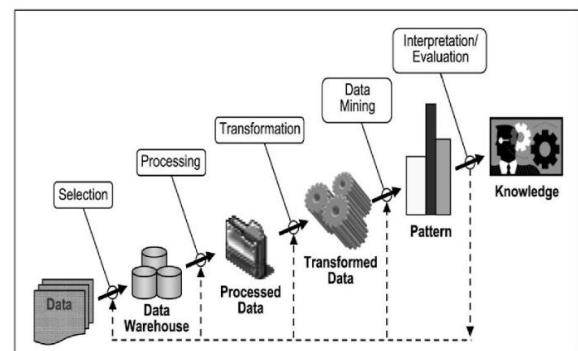
Mengacu dari masalah tersebut maka penulis tertarik melakukan penelitian untuk menerapkan data mining untuk membantu koperasi menentukan besar pinjaman kepada anggota sehingga nantinya tidak terjadi resiko kredit macet yang tidak diinginkan. Pada penelitian penerapan data mining ini, data yang akan digunakan merupakan data simpan pinjam Koperasi Wanita Potre Koneng. Tujuan pada penelitian ini adalah melakukan penerapan data mining untuk menentukan besar pinjaman yang akan diberikan oleh Koperasi Wanita XYZ kepada anggota menggunakan algoritma C4.5 yang kemudian akan menghasilkan sebuah pohon keputusan sehingga dapat membantu Koperasi Wanita XYZ dalam menentukan pemberian besar pinjaman.

## II. LANDASAN TEORI

### A. Data Mining

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visulisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar[4].

Istilah data mining dan Knowledge Discovery in Database (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan dengan satu sama yang lain. Salah satu tahapan dalam keseluruhan proses KDD adalah data mining. Proses *Knowledge Discovery in Database* dapat dilihat pada Gambar 1.



Gambar 1. Data Mining sebagai Salah Satu Tahap KDD

Alur kerja dari Data Mining berisi beberapa tahap, yaitu [5] :

#### 1. Data Collection

Pada tahap pengumpulan, data sering disimpan di database atau data warehouse untuk selanjutnya di proses.

#### 2. Feature extraction and data cleaning

Ketika data sudah terkumpul, banyak bentuk data yang tidak cocok digunakan untuk diolah. Untuk membuat data sesuai dengan format agar data dapat diolah, penting untuk mengubah data menjadi format yang sesuai dengan algoritma data mining. Hasil akhir dari prosedur ini adalah sebuah kumpulan data set yang terstruktur, yang dapat secara efektif digunakan oleh program komputer.

#### 3. Analytical processing and algorithms

Bagian akhir dari proses Data Mining adalah untuk merancang metode analisis yang efektif dari data yang diolah.

#### B. Preprocessing

Preprocessing adalah suatu proses atau langkah yang dilakukan untuk membuat data mentah menjadi data yang berkualitas serta menjadi input yang baik bagi data mining tools. Tahap *preprocessing* terbagi ke dalam beberapa tahap, yaitu :

##### 1. Data Cleaning

*Data Cleaning* berfungsi untuk membersihkan data dengan mengisi nilai – nilai yang hilang, *smoothing noise* data, mengidentifikasi dan menghapus *outlier*, serta menyelesaikan inkonsistensi [6].

##### 2. Data Integration

*Data Integration* merupakan penggabungan data dari berbagai Database ke dalam satu Database baru. Data Integration dapat membantu mengurangi dan menghindari redundansi data dan inkonsistensi dalam kumpulan data yang dihasilkan. Ini dapat membantu meningkatkan akurasi dan kecepatan proses data mining berikutnya.

##### 3. Data Reduction

*Data Reduction* dapat diterapkan untuk mengurangi ukuran data dengan menggabungkan atau menghilangkan data yang tidak dibutuhkan [6].

##### 4. Data Transformation

*Data Transformation* berfungsi mengubah data ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan.

### C. Pohon Keputusan

Pohon keputusan merupakan metode klasifikasi dan prediksi yang mengubah data yang sangat besar menjadi pohon keputusan yang menghasilkan sebuah aturan. Pohon keputusan juga dapat diterapkan dalam bentuk bahasa basis data untuk mencari *record* pada kategori tertentu [7].

Pohon keputusan menggunakan struktur seperti sebuah pohon (Tree) dimana setiap simpul (Node) mempresentasikan atribut, cabang mempresentasikan nilai dari atribut, dan daun mempresentasikan kelas [8]. Terdapat beberapa algoritma yang dapat digunakan untuk membangun pohon keputusan seperti ID3, C4.5, dan CART.

### D. Algoritma C4.5

Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal : bisa mengatasi missing data, bisa mengatasi data kontinyu, pruning, dan menangani atribut numerik [9].

