

## ALGORITMA C4.5 UNTUK KLASIFIKASI PEMILIHAN CALON KETUA RW

Erlin Elisa<sup>1</sup> Rika Harman<sup>2</sup>

Universitas Putera Batam<sup>12</sup>

Jl.R.Suprpto Muka Kuning Batam

E-mail : Erlin.Elisa@puterabatam.ac.id<sup>1</sup>,Rika.Harman@puterabatam.ac.id<sup>2</sup>

### ABSTRAK

Indonesia merupakan sebuah negara yang menjunjung tinggi demokrasi dalam pemerintahannya. Salah satu bentuk demokrasi di Indonesia di adakannya Pemilu (Pemilihan Umum) sekali dalam lima tahun. Pemilihan umum tidak hanya dilakukan pada pemerintahan pusat tetapi juga sampai kepada pemerintahan daerah bahkan pemerintahan ditempat tinggal kita. Keberadaan lembaga masyarakat bertujuan untuk mempercepat masyarakat yang tertib, aman dan sejahtera. Penelitian ini dilakukan untuk mengklasifikasikan kriteria pemilihan calon ketua RW (Rukun Warga) pada sebuah daerah dikota batam, permasalahan yang terjadi selama ini adalah proses pemilihan atau penetapan calon ketua RW belum memiliki kriteria yang harus dijadikan landasan dalam penetapan calon, begitu juga pada saat pemilihan berlangsung para pemilih tidak mengetahui dengan jelas siapa yang dipilih atau mungkin melakukan pemihan atas dasar kenal saja dengan calon, sehingga tujuan sebenarnya diadakan pemungutan suara untuk ketua RW tidak tercapai. Dalam pemilihan partisipasi warga sangat diperlukan akan tetapi tidak semua warga ikut berpartisipasi pada perhelatan ini, tentunya hal ini menimbulkan permasalahan setelah dilakukan pemungutan suara, banyak dari warga yang tidak kenal siapa RW yang terpilih sampai kepada tidak mengetahui latar belakang dari pemimpin diperumahannya, selain itu para calon ketua RW tidak melakukan sosialisasi secara menyeluruh tentang visi-misi kedepan dalam pembangunan masyarakat, situasi ini menimbulkan keluhan dari masyarakat seperti pada saat melakukan pemilihan mereka tidak memilih berdasarkan kriteria yang baik untuk dijadikan wakil pimpinan tempat mereka tinggal. Berdasarkan permasalahan diatas maka peneliti akan menggunakan sebuah teknik *datamining* dengan algoritma C4.5 yang bertujuan untuk mengklasifikasikan kriteria calon ketua RW (Rukun Warga) sehingga dapat membuat keputusan terhadap penetapan calon RW. Berdasarkan Hasil perhitungan C4.5 dengan pengaplikasian *decision tree* menggunakan aplikasi *weka* kriteria yang diharapkan oleh masyarakat untuk menjadi calon ketua RW yang akan dipilih dengan nilai *gain* tertinggi adalah memiliki visi dan misi sebagai calon, mempunyai tingkat solidaritas terhadap tertangga dan memiliki jiwa kepemimpinan yang baik.

Kata kunci : Pemilihan, *Datamining*, Algoritma C4.5 Pohon Keputusan.

### ABSTRACTS

Indonesia is a country that upholds democracy in its government. One form of democracy in Indonesia is to hold an election (General Election) once every five years. Elections are not only carried out at the central government but also to the regional government and even the government where we live. The existence of community institutions aims to accelerate an orderly, safe and prosperous community. This research was conducted to classify the criteria for the selection of candidates for RW (Rukun Warga) candidates in an area in the city of Batam, the problem that has occurred so far is the process of selecting or determining prospective RW heads. which must be used as a basis for determining candidates, as well as voters do not know clearly who is elected or may vote on the basis of just knowing the candidate, so that the actual purpose of voting for the RW chairman is not achieved. In the selection of citizen participation is very necessary but not all residents participate in this event, of course this raises problems after the vote, many of the residents who do not know who the RW is chosen to do not know the background of the leader in their housing, besides the candidates RW chairman did not conduct a comprehensive socialization of the future vision and mission in community development, this situation led to complaints from the community such as when making an election they did not vote based on good

criteria to be the deputy leader where they lived. Based on the above problems, the researcher will use a data lamination technique with C4.5 algorithm which aims to classify the criteria of prospective RW heads so that they can make decisions about the determination of RW candidates. Based on the results of the C4.5 calculation by applying the decision tree criteria expected by the community to become the candidate for RW leader who will be chosen with the highest gain value is to have a vision and mission as a candidate, have a level of solidarity with the proud and have a good leadership spirit.

**Keywords:** Selection, Datamining, C4.5 Decision Tree Algorithm.

## 1. PENDAHULUAN

Demokrasi memiliki pandangan sebagai sesuatu yang penting karena nilai yang terkandung didalamnya sangat diperlukan sebagai acuan dalam kehidupan berbangsa yang baik. Indonesia merupakan sebuah negara yang menjunjung tinggi demokrasi dalam pemerintahannya dimana setiap pembentukan pemerintahan baru dilakukan secara demokrasi melalui pemilihan umum atau yang sering disebut dengan istilah PEMILU. penelitian sebelumnya mengatakan bahwa pemilihan umum merupakan mekanisme utama yang terdapat dalam tahapan penyelenggaraan negara dan pembentukan pemerintahan. Pemilihan umum dipandang sebagai bentuk paling nyata dari kedaulatan yang berada di tangan rakyat dalam penyelenggaraan Negara [1].

Pemilihan umum tidak hanya dilakukan pada pemerintahan pusat seperti pemilihan calon legislatif, presiden kepada tingkat propinsi, daerah saja tetapi ketingkat pemerintahan atau organisasi paling bawah perlu dilakukan pemilihan sebagai wujud terlaksananya suatu kesepakatan yang bernilai demokratis, salah satunya pemerintahan atau organisasi kemasyarakatan tempat tinggal kita. Keberadaan lembaga masyarakat dilingkungan tempat tinggal bertujuan untuk mempercepat masyarakat yang tertib, aman dan sejahtera. Selain itu juga bertujuan untuk meningkatkan potensi masyarakat agar lebih inovatif dan kreatif sehingga dapat berkarya sesuai dengan kemampuan yang dimiliki oleh masyarakat seperti RW (Rukun warga). Rukun Warga (RW) adalah sebuah divisi dari wilayah di Indonesia di bawah Dusun atau Lingkungan. Rukun Warga tidak termasuk Pembagian administrasi pemerintahan, dan pembentukan melewati warga pertemuan. Hal ini dibentuk untuk sosial layanan yang diperbaiki oleh Desa atau Kelurahan. RW terdiri dari beberapa RT [2].

Penelitian ini di lakukan untuk mengklasifikasikan kriteria pemilihan calon

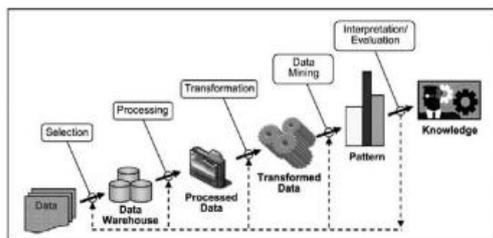
ketua RW (Rukun Warga) pada sebuah daerah dikota batam, permasalahan yang terjadi selama ini adalah proses pemilihan atau penetapan calon ketua RW belum memiliki kriteria yang harus dijadikan landasan dalam penetapan calon, begitu juga pada saat pemilihan berlangsung para pemilih tidak mengetahui dengan jelas siapa yang dipilih atau mungkin melakukan pemihan atas dasar kenal saja dengan calon, sehingga tujuan sebenarnya diadakan pemungutan suara untuk ketua RW tidak tercapai. Dalam pemilihan partisipasi warga sangat diperlukan akan tetapi tidak semua warga ikut berpartisipasi pada perhelatan ini tentunya hal ini menimbulkan permasalahan setelah dilakukan pemungutan suara, banyak dari warga yang tidak kenal siapa RW yang terpilih sampai kepada tidak mengetahui latar belakang dari pemimpin diperumahannya, selain itu para calon ketua RW tidak melakukan sosialisasi secara menyeluruh tentang visi-misi kedepan dalam pembangunan masyarakat, situasi ini menimbulkan keluhan dari masyarakat seperti pada saat melakukan pemilihan mereka tidak memilih berdasarkan kriteria yang baik untuk dijadikan wakil pimpinan tempat mereka tinggal.

*Datamining* akan menggali informasi yang berharga selama ini untuk dijadikan solusi permasalahan kemudian dengan teknik klasifikasi algoritma C4.5 akan menjawab persoalan pada kasus penelitian ini karena Algoritma C4.5 atau pohon keputusan mirip sebuah pohon dimana terdapat *node* internal (bukan daun) yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji, dan setiap daun menggambarkan kelas. Pohon keputusan dengan mudah dapat dikonversi ke aturan klasifikasi [3]. Hasil klasifikasi ini akan menentukan kriterria dari calon ketua RW yang layak dipilih oleh warga.

1.1 *Knowledge Discovery in Databases* (KDD)  
Menambang data dari sekumpulan data untuk

menemukan informasi baru yang berguna untuk membuat keputusan bagi perusahaan disebut dengan *datamining*. *Datamining* adalah bagian terintegrasi dari *Knowledge Discovery Database* (KDD), di mana KDD menggambarkan tahapan atau langkah pemrosesan data sebagai penambangan [4].

Untuk lebih jelasnya tahapan dari KDD ini dapat dilihat pada gambar dibawah ini.



Gambar 1. Proses Knowledge Discovery In Database

### 1.2 Datamining

Pada penelitian sebelumnya untuk mengklasifikasi kan resiko kredit mengatakan dalam referensinya bahwa *datamining* Menurut *Gartner Group*, data mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika. *Datamining* merupakan gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar [5].

Penelitian lain juga mengatakan *datamining* adalah salah satu cabang ilmu komputer yang banyak menarik perhatian masyarakat. Data mining adalah tumpukan data yang sudah tersimpan selama bertahun-tahun yang dimasukkan kedalam *database* tetapi tidak digunakan kembali atau disebut “data sampah”. *Datamining* digunakan untuk menggali dan mendapatkan informasi dari data dengan jumlah besar [6].