Algoritma C4.5 adalah pengembangan dari algoritma ID3 yang mempunyai prinsip dasar kerja yang sama. Hanya saja dalam algoritma C4.5 pemilihan atribut dilakukan dengan menggunakan Gain Ratio sedangkan untuk algoritma ID3 hanya menggunakan nilai gain tertinggi untuk menentukan simpul (Node) pohon keputusan. Atribut dengan nilai Gain Ratio tertinggi dipilih sebagai atribut test untuk simpul [10]. Dibawah ini merupakan rumus GainRatio :

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)}$$

Gambar 2. Rumus GainRatio

$S$  = Dataset  
 $A$  = Atribut  
 $Gain(S, A)$  = Information Gain pada atribut A  
 $SplitInfo(S, A)$  = Split Information pada atribut A

Sebelum menghitung nilai GainRatio, terlebih dahulu dihitung nilai gain setiap atribut.

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

Gambar 3. Rumus Gain

$S$  = Dataset  
 $A$  = Atribut  
 $n$  = Jumlah partisi atribut A  
 $|S_i|$  = Proporsi  $S_i$  terhadap  $S$   
 $|S|$  = Jumlah kasus dalam  $S$

Sebelum menghitung nilai gain dari setiap atribut, terlebih dahulu menghitung nilai entropy untuk setiap atribut. Dibawah ini rumus Entropy :

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

Gambar 4. Rumus Entropy

$S$  = Dataset  
 $n$  = Jumlah partisi  $S$   
 $p_i$  = Proporsi  $S_i$  terhadap  $S$

Setelah mendapatkan nilai gain , selanjutnya menghitung nilai SplitInfo dari setiap atribut. SplitInfo menyatakan entropi atau informasi potensial dari atribut. Dibawah ini merupakan rumus SplitInfo :

$$SplitInfo(S, A) = -\sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

Gambar 5. Rumus SplitInfo

$S$  = Dataset  
 $A$  = Atribut  
 $S_i$  = Jumlah Sampel untuk Atribut  $i$

Atribut dengan nilai Gain Ratio tertinggi dipilih sebagai atribut test untuk simpul (Node) yang digunakan untuk membangun pohon keputusan. Berdasarkan rumus – rumus yang telah dijelaskan diatas, data yang diperoleh dapat dimasukkan dan diproses dengan algoritma C4.5 untuk proses pembuatan pohon keputusan [11].

### E. WEKA

WEKA (*Weka Environment for Knowledge Analysis*) adalah aplikasi *Data Mining Open Source* berbasis Java. J48 adalah salah satu kelas di paket classifiers pada sistem Weka yang mengimplementasikan C4.5. Paket classifiers berisi implementasi dari hampir semua algoritma untuk klasifikasi dan prediksi. Kelas yang paling penting di paket ini adalah Classifier, yang mendeklarasikan struktur umum dari skema klasifikasi dan prediksi.

Semua kelas yang mengimplementasikan algoritma klasifikasi menginduk ke kelas Classifier, termasuk kelas J48. J48 yang menangani himpunan data dalam format ARFF, tidak mengandung kode untuk mengkonstruksi pohon keputusan. Kelas ini mereferensi kelas kelas lain, kebanyakan di paket weka.classifiers.j48, yang mengerjakan semua proses konstruksi pohon [12].

### F. Percentage Split

Percentage split adalah salah satu test option yang disediakan oleh WEKA. Hasil klasifikasi percentage split akan dites dengan menggunakan K% dari data set sebagai data training [13]. Dalam percentage split, database dibagi secara acak menjadi dua kumpulan data yang terpisah. Set pertama, dimana sistem data mining mencoba mengekstrak

pengetahuan dari data training. Pengetahuan yang diekstrak dapat diuji terhadap set kedua yang disebut uji coba, biasanya sistem mebagi secara acak kumpulan data menjadi 2 bagian. Secara umum pembagian dilakukan dengan persentase 66% objek dari dataset asli sebagai data training dan sisa dataset sebagai data testing [14].

#### G. Koperasi Simpan Pinjam

Koperasi adalah perkumpulan otonom dari orang – orang yang bersatu secara sukarela untuk memenuhi kebutuhan – kebutuhan dan aspirasi ekonomi, sosial dan budaya bersama – sama melalui perusahaan koperasi yang dimiliki bersama dan dikendalikan secara demokratis [15].

Koperasi simpan pinjam didirikan untuk memberi kesempatan kepada anggotanya memperoleh pinjaman dengan mudah dan bunga ringan. Koperasi simpan pinjam menghimpun dana dari para anggotanya yang kemudian menyalurkan kembali dana tersebut kepada para anggotanya.

#### H. Sistem Tanggung Renteng

Sistem tanggung renteng adalah tanggung jawab bersama di satu kelompok atas segala kewajiban terhadap koperasi dengan dasar keterbukaan dan musyawarah kelompok. Kelebihan dari sistem ini adalah semua ikut berpikir, bekerja, dan memantau. Tanggung jawab dibagi secara merata sehingga semua berhak mendapatkan akses informasi atas perkembangan usaha dan hasil dari usaha yang dimaksud.