### 1.3 Klasifikasi

Klasifikasi dapat digambarkan sebagai berikut. Data input, disebut juga training set, terdiri atas banyak contoh (*record*), yang masing-masing memiliki beberapa atribut. Selanjutnya, tiap

contoh diberi sebuah label *class* khusus teori ini dilakukan pada penelitian sebelumnya untuk memprediksi tingkat kelulusan siswa [7].

Lebih lanjut, *input* didefinisikan sebagai sekumpulan *record* (*training set*), dan setiap *record* terdiri atas sekumpulan atribut, salah satu atribut adalah klas. Adapun model klasifikasi digunakan untuk antara lain :

1. Pemodelan Deskriptif sebagai perangkat penggambaran untuk membedakan objek-objek dari klas berbeda.
2. Pemodelan Prediktif digunakan untuk memprediksi label klas untuk *record* yang tidak diketahui atau tidak dikenal [6].

### 1.4 Pohon Keputusan

Manusia selalu dihadapkan dengan berbagai macam masalah dari berbagai bidang kehidupan. Masalah ini juga memiliki variasi tingkat kesulitannya. Untuk menghadapi masalah ini manusia mulai mengembangkan sebuah sistem untuk membantu mereka menyelesaikan masalah-masalah ini, salah satu sistem tersebut adalah pohon keputusan. Pohon keputusan adalah metode klasifikasi dan prediksi yang sudah terbukti *powerfull* dan sangat terkenal. Metode ini berfungsi untuk mengubah fakta menjadi pohon keputusan yang merepresentasikan aturan yang dapat mudah dimengerti dengan bahasa alami. Proses dari pohon keputusan ini dimulai dari node akar hingga *node* daun yang dilakukan secara rekursif dimana setiap percabangan menyatakan kondisi dan setiap ujung pohon akan menyatakan keputusan, hal ini diungkap pada penelitian sebelumnya untuk memprediksi calon pegawai baru pada sebuah perusahaan [8].

Penelitian terdahulu dalam referensinya mengatakan *Decision tree* merupakan metode yang mengubah data menjadi pohon keputusan (*decision tree*) dan aturan-aturan keputusan [9]. yang dijelaskan pada gambar berikut :



Gambar 2. Konsep Decision Tree

Pada *decision tree* terdapat 3 jenis *node*, yaitu:

- a. *Root node*: merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.

- b. *Internal node*: merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
- c. *Leaf node* atau *terminal node*: merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output*.

### 1.5 Algoritma C4.5

Algoritma C4.5 adalah salah satu metode untuk membuat decision tree berdasarkan training data yang telah disediakan. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi missing value, bisa mengatasi data kontinu, dan pruning. Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan yang merupakan pengembangan dari algoritma ID3, algoritma tersebut digunakan untuk membentuk pohon keputusan. Pohon keputusan dianggap sebagai salah satu pendekatan yang paling populer, dalam klasifikasi pohon keputusan terdiri dari sebuah node yang membentuk akar, node akar tidak memiliki inputan [10].

Sebuah objek yang diklasifikasikan dalam pohon harus dipes nilai *Entropy* -nya. *Entropy* adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari *impurity* dan *homogeneity* dari kumpulan data. Dari nilai *Entropy* tersebut kemudian dihitung nilai *information gain* (IG) masing-masing atribut. *Entropy* (*S*) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel *S*. *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai *Entropy* maka akan semakin *Entropy* digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidakkaslarian *S*.

Untuk memilih atribut akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti yang tertera dalam persamaan berikut.

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

Di mana :

- S* : himpunan kasus
- A* : atribut
- N* : jumlah partisi atribut *A*

*|S<sub>i</sub>|* : jumlah kasus pada partisi ke-*i*

*|S|* : jumlah kasus dalam *S*

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 2 berikut.

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

Di mana :

*S* : himpunan kasus

*A* : fitur

*N* : jumlah partisi *S*

*p<sub>i</sub>* : proporsi dari *S<sub>i</sub>* terhadap *S*

### 1.6 Weka (Waikato Environment for Knowledge Analysis)

*Waikato Environment for Knowledge Analysis*, yang dibuat di Universitas *Waikato, New Zealand* untuk penelitian, pendidikan dan berbagai aplikasi. *Weka* mampu menyelesaikan masalah-masalah *data mining* di dunia nyata, khususnya klasifikasi yang mendasari pendekatan-pendekatan machine learning. Perangkat lunak ini ditulis dalam hirarki class Java dengan metode berorientasi objek dan dapat berjalan hampir di semua *platform*. *Weka* mudah digunakan dan diterapkan pada beberapa tingkatan yang berbeda. Tersedia implementasi algoritma-algoritma pembelajaran *state-of-the-art* yang dapat diterapkan pada dataset dari *command line*. *WEKA* mengandung *tools* untuk *pre-processing* data, klasifikasi, regresi, clustering, aturan asosiasi, dan visualisasi. *User* dapat melakukan *preprocess* pada data, memasukkannya dalam sebuah skema pembelajaran, dan menganalisa classifier yang dihasilkan dan performansinya semua itu tanpa menulis kode program sama sekali. Ada banyak metodologi *Data mining* yang dapat dilakukan pada *WEKA*, salah satu yang populer adalah pohon keputusan (*decision tree*).

## 2. METODE PENELITIAN

### 2.1 Metode Pengumpulan Data

Teknik pengumpulan data dalam analisis klasifikasi pemilihan calon ketua RW adalah sebagai berikut :

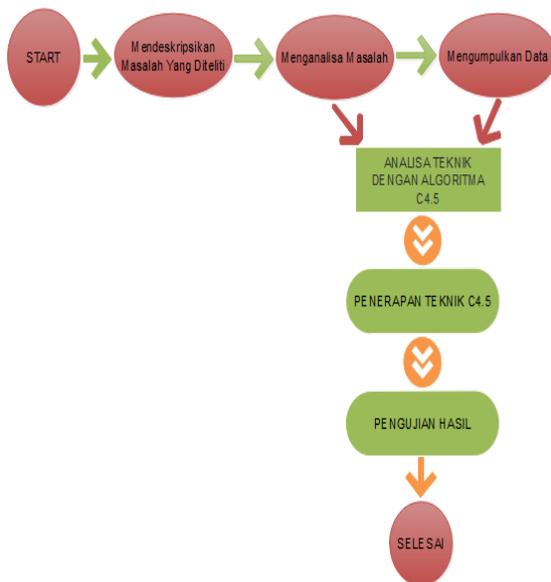
#### 1. Observasi

Pengumpulan data yang diperoleh dari pengamatan langsung dilapangan tentang permasalahan yang dihadapi dalam penentuan calon ketua RW pada objek penelitian dalam hal ini dilaksanakan pada perumahan cipta asri group dan berbagai aspek yang berhubungan dengan

- penelitian.
2. Studi Literatur  
Mempelajari konsep-konsep tentang datamining dengan algoritma C4.5. Sumber literatur dapat berupa buku, jurnal, *paper* dan *website*.
  3. Wawancara  
Untuk mendapatkan data yang akurat diperlukan sebuah tanya jawab dengan daftar pertanyaan yang telah disiapkan dengan pihak-pihak yang terkait dengan objek penelitian. dilakukan untuk mendapatkan data yang tidak diperoleh pada saat observasi lapangan.

### 2.2 Desain Analisis Penelitian

Desain penelitian ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun desain penelitian ini dapat dilihat pada gambar berikut.



Gambar 3. Desain Penelitian

Berdasarkan desain penelitian pada gambar diatas maka masing-masing langkahnya dapat diuraikan seperti berikut ini :

1. Mendeskripsikan Masalah  
Mendeskripsikan masalah yang akan teliti perlu ditentukan terlebih dahulu. Mendeskripsikan masalah dalam penelitian dengan menentukan dan mendefinisikan batasan masalah yang akan diteliti, sehingga

membantu dalam mendapatkan suatu solusi yang terbaik dari masalah tersebut. Jadi, langkah pertama ini adalah langkah awal yang terpenting dalam penelitian ini.

2. Analisa Masalah  
Langkah analisis masalah merupakan langkah untuk dapat memahami masalah yang telah ditentukan ruang lingkup atau batasannya. Dengan menganalisis masalah yang telah ditentukan tersebut, maka diharapkan masalah dapat dipahami dengan baik.
3. Mempelajari Literatur  
Untuk mencapai tujuan, maka dipelajari beberapa literatur-literatur yang diperkirakan dapat digunakan. Kemudian literatur-literatur yang dipelajari tersebut diseleksi untuk dapat ditentukan literatur-literatur mana yang akan digunakan dalam penelitian.
4. Mengumpulkan Data  
Dalam pengumpulan data dilakukan observasi yaitu pengamatan secara langsung pada objek penelitian ini sehingga permasalahan yang ada dapat diketahui secara jelas. Kemudian dilakukan *interview* yang bertujuan untuk mendapatkan informasi atau data yang dibutuhkan. Selain itu juga dilakukan studi kepustakaan yaitu dengan membaca buku-buku yang menunjang dalam melakukan analisis terhadap data dan informasi yang didapat. Analisa teknik pengolahan data menggunakan algoritma C4.5. Data yang diperoleh dari tempat penelitian selanjutnya dilakukan analisa dan pengolahan menggunakan algoritma C4.5.
5. Analisa Teknik Pengolahan Data dengan Algoritma C4.5  
Pada tahap ini akan dilakukan proses perancangan dari model sistem dengan algoritma C4.5 sehingga membentuk pohon keputusan (*decision tree*) dan menghasilkan suatu *rule* klasifikasi pemilihan calon Ketua RW.
6. Implementasi Algoritma C4.5  
Adapun langkah-langkah dalam tahapan ini adalah :
  - a. Menentukan atribut sebagai akar dan menghitung nilai informasi *gain* atribut.
  - b. Menyusun *Tree* awal
  - c. Mengubah *Tree* menjadi *rule*
7. Pengujian Hasil  
Pada tahap ini, penulis melakukan pengujian dan hasil perancangan sistem menggunakan

software Data Mining open source WEKA. Sistem diuji dengan prosedur-prosedur untuk melakukan eksplorasi dan permodelan dari data-data yang ada sehingga mendapatkan suatu hubungan tersembunyi dari data tersebut.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

Data yang telah dikumpulkan Data yang telah dikumpulkan dan ditransformasikan kedalam bentuk tabel yang akan di mining dengan variabel ukuran berjumlah 6 variabel yaitu : Pendidikan, Peminatan Terhadap Pencalonan, Memiliki Solidaritas Terhadap tetangga/Warga, Memiliki Visi dan Misi Sebagai Calon, Memiliki Jiwa Kepemimpinan, Tanggung Jawab dengan dua tingkat klasifikasi keputusan yaitu Dipilih dan Tidak Dipilih. Kemudian data tersebut akan dimasukan kedalam tabel yang di sebut dengan tabel Pra-proses untuk lebih jelasnya dapat dilihat pada tabel dibawah ini.