Untuk itu syarat utama dalam sistem tanggung renteng adalah anggota harus berkelompok untuk memahami hak dan kewajiban masing – masing anggota. setiap anggota minimal terdiri dari 15 anggota dan maksima 30 anggota. Semua proses pengambilan keputusan harus melalui musyawarah karena apapun yang diputuskan akan menjadi tanggung jawab seluruh anggota kelompok. Sistem tanggung renteng akan sangat membantu koperasi yang mempunyai unit usaha simpan pinjam untuk terus sifat mempertahankan dan mengembangkan asetnya

### III. ANALISIS DAN DESAIN

#### A. Analisa Sistem

Berikut ini adalah sistem simpan pinjam yang berlaku di Koperasi Wanita XYZ :

1. Anggota dapat melakukan simpan pinjam setelah tiga bulan menjadi anggota koperasi dengan maksimal peminjaman adalah Rp. 2,000,000.00.
2. Anggota melalui kelompok mengajukan pinjaman dengan cara mengisi Surat Pengajuan Pinjaman (SPP) sesuai dengan kemampuan masing-masing peminjam.
3. Surat Pengajuan Pinjaman ditandatangani oleh suami serta 50% + 1 dari jumlah anggota kelompok. Surat Pengajuan Pinjaman yang telah disetujui oleh kelompok diserahkan oleh penanggung jawab kelompok kepada pengurus koperasi.

4. Jika status anggota belum menikah, maka SPP ditandatangani oleh orang tua anggota. Anggota baru yang belum menikah harus mempunyai pekerjaan dan penghasilan setiap bulan.
5. Surat Pengajuan Pinjaman diteliti oleh pengurus koperasi, jika SPP tidak memenuhi syarat maka akan dikembalikan kepada anggota dan jika layak mendapatkan pinjaman maka realisasi pinjaman akan diproses.
6. Pengurus koperasi membuat Surat Pengakuan Hutang (SPH) sesuai dengan SPP yang disetujui jumlah pinjamannya oleh pengurus koperasi.
7. Pinjaman akan diberikan kepada anggota setelah angsuran dari pinjaman sebelumnya telah dilunasi.

#### B. Hasil Wawancara

Berdasarkan wawancara dengan ketua Koperasi Wanita Potre Koneng didapatkan kesimpulan seperti berikut :

1. Pencatatan transaksi simpan pinjam di Koperasi Wanita Potre Koneng masih menggunakan sistem pencatatan secara manual.
2. Keputusan pemberian pinjaman dilakukan oleh pengurus koperasi berdasarkan beberapa pertimbangan, diantaranya : referensi dari kelompok, referensi dari petugas lapangan Koperasi Wanita Potre Koneng, penghasilan perbulan anggota, keperluan pinjaman, pekerjaan dari anggota.
3. Penurunan jumlah realisasi pinjaman terjadi apabila kelompok terlambat membayar angsuran pinjaman kepada pihak koperasi. Minimal penurunan pinjaman adalah senilai Rp. 1,000,000.00 dan maksimal penurunan pinjaman adalah Rp. 4,000,000.00. Jika kelompok selalu terlambat membayar angsuran 4 kali berturut-turut, maka pihak koperasi tidak memberikan realisasi pinjaman kepada anggota.
4. Setiap anggota yang melakukan transaksi simpan pinjam harus mencapai jumlah plafond pinjaman yang ditetapkan dengan rumus enam kali simpanan pokok ditambah simpanan wajib. Pinjaman anggota yang belum mencapai plafond dikenakan wajib simpan.

#### C. Analisa Data

Dalam penelitian ini, dicari atribut-atribut dari data yang dikumpulkan untuk menentukan besar pinjaman pada Koperasi Wanita XYZ. Data yang digunakan dalam penelitian ini merupakan data simpan pinjam Koperasi Wanita XYZ selama 10 bulan, yaitu dari bulan Januari 2016 sampai dengan Oktober 2016. Class dalam penelitian ini adalah jumlah realisasi. Sumber data simpan pinjam diperoleh dari buku simpan pinjam yang disimpan oleh pengurus koperasi. Jumlah data yang diperoleh adalah 1036 data dengan 14 kolom atribut data. Adapun atribut dari data dijelaskan dalam Tabel 1.

TABEL 1  
ATRIBUT DATA SIMPAN PINJAM

Atribut	Keterangan
No Kelompok	No kelompok anggota koperasi
Jumlah Anggota Kelompok	Jumlah anggota pada kelompok
Tanggal Lahir	Tanggal lahir anggota koperasi
Tanggal Masuk	Tanggal masuk anggota menjadi anggota koperasi
Alamat	Alamat anggota koperasi
Kecamatan	Kecamatan dari alamat anggota koperasi
Status	Status pernikahan dari anggota koperasi
Pekerjaan	Pekerjaan anggota koperasi
Jumlah Permintaan Pinjaman	Jumlah permintaan pinjaman yang diajukan oleh anggota koperasi
Keperluan	Keperluan pinjaman
Lama Angsuran	Lama angsuran pinjaman
Pinjaman Sebelumnya	Data pinjaman sebelum transaksi pinjaman saat ini
Pendapatan	Pendapatan per bulan dari anggota koperasi
Jumlah Realisasi	Jumlah realisasi pinjaman yang telah ditentukan oleh pengurus koperasi

Berikut ini adalah tipe data atribut yang dipakai :

TABEL 2  
TIPE DATA ATRIBUT

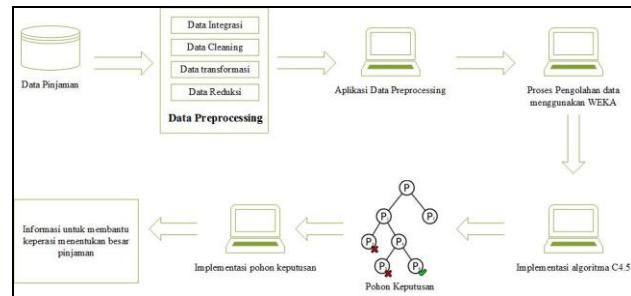
Atribut	Tipe Data
No Kelompok	Nominal
Jumlah Anggota Kelompok	Ordinal
Tanggal Lahir	Numerik
Tanggal Masuk	Numerik
Alamat	Nominal
Kecamatan	Nominal
Status	Nominal
Pekerjaan	Nominal
Jumlah Permintaan Pinjaman	Numerik
Keperluan	Nominal
Lama Angsuran	Ordinal
Pinjaman Sebelumnya	Numerik
Pendapatan	Ordinal
Jumlah Realisasi	Nominal

Class atribut yang dipakai dalam penelitian ini adalah jumlah realisasi. Tabel 3 dibawah ini merupakan variabel class jumlah realisasi.

TABEL 3  
VARIABEL CLASS JUMLAH REALISASI

No	Variabel
1	Terima
2	Kurangi 1
3	Kurangi 2
4	Kurangi 3
5	Kurangi 4
6	Tolak

#### D. Desain Arsitektur



Gambar 6. Desain Arsitektur

##### 1. Pengumpulan Data

Tahap pertama dalam penelitian ini adalah mengumpulkan data simpan pinjam

##### 2. Data Preprocessing

Pada tahap ini dilakukan proses preprocessing dimulai dari data cleansing, data integration, data reduction, dan data transformation. Format output dari tahap preprocessing adalah .csv (comma separated value).

##### 3. Pengolahan data Menggunakan WEKA

Pada tahap ini, output yang dihasilkan dari tahap preprocessing diolah menggunakan aplikasi WEKA.

##### 4. Implementasi C4.5

Implementasi algoritma C4.5 dilakukan menggunakan aplikasi WEKA dengan nama kelas j48 untuk menghasilkan pohon keputusan.

##### 5. Implementasi Pohon Keputusan

Pada tahap ini hasil pohon keputusan diimplementasikan ke dalam aplikasi penentuan besar pinjaman berbasis web.

#### E. Desain Aplikasi Preprocessing

Aplikasi preprocessing merupakan aplikasi yang dibuat untuk mengubah data dari format excel menjadi format yang dapat diterima oleh WEKA. Cara kerja dari aplikasi preprocessing adalah sebagai berikut :

1. Memilih file excel untuk dirubah menjadi format yang dapat digunakan saat mengolah data menggunakan WEKA.
2. Aplikasi akan melakukan pengecekan jumlah sheets, kolom, dan row yang ada dalam file excel tersebut.
3. Setelah melakukan pengecekan, aplikasi melakukan insert data ke dalam satu tabel database dan ditampilkan dalam bentuk datagridview.
4. Ketika aplikasi menjalankan proses preprocessing, maka aplikasi melakukan proses data cleaning dan data transformation.
5. Setelah itu aplikasi melakukan export file menjadi format .csv sehingga dapat diterima oleh WEKA.

#### IV. IMPLEMENTASI

##### A. Aplikasi Preprocessing

Pada bagian Aplikasi preprocessing berisi tahap preprocessing pada data simpan pinjam.

###### 1. Data Integration

*Data Integration* melakukan input data simpan pinjam dari file excel ke dalam database. Proses Pertama adalah memilih data simpan pinjam dalam file excel yang akan diproses. Lalu, dilakukan proses *data Integration*.

###### 2. Data Cleaning

*Data Cleaning* adalah proses untuk untuk menghapus missing values pada data simpan pinjam. Jika, ada salah satu data yang kosong pada salah satu *row* pada file, maka *row* tersebut dihapus.

###### 3. Data Transformation

*Data Transformation* merupakan proses pengubahan data ke dalam format yang sesuai untuk diproses dalam data mining. Pada penelitian ini terdapat tujuh pengubahan data yang dilakukan. Sebanyak tujuh atribut yang mengalami pengubahan yaitu :

###### a. Jumlah Anggota Kelompok

Pada *data transformation* data jumlah anggota kelompok dikelompokkan menjadi tiga kategori, yaitu kategori 1, 2, dan 3. Tipe data jumlah anggota kelompok diubah dari tipe data numerik menjadi tipe data ordinal.

###### b. Tanggal Lahir

Pada proses *data transformation*, data tanggal lahir diubah menjadi data usia.