Tabel 1. Data Pra-Proses

Pendidikan	Peminatan terhadap Pencalonan	Solidaritas Terhadap Warga/Tetangga	Memiliki Visi Misi Sebagai Calon	Memiliki Jiwa Kepemimpinan	Tanggung Jawab	Klasifikasi Calon
P	P	P	P	P	P	Dipilih
P	Tperli	P	P	TP	Tperli	Dipilih
P	Tperli	TP	P	TP	Tperli	Tidak Dipilih
P	P	P	P	TP	P	Dipilih
P	P	P	P	P	P	Tidak Dipilih
P	Tperli	TP	Tperli	P	Tperli	Tidak Dipilih
P	P	P	P	TP	P	Tidak Dipilih
P	P	TP	P	P	P	Dipilih
P	Tperli	TP	P	P	Tperli	Tidak Dipilih
TP	P	P	P	P	P	Dipilih
P	P	TP	P	P	P	Tidak Dipilih
TP	P	P	P	TP	P	Dipilih
P	Tperli	TP	Tperli	P	Tperli	Tidak Dipilih
TP	P	P	P	P	P	Dipilih
P	Tperli	TP	Tperli	P	Tperli	Tidak Dipilih
TP	P	P	P	P	P	Dipilih
TP	Tperli	TP	P	TP	Tperli	Tidak Dipilih
TP	P	P	P	P	P	Dipilih
P	P	P	P	KP	P	Dipilih
P	Tperli	TP	P	P	Tperli	Dipilih
P	P	TP	P	P	P	Tidak Dipilih
TP	P	TP	P	KP	P	Tidak Dipilih
TP	Tperli	TP	P	P	Tperli	Dipilih
P	P	TP	P	KP	P	Tidak Dipilih
P	P	P	P	P	P	Dipilih
TP	Tperli	TP	Tperli	P	Tperli	Dipilih
P	P	P	P	TP	P	Dipilih
TP	P	P	P	KP	P	Tidak Dipilih
P	P	P	P	P	P	Dipilih
P	Tperli	TP	P	P	Tperli	Dipilih
TP	P	P	P	P	P	Tidak Dipilih
P	P	P	P	P	P	Dipilih
P	P	P	P	TP	P	Dipilih
TP	P	P	P	P	P	Dipilih
P	P	P	P	P	P	Dipilih
TP	P	TP	P	P	P	Tidak Dipilih
P	P	P	P	P	P	Dipilih
TP	Tperli	TP	P	P	Tperli	Dipilih
TP	P	P	P	TP	P	Dipilih
TP	P	P	P	P	P	Dipilih
TP	P	P	P	TP	P	Dipilih
TP	Tperli	TP	P	P	Tperli	Dipilih
TP	Tperli	TP	P	TP	Tperli	Tidak Dipilih
P	Tperli	P	P	P	Tperli	Dipilih

Dari tabel pra-proses diatas maka akan dicari klasifikasi calon ketua RW dengan mencari nilai Entropy dan Gain masing-masing variabel dengan menggunakan rumus 1 dan 2 pada penjelasan tinjauan pustaka diatas.

1. Nilai Entropy tiap-tiap atribut :

$$Entropy(\text{Total}) = \left(-\frac{34}{50} * \log_2\left(\frac{34}{50}\right)\right) + \left(-\frac{16}{50} * \log_2\left(\frac{16}{50}\right)\right) = 0.904381$$

a. Atribut Pendidikan

$$Entropy(\text{Penting}) = \left(-\frac{17}{27} * \log_2\left(\frac{17}{27}\right)\right) + \left(-\frac{10}{27} * \log_2\left(\frac{10}{27}\right)\right) = 0.95095$$

$$Entropy(\text{T. Penting}) = \left(-\frac{17}{23} * \log_2\left(\frac{17}{23}\right)\right) + \left(-\frac{6}{23} * \log_2\left(\frac{6}{23}\right)\right) = 0.82805$$

b. Atribut Peminatan Terhadap Pencalonan

$$Entropy(\text{Perlu}) = \left(-\frac{25}{34} * \log_2\left(\frac{25}{34}\right)\right) + \left(-\frac{9}{34} * \log_2\left(\frac{9}{34}\right)\right) = 0.83376$$

$$Entropy(\text{T. Perlu}) = \left(-\frac{9}{16} * \log_2\left(\frac{9}{16}\right)\right) + \left(-\frac{7}{16} * \log_2\left(\frac{7}{16}\right)\right) = 0.98869$$