###### c. Tanggal Masuk

Pada proses *data transformation*, data tanggal masuk diubah menjadi data lama keanggotaan.

###### d. Status

Pada *data transformation*, data status dikelompokkan menjadi dua kategori, yaitu menikah dan belum menikah.

###### e. Pekerjaan

Pada *data transformation*, data pekerjaan dikelompokkan menjadi lima kategori, yaitu ibu rumah tangga, PNS, PS, wirausaha, dan guru.

###### f. Keperluan

Pada *data transformation*, data keperluan dikelompokkan menjadi dua kategori, yaitu usaha dan rumah tangga.

###### g. Pendapatan

Pada *data transformation*, data pendapatan dikelompokkan menjadi 5 kategori, yaitu kategori 1, 2, 3, 4, dan 5. Tipe data jumlah anggota kelompok diubah dari tipe data numerik menjadi tipe data ordinal.

###### 4. Export File CSV

Setelah proses preprocessing selesai dilakukan maka

data hasil preprocessing di export ke dalam format data yang dapat diterima oleh WEKA. Salah satu format yang dapat dipakai adalah format comma separated value(CSV).

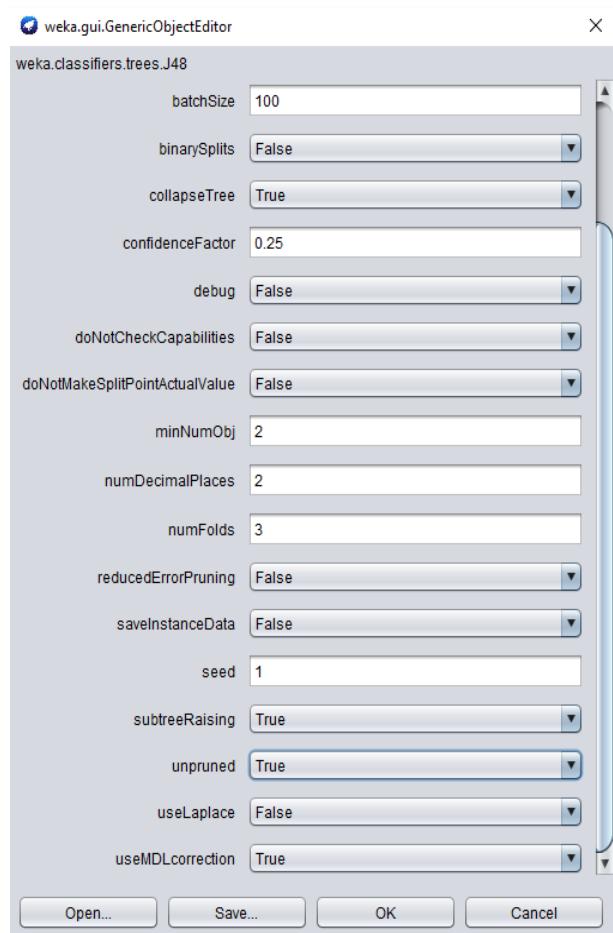
##### B. Proses mengolah data menggunakan WEKA

Sebelum melakukan proses pengolahan data menggunakan WEKA, dilakukan data reduction yaitu menghapus atribut no kelompok dan kecamatan yang dihapus pada panel preprocess pada aplikasi WEKA. Atribut no kelompok dihapus dikarenakan atribut no kelompok mempunyai korelasi yang tinggi dengan class jumlah realisasi. Sehingga jika atribut no kelompok digunakan dalam proses pembentukan pohon keputusan maka atribut – atribut lainnya tidak memiliki kontribusi dalam pembentukan pohon keputusan.

Atribut kecamatan dihapus karena atribut kecamatan dipakai hanya untuk disimpan menjadi catatan di buku anggota Koperasi Wanita XYZ, selain itu jika atribut kecamatan dihapus akurasi hasil pohon keputusan dari WEKA mempunyai tingkat akurasi lebih tinggi daripada atribut kecamatan digunakan dalam pengolah data.

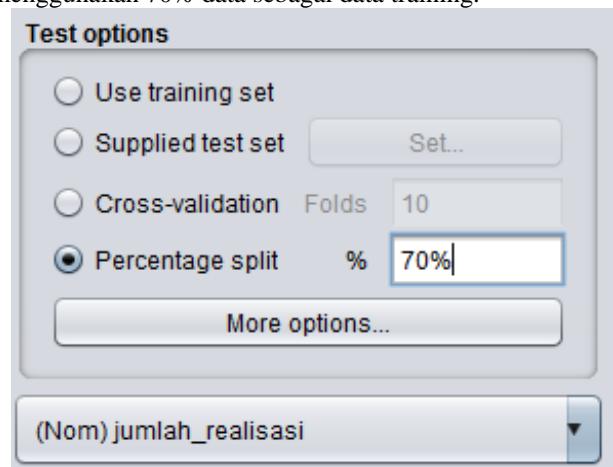
Setelah melakukan proses preprocessing yang terdiri dari tahap *data integration*, *data cleaning*, *data transformation*, dan *data reduction*. maka langkah selanjutnya adalah melakukan proses mining menggunakan aplikasi WEKA. Sebelum melakukan implementasi terhadap algoritma C4.5 dalam WEKA, terlebih dahulu penulis melakukan data reduction yang dilakukan di WEKA.

Setelah proses data reduction dilakukan selanjutnya proses berpindah ke panel Classify. Sebelum proses dimulai terlebih dahulu menentukan algoritma yang dipakai. Dalam panel classify terdapat beberapa algoritma yang dapat dipilih dan memiliki fungsi serta cara kerja yang berbeda. Pada penelitian ini penulis memilih untuk menggunakan algoritma C4.5. algoritma C4.5 menghasilkan pohon keputusan. Dalam WEKA algoritma C4.5 mempunyai nama J48. Sebelum proses dimulai terdapat parameter yang harus diatur, yaitu atribut unpruned. Parameter unpruned diubah menjadi dari false menjadi true.



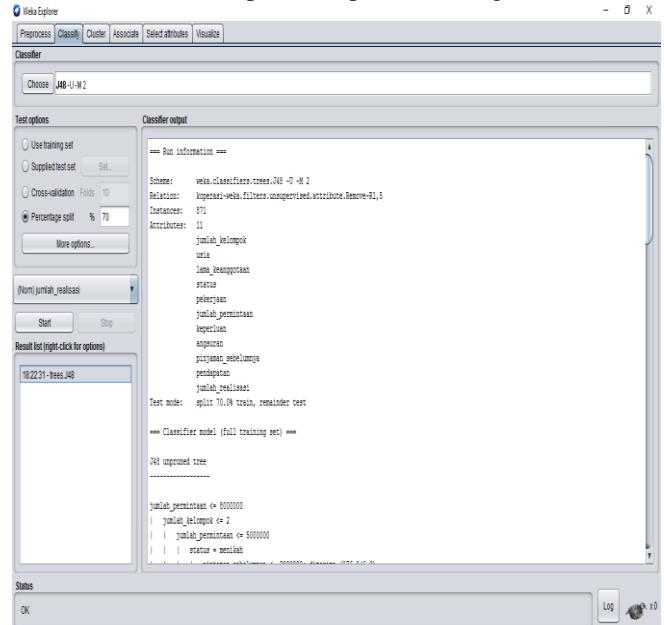
Gambar 7. Parameter Algoritma C4.5

Setelah parameter algoritma C4.5 diatur maka langkah selanjutnya adalah menentukan pilihan test yang akan digunakan. Terdapat beberapa teknik test yang dapat dipilih, pada penelitian ini penulis menggunakan pilihan teknik percentage split yang hasil klasifikasinya akan dites dengan menggunakan K% dari data tersebut sebagai data training. K merupakan input dari user, pada penelitian ini penulis menggunakan 70% data sebagai data training.



Gambar 8. Pemilihan Teknik Percentage Split

Setelah input configurasi diatas telah dilakukan, selanjutnya proses pencarian dengan menggunakan WEKA. Seluruh proses pencarian akan dilakukan oleh aplikasi WEKA dengan menggunakan Algoritma C4.5 secara otomatis. Gambar 7 dibawah ini menunjukkan tampilan weka saat melakukan proses implementasi algoritma C4.5.



Gambar 9. Implementasi Algoritma C4.5

### C. Pohon Keputusan

Setelah mengolah data tersebut, WEKA akan memberikan output berupa pohon keputusan. Pohon keputusab dapat dilihat pada Segmen 1 dibawah ini.

```
jumlah_permintaan <= 8000000
| jumlah_kelompok <= 2
| | jumlah_permintaan <= 5000000
| | | status = menikah
| | | | pinjaman_sebelumnya
<= 2000000: diterima (176.0/6.0)
| | | | | pinjaman_sebelumnya >
2000000
| | | | |
pinjaman_sebelumnya <= 8000000:
diterima (16.0/2.0)
| | | | |
pinjaman_sebelumnya > 8000000:
ditolak (2.0)
| | | status = belum menikah
| | | | pendapatan <= 1:
ditolak (2.0)
| | | | pendapatan > 1:
diterima (6.0)
| | jumlah_permintaan > 5000000
| | | jumlah_kelompok <= 1
| | | | keperluan = usaha:
diterima (9.0)
| | | | keperluan = rumah
tanqqa
```

```

| | | | |
pinjaman_sebelumnya <= 0
| | | | | usia <= 47:
dikurangi 3 (3.0)
| | | | | usia > 47:
diterima (3.0)
| | | | |
pinjaman_sebelumnya > 0: diterima
(18.0/1.0)
| | | jumlah_kelompok > 1
| | | | pekerjaan = Ibu rumah
tangga: diterima (25.0/1.0)
| | | | pekerjaan = Guru:
diterima (5.0)
| | | | pekerjaan = PNS:
diterima (8.0)
| | | | pekerjaan = Wirausaha
| | | | keperluan =
usaha: diterima (5.0)
| | | | keperluan = rumah
tangga
| | | | | usia <= 44:
ditolak (2.0)
| | | | | usia > 44:
diterima (2.0)
| | | | pekerjaan = PS:
diterima (6.0)
| jumlah_kelompok > 2
| | jumlah_permintaan <= 6000000:
diterima (22.0)
| | jumlah_permintaan > 6000000
| | | lama_keanggotaan <= 5:
diterima (20.0/3.0)
| | | lama_keanggotaan > 5
| | | | lama_keanggotaan <=
8: ditolak (3.0/1.0)
| | | | lama_keanggotaan > 8:
diterima (3.0/1.0)
jumlah_permintaan > 8000000
| pendapatan <= 1
| | jumlah_kelompok <= 1
| | | pekerjaan = Ibu rumah
tangga
| | | | keperluan = usaha:
diterima (19.0/2.0)
| | | | keperluan = rumah
tangga
| | | | | usia <= 38:
dikurangi 2 (2.0/1.0)
| | | | | usia > 38:
diterima (15.0/1.0)
| | | | pekerjaan = Guru:
diterima (0.0)
| | | | pekerjaan = PNS: diterima
(2.0)
| | | | pekerjaan = Wirausaha:
diterima (16.0)
| | | | pekerjaan = PS: diterima
(2.0)