c. Atribut Memiliki Visi dan Misi Sebagai Calon

$$Entropy(\text{Perlu}) = \left(-\frac{23}{27} * \log_2\left(\frac{23}{27}\right)\right) + \left(-\frac{4}{27} * \log_2\left(\frac{4}{27}\right)\right) = 0.60518$$

$$Entropy(\text{T. Perlu}) = \left(-\frac{11}{23} * \log_2\left(\frac{11}{23}\right)\right) + \left(-\frac{12}{23} * \log_2\left(\frac{12}{23}\right)\right) = 0.99863$$

d. Atribut Solidaritas Terhadap Warga

$$Entropy(\text{Penting}) = \left(-\frac{33}{46} * \log_2\left(\frac{33}{46}\right)\right) + \left(-\frac{13}{46} * \log_2\left(\frac{13}{46}\right)\right) = 0.85898$$

$$Entropy(\text{T. Penting}) = \left(-\frac{1}{4} * \log_2\left(\frac{1}{4}\right)\right) + \left(-\frac{3}{4} * \log_2\left(\frac{3}{4}\right)\right) = 0.81127$$

e. Atribut Memiliki Jiwa Kepemimpinan

$$Entropy(Penting) = \left(-\frac{25}{34} * \log_2\left(\frac{25}{34}\right)\right) +$$

$$\left(-\frac{9}{34} * \log_2\left(\frac{9}{34}\right)\right) = 0.83376$$

$$Entropy(K. Penting) = \left(-\frac{1}{4} * \right.$$

$$\log_2\left(\frac{1}{4}\right) + \left(-\frac{3}{4} * \right.$$

$$\log_2\left(\frac{3}{4}\right) = 0.81127$$

$$Entropy(T. Penting) = \left(-\frac{8}{12} * \right.$$

$$\log_2\left(\frac{8}{12}\right) + \left(-\frac{4}{12} * \right.$$

$$\log_2\left(\frac{4}{12}\right) = 0.91829$$

f. Atribut Tanggung Jawab

$$Entropy(Perlu) = \left(-\frac{25}{34} * \log_2\left(\frac{25}{34}\right)\right) +$$

$$\left(-\frac{9}{34} * \log_2\left(\frac{9}{34}\right)\right) = 0.83376$$

$$Entropy(T. Perlu) = \left(-\frac{9}{16} * \log_2\left(\frac{9}{16}\right)\right) +$$

$$\left(-\frac{7}{16} * \log_2\left(\frac{7}{16}\right)\right) = 0.98869$$

Dari perhitungan nilai *entropy* dan *gain* diatas dimasukkan kedalam tabel sehingga dibentuk tabel nilai *node* 1.1 sebagai berikut.

Tabel 2. Node 1.1

Klasifikasi Penilaian	Jumlah Kasus	Tidak Dipilih		Entropy	Gain
		S1	S2		
Total	50	34	16	0.904381458	
Pendidikan					
Penting	27	17	10	0.950956048	0.009959558
Tidak Penting	23	17	6	0.828055725	
Peminatan terhadap Pencalonan					
Perlu	34	25	9	0.833764907	0.02103751
Tidak Perlu	16	9	7	0.988699408	
Memiliki Visi Misi Sebagai Calon					
Perlu	27	23	4	0.605186576	0.118208163
Tidak Perlu	23	11	12	0.998635964	
Solidaritas Terhadap Warga/Tetangga					
Penting	46	33	13	0.858981037	0.049216654
Tidak Penting	4	1	3	0.811278124	
Memiliki Jiwa Kepemimpinan					
Pernting	34	25	9	0.833764907	0.052128071
Kurang Penting	4	1	3	0.811278124	
Tidak Penting	12	8	4	0.918295834	
Tanggung Jawab					
Perlu	34	25	9	0.833764907	0.02103751
Tidak Perlu	16	9	7	0.988699408	

Dari hasil perhitungan *node* 1.1 dapat dilihat nilai *gain* tertinggi berada pada variabel memiliki visi dan misi sebagai calon dimana nilai *gain* nya 0.1182082, sehingga hasil ini menjadikan variabel tersebut menjadi akar pertama dari pohon keputusan, kemudian pencarian akan dilanjutkan lagi dengan perhitungan *node* 1.2 untuk menentukan akar selanjutnya, hasil perhitungan dapat dilihat pada tabel dibawah ini.

Tabel 3. Node 1.2

Klasifikasi Penilaian	Jumlah Kasus	Tidak Dipilih		Entropy	Gain
		S1	S2		
Total	46	33	13	0.858981037	
Pendidikan					
Penting	24	17	7	0.870864469	0.000318692
Tidak Penting	22	16	6	0.845350937	
Peminatan terhadap Pencalonan					
Perlu	34	25	9	0.833764907	0.003164584
Tidak Perlu	12	8	4	0.918295834	
Solidaritas Terhadap Warga/Tetangga					
Penting	27	23	4	0.605186577	0.091545073
Tidak Penting	19	10	9	0.998000884	
Memiliki Jiwa Kepemimpinan					
Pernting	30	24	6	0.721928095	0.078057008
Kurang Penting	4	1	3	0.811278124	
Tidak Penting	12	8	4	0.918295834	
Tanggung Jawab					
Perlu	34	25	9	0.833764907	0.003164584
Tidak Perlu	12	8	4	0.918295834	

Berdasarkan hasil perhitungan *Node* 1.2 pada tabel diatas dapat dilihat hasil perhitungan *Gain* tertinggi ada pada variabel solidaritas terhadap tetangga/ warga dengan nilai *gain* 0.0915451, dengan demikian akar selanjutnya dari pohon keputusan yang terbentuk adalah solidaritas sebagai calon terhadap warga dimana klasifikasi bernilai pernting sudah lebih banyak dari tidak penting sedangkan klasifikasi tidak penting belum dapat diketahui maka akan dilanjutkan dengan perhitungan *node* 1.3 pada tabel dibawah ini.