```

```

| | | jumlah_kelompok > 1
| | | | pekerjaan = Ibu rumah
tangga: diterima (64.0/8.0)
| | | | pekerjaan = Guru:
diterima (6.0)
| | | | pekerjaan = PNS: diterima
(5.0/1.0)
| | | | pekerjaan = Wirausaha
| | | | | jumlah_permintaan <=
10000000: diterima (6.0/1.0)
| | | | | jumlah_permintaan >
10000000: ditolak (3.0/1.0)
| | | | pekerjaan = PS: diterima
(0.0)
| | | | pendapatan > 1: diterima
(393.0/17.0)

```

#### Segmen 1. Pohon Keputusan

#### D. Website Penentuan Besar Pinjaman

Website penentuan besar pinjaman merupakan website hasil implementasi dari pohon keputusan yang terbentuk. Terdapat 4 halaman website, yaitu :

##### 1. Halaman Login

Pada halaman login user melakukan login untuk dapat masuk ke dalam halaman permintaan pinjaman.

##### 2. Halaman Permintaan Pinjaman

Pada halaman ini user dapat menyimpan data transaksi pinjaman ke dalam database dan dapat melihat hasil rekomendasi dari pohon keputusan yang telah dihasilkan.

##### 3. Halaman Pohon Keputusan

Halaman pohon keputusan pada website penentuan besar pinjaman merupakan implementasi dari pohon keputusan yang telah dihasilkan dari algoritma C4.5. Pada halaman ini user tidak dapat melakukan input data ke dalam database. Halaman pohon keputusan ini memberikan hasil rekomendasi berupa jumlah realisasi dengan keyakinan atas hasil rekomendasi tersebut.

## V. HASIL PENGUJIAN

#### A. Pengujian Akurasi Data

Pengujian akurasi data pada penelitian ini dilakukan untuk mengukur ketepatan hasil rekomendasi yang dihasilkan dari pohon keputusan, dengan hasil keputusan yang dibuat oleh pengurus koperasi. pengujian akurasi data dilakukan dengan menggunakan data transaksi simpan pinjam Koperasi Wanita XYZ yang terjadi pada bulan November 2016 hingga Desember 2017 dengan total data pengajuan simpan pinjam 225 data. Tabel 2 dibawah ini merupakan detail jumlah data pengajuan simpan pinjam yang digunakan untuk pengujian akurasi data.

TABEL 4  
DATA PENGUJIAN

Bulan	Jumlah Data
November 2016	109 Pengajuan
Desember 2016	116 Pengajuan

Perhitungan akurasi data pada penelitian ini menggunakan *confusion matrix* seperti tabel 5 dibawah ini.

TABEL 5  
CONFUSION MATRIX

Pred ik-si	Kenyataan						
	class	Teri ma	Ku rangi 1	Kur angi 2	Kura ngi 3	Kura ngi 4	Tol ak
Teri ma	208	0	3	2	0	5	
Kura ngi 1	0	0	0	0	0	0	
Kura ngi 2	0	0	1	0	0	1	
Kura ngi 3	0	0	0	0	0	0	
Kura ngi 4	0	0	0	0	0	0	
Tolak	0	1	0	0	0	4	

Berdasarkan confusion matrix pada Tabel 5.2 diatas dapat dilihat bahwa prediksi class diterima dan kenyataan class diterima memiliki total 208 data, prediksi class diterima dan kenyataan class dikurangi 2 memiliki total 3 data, prediksi class diterima dan kenyataan class dikurangi 3 memiliki total 2 data, prediksi class diterima dan kenyataan class ditolak memiliki total 5 data, prediksi class dikurangi 2 dan kenyataan class dikurangi 2 memiliki total 1 data, prediksi class dikurangi 2 dan kenyataan ditolak memiliki total 1 data, prediksi class ditolak dan kenyataan class dikurangi 1 memiliki total 1 data, dan prediksi class ditolak dan kenyataan class ditolak memiliki total 4 data. Berdasarkan hasil pengujian akurasi data yang telah dilakukan menggunakan data pada tabel 5.2, diperoleh nilai akurasi data yang cocok antara hasil rekomendasi dari pohon keputusan dengan hasil keputusan dari pengurusan koperasi adalah 94,67%.