Tabel 4. Node 1.3

Klasifikasi Penilaian	Jumlah Kasus	Dipilih S1	Tidak Dipilih S2	Entropy	Gain
Total	19	10	9	0.998000884	
Pendidikan					
Penting	11	6	5	0.994030211	0.001457077
Tidak Penting	8	4	4	1	
Peminatan terhadap Pencalonan					
Perlu	10	5	5	1	0.002228013
Tidak Perlu	9	5	4	0.99107606	
Memiliki Jiwa Kepemimpinan					
Penting	13	9	4	0.89049164	0.217921735
Kurang Penting	2	0	2	0	
Tidak Penting	4	1	3	0.811278124	
Tanggung Jawab					
Perlu	10	5	5	1	0.002228013
Tidak Perlu	9	5	4	0.99107606	

Pada tabel *node* 1.3 ini merupakan perhitungan nilai *entropy* dan *gain* terakhir karena telah ditemukan klasifikasi akhir dari pohon keputusan yaitu memiliki jiwa kepemimpinan dengan nilai gain 0.217921735.

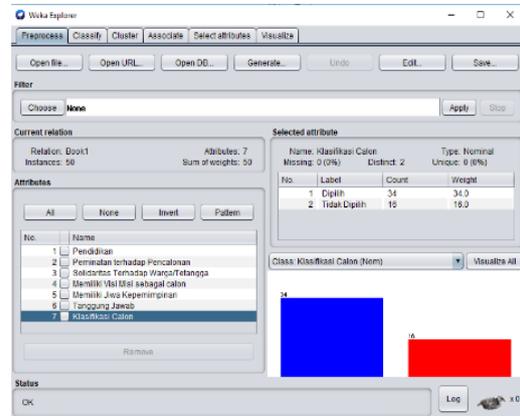
Berdasarkan analisa *datamining* yang telah dilakukan dengan perhitungan algoritma C4.5 maka didapat hasil bahwasanya klasifikasi untuk kriteria sebagai calon ketua RW yang diinginkan oleh masyarakat sebagai calon yang nantinya mereka pilih adalah :

1. Memiliki Visi dan Misi sebagai calon  
Setiap calon yang akan mengikuti pemilihan harus memiliki atau penting memiliki visi dan misi yang akan membangun masyarakat kedepannya.
2. Solidaritas Terhadap Warga/Tetangga  
Para calon yang akan mengajukan diri dan akan dipilih haruslah orang yang punya solidaritas yang tinggi karena ini merupakan kepala masyarakat tentunya harus memiliki contoh yang baik bagi warganya.
3. Memiliki Jiwa Kepemimpinan

### 3.2 Pembahasan

Pada pembahasan proses *mining* yang telah dilakukan pada hasil penelitian ini akan dilakukan pengujian dengan aplikasi *weka* untuk melihat pohon keputusan yang dibentuk serta menguji kebenaran hasil dari perhitungan manual.

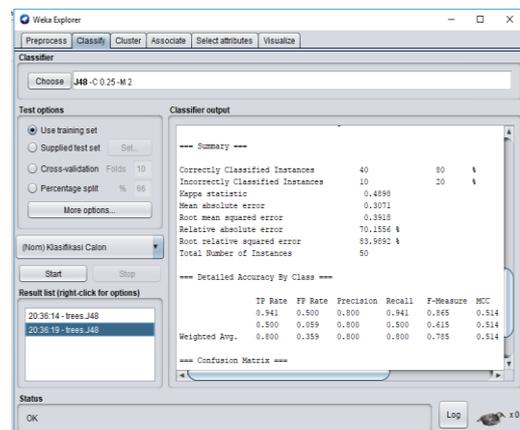
Proses *mining* dengan *weka* ini perlu dilakukan transformasi karena *weka* hanya membaca file yang memiliki format *Csv*. kemudian akan di *explorer* ke dalam aplikasi untuk pemilihan variabel yang akan di *mining* seperti gambar dibawah ini.



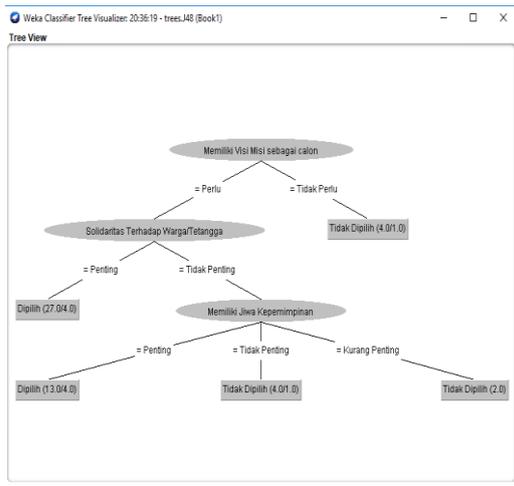
Gambar 4. Pemilihan Variabel Format Data Csv

Dari gambar diatas dapat dilihat klasifikasi dari data yang diperoleh dan dilakukan praproses dalam bentuk format data yang dapat dibaca oleh *Weka* yaitu terdapat klasifikasi 34 dipilih dan 16 tidak dipilih.