## VI. KESIMPULAN DAN SARAN

Kesimpulan yang didapatkan pada penelitian ini adalah sebagai berikut :

1. Aplikasi preprocessing pada penelitian ini mampu melakukan data integration dengan melakukan insert data pada database, data cleaning dengan menghapus row data yang kosong, dan data transformation dengan mengubah tipe data sesuai dengan format.
2. Aplikasi preprocessing mampu melakukan export data dalam bentuk file .CSV sehingga dapat diproses di software WEKA.
3. Data reduction menghapus atribut no kelompok dan kecamatan.

4. Hasil pengujian akurasi data pada penelitian ini menggunakan confusion matrix dan memiliki akurasi 94.67%.
  5. Website mampu menampilkan hasil rekomendasi keputusan untuk pemberian besar pinjaman.
- Saran pada penelitian ini adalah sebagai berikut :
1. Pengembangan pada Aplikasi Preprocessing dan Webiste hasil implementasi, sehingga dapat diterapkan pada koperasi lainnya.
  2. Pengembangan untuk fitur simpanan pada website permintaan pinjaman.
  3. Data penelitian ditambah periodenya untuk hingga 2 periode sehingga dapat melihat siklus bisnis / pinjaman pada Koperasi wanita Potre Koneng.
  4. Pada penelitian selanjutnya dapat ditambahkan atribut pengeluran dan tanggungan dari anggot koperasi.
  5. Penelitian ini dapat dijadikan acuan dan refensi bagi para peneliti berikutnya yang akan melakukan proses data mining dengan menggunakan algoritma C4.5.

## DAFTAR PUSTAKA

- [1] A. Elcholqi, B. Noranita and I. Waspada, "Penentuan Besar Pinjaman di Koperasi Simpan Pinjam dengan Algoritma K-Nearest Neighbor (Studi Kasus di Koperasi Simpan Pinjam BMT Bina Insani Pringapus)," *Jurnal Masyarakat Informatika*, p. 15, 2012.
- [2] S. G. Linoff and M. J. Berry, *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*, 3rd Edition, USA: Wiley Publishing, Inc, 2004.
- [3] R. D. Jayanti and Noeryanti, "Aplikasi Metode K-Nearest Neighbor dan Analisis Diskriminasi untuk Analisis Resiko Kredit pada Koperasi Simpan Pinjam di Kopinkra Sumber Rejeki," in Prosiding Seminar Nasional Aplikasi Sains & Teknologi, Yogyakarta, 2014.
- [4] D. T. Larose, *Data Mining Methods and Models*, New Jersey: Jhon Wiley & Sons, Inc, 2005.
- [5] C. C. Aggarwal, *Data Mining*, New York: IBM T.J. Watson Research Center, 2015.
- [6] J. Han, M. Kamber and J. Pei, *DATA MINING Concepts and Techniques*, United States: Morgan Kaufmann Publishers, 2012.
- [7] Kusriini and E. T. Luthfi, *Algoritma Data Mining*, Yogyakarta: C.V ANDI OFFSET, 2009.
- [8] A. R. Wicaksono and Y. Rahayu, "Klasifikasi Data Mining Untuk Menentukan Potensi Kredit Macet pada Koperasi Simpan Pinjam Prinkoveri Waleri Menggunakan Algoritma Decision Tre C4.5," *1,2Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang*, p. 3, 2015.
- [9] I. H. Witten, E. Frank and M. A. Hall, *Data Mining : Practical Machine Learning Tools and Techniques*, USA: Morgan Kaufmann Publishers, 2011.
- [10] s. Defiyanti, "Perbandingan Kinerja Algoritma ID3 dan C4.5 dalam Klasifikasi Spam-Mail," 2016.
- [11] F. F. Harryanto and S. Hansun, "Penerepan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE," *Jatisi*, p. 95, 2017.
- [12] A. Raditya, "Implementasi Data Mining Classification untuk Mencari Pola Prediksi Hujan dengan Menggunakan Algoritma C4.5," p. 3, 2012.
- [13] R. R. Bouckaert, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald and D. Scuse, *WEKA Manual for Version 3-7-8*, New Zealand: General Public License, 2013.

- [14] T. C. Sharma and M. Jain, "WEKA Approach for Comparative Study of Classification Algorithm," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 4, p. 5, April 2013.
- [15] A. Supriyanto, TATA KELOLA KOPERASI KREDIT ATAU KOPERASI SIMPAN PINJAM, Yogyakarta: CV. ANDI OFFSET, 2015.