Kemudian dilakukan proses klasifikasi terhadap data yang dimasukkan dengan menggunakan pengklasifikasi (*classifier*) berupa *trees classifier* berbasis algoritma J4.8, yang merupakan implementasi dari algoritma C4.5 pada WEKA. Untuk lebih jelasnya dapat dilihat pada gambar dibawah ini.



Gambar 5. Using Training Set J4.8



Gambar 6. Pohon Keputusan (*Decision Tree*)

Dari hasil pohon keputusan yang terbentuk oleh aplikasi weka maka dapat dilihat yang menjadi akar utamanya yaitu memiliki visi dan misi sebagai calon kemudian diikuti solidaritas dan terakhir memiliki jiwa kepemimpinan hal ini menunjukkan hasil perhitungan manual yang dilakukan sama dengan aplikasi *weka*.

#### 4. KESIMPULAN

##### 4.1 Simpulan

Analisa datamining dengan algoritma C4.5 dapat membantu mengklasifikasikan kriteria Calon Ketua RW yang akan dipilih oleh masyarakat dimana pada hasil pengolahan didapat kriteria dengan nilai *gain* tertinggi yang menjadi akar pohon keputusan yaitu Calon yang memenuhi kriteria klasifikasi dalam memiliki visi dan misi sebagai calon karena merupakan hal yang penting bagi masyarakat mengenal calonnya melalui visi dan misi yang mereka ajukan untuk menjadi ketua, kemudian memiliki solidaritas yang tinggi terhadap warga sekitar ini merupakan sesuatu yang perlu dimiliki oleh calon karena pemimpin masyarakat walaupun di tingkat organisasi paling bawah dalam pemerintahan harus memiliki sikap solidaritas terhadap sesama tanpa membedakan suku, agama dan golongan terakhir hasil dari analisa menunjukkan seorang calon yang akan dipilih harus masuk dalam kriteria klasifikasi memiliki jiwa kepemimpinan yang baik karena warga butuh seorang calon yang nanti memimpin mereka adalah orang yang bisa mengayomi masyarakat dalam hal kegiatan-kegiatan berwarga.

Pengujian dengan software weka juga

menjelaskan bahwa hasil perhitungan manual dengan *software* mempunyai kesamaan.

#### 4.2 SARAN

Penambahan variabel diperlukan untuk penelitian lebih lanjut baik di lapangan maupun berdasar data yang ada guna menghasilkan *rule* yang tepat dan untuk mendapatkan hasil prediksi yang lebih baik perlu menggunakan dan menggabungkan metode klasifikasi yang lain

#### DAFTAR PUSTAKA

- [1] B. Hardiyanto, Suharso, "Pemilihan Umum Kepala Daerah Periode 2015/2020 (Studi Politik Hukum Calon Tunggal)," vol. 12, no. 1, pp. 204–221, 2016.
- [2] Yanuardi, "Pelaksanaan Tugas Rukun Tetangga dan Rukun Warga (RT/RW) Kelurahan Delima Pekanbaru," vol. 2, no. c, pp. 1–4, 2015.
- [3] H. A. Evicienna, "Algoritma c4.5 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI Jakarta," Techno Nusa Mandiri, vol. IX, no. 1, pp. 48–56, 2013.
- [4] E. N. Wahyudi, "Teknik Klasifikasi untuk Melihat Kecenderungan Calon Mahasiswa Baru dalam Memilih Jenjang Pendidikan Program Studi di Perguruan Tinggi," vol. 18, no. 1, pp. 55–64, 2013.
- [5] Y. Mardi, "Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database ( KDD ) . Jurnal Edik Informatika," J. Edik Inform., pp. 213–219, 2017.
- [6] Eviciana and H. Amalia, "Algoritma C4.5 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI Jakarta," Techno Nusa Mandiri, vol. IX, no. 1, pp. 48–56, 2013.
- [7] G. Indrawan, "Penerapan Metode Decision Tree (Datamining) Untuk Memprediksi Tingkat Kelulusan Siswa SMPN 1," pp. 35–44,

2016.

- [8] F. F. Harryanto and S. Hansun, "Penerapan Algoritma C4 . 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE," *Jatani*, vol. 3, no. 2, pp. 95–103, 2017.
- [9] T. Thi et al., "Implementasi Iterative Dichotomiser 3 Pada Data Kelulusan Mahasiswa S1 Di Universitas Sebelas Maret," vol. 4, no. 2, pp. 84–91, 2015.
- [10] P. A. Sularno, "Penerapan Algoritma C4.5 Untuk Klasifikasi Tingkat Keganasan Hama Pada Tanaman Padi," *J. Sains Dan Inform.*, vol. 2, pp. 54–60, 2